Algorithms Seminar 2000–2001,
F. Chyzak (ed.), INRIA, (2002), pp. 121–126.

Available online at the URL
http://algo.inria.fr/seminars/.

Cover Time and Favourite Points for Planar Random Walks

Amir Dembo
Mathematics and Statistics Department, Stanford University (USA)

June 18, 2001

Summary by Christine Fricker and Pierre Nicodème

Abstract

In this talk, Amir Dembo considers random walks on $\mathbb{Z}^2$ and presents a proof of the Erdős–Taylor conjecture related to frequently covered points. The Kesten–Révész conjecture on the covering time of the two-dimensional torus $\mathbb{Z}_n^2 = \mathbb{Z}^2/n\mathbb{Z}^2$ is also solved. These results are a common work of Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni.

1. Introduction

Let (X_n) be a simple random walk on $\mathbb{Z}^2$ and $T_n(x) = \sum_{j=1}^n 1_{\{X_j = x\}}$ be the number of visits to x before time n. Let $T_n^\star = \max_{x\in\mathbb{Z}^2} T_n(x)$ be the number of visits to the most visited point. The *Erdős–Taylor conjecture* asserts that

$$\lim_{n\to\infty} \frac{T_n^\star}{(\log n)^2} = \frac{1}{\pi}, \quad \text{almost surely.} \tag{1}$$

Erdős and Taylor [7] proved the upper bound $1/\pi$ and a lower bound $1/(4\pi)$. The main result of the talk is that the Erdős–Taylor conjecture is true.

Let $(\widetilde{X}_j)$ be a simple random walk on the two-dimensional torus $\mathbb{Z}_n^2 = \mathbb{Z}^2/n\mathbb{Z}^2$. Consider $\mathcal{T}(x) = \min\{\, j \geq 0 \mid \widetilde{X}_j = x \,\}$, the time to attain the point x for the first time and

$$\mathcal{T}_n = \max_{x\in\mathbb{Z}_n^2} \mathcal{T}(x),$$

the covering time of the torus. The *Aldous-Lawler conjecture* asserts than

$$\lim_{n\to\infty} \frac{\mathcal{T}_n}{(n\log n)^2} = \frac{4}{\pi}, \quad \text{in probability.} \tag{2}$$

Kesten, Révész, Lawler, and Aldous proved an upper bound $4/\pi$ (see [1, Corollary 25, Chapter 7]) and a lower bound $2/\pi$. A related question is the Kesten–Révész conjecture for the simple random walk on $\mathbb{Z}^2$ (see [4]).

The proofs for the upper bounds rely on the second moment method, the approximation of random walks by Brownian motions, and an underlying tree structure for the occupation of small disks by a Brownian motion. We give here a sketch of the proofs; see [4, 5] for complete proofs.

2. The Second Moment Method

Janson gives a short account of the second moment method in [2]. Basically, we consider a sequence of non-negative random variables X_n, and we want to estimate $\mathbf{P}(X_n > 0)$. The second moment method asserts that if

(3) $$\frac{\mathbf{Var}(X_n)}{(\mathbf{E}X_n)^2} \to 0, \qquad \text{or equivalently,} \qquad \frac{\mathbf{E}X_n^2}{(\mathbf{E}X_n)^2} \to 1 \qquad (\text{as} \quad n \to \infty),$$

then

(4) $$\mathbf{P}(X_n > 0) \to 1.$$

The method is frequently used in the context of random graphs; for example, this method proves the existence of a Hamilton cycle in random graphs satisfying suitable conditions.

The second moment method is a consequence of the Chebyshev inequality,

$$\mathbf{P}(|X| > t) \leq \frac{1}{t^2}\mathbf{E}(X^2).$$

As a consequence of the latter,

$$\mathbf{P}(X = 0) \leq \mathbf{P}\big(|X - \mu| \geq \mu\big) \leq \frac{\mathbf{Var}(X)}{\mu^2}, \qquad \text{for } \mu = \mathbf{E}X.$$

3. Proof of the Erdős–Taylor Conjecture

3.1. **Upper bound.** By definition, the truncated Green function $G_n(x,y)$ is the expectation of the number of passages at y in n steps, when starting from x.

We have

$$G_n(0,0) = \sum_{j=0}^{n} \mathbf{E}\left(1_{\{X_j=0\}}\right) = \sum_{j=0}^{n} \mathbf{P}(X_j = 0) \sim \frac{\log n}{\pi}.$$

(See Feller [8, p. 361].) Applying [3, Theorem 8.7.3] for the renewal sequence $u_n = \mathbf{P}(X_n = 0)$, we deduce that for large n, and fixed small $\delta > 0$,

$$\mathbf{P}\big(X_j \neq 0 \text{ for } j = 1, \ldots, n-1\big) \leq \frac{(1-\delta)\pi}{\log n}.$$

This implies by the strong Markov property that

(5) $$\mathbf{P}\left(T_n(0) \geq \alpha\pi(\log n)^2\right) \leq \left(1 - \frac{(1-\delta)\pi}{\log n}\right)^{\alpha(\log n)^2} \leq e^{-\alpha\pi(\log n)(1-\delta)} = n^{-(1-\delta)\alpha\pi}.$$

We now consider the disk of center zero and radius $n^{(1+\delta)/2}$. The probability that the random walk exits this disk before time n tends to zero as n tends to infinity, and the number of points of $\mathbb{Z}^2$ inside this disk is close to $\pi n^{(1+\delta)}$. From Equation (5), we then get

(6) $$\mathbf{P}_n^\alpha \leq \mathbf{P}\left(\max_{0 \leq i \leq n} |X_i| > n^{(1+\delta)/2}\right) + \pi n^{(1+\delta)} n^{-(1-\delta)\alpha\pi},$$

where $\mathbf{P}_n^\alpha = \mathbf{P}\big(T_n^\star \geq \alpha(\log n)^2\big)$. The first term of the right member of Equation (6) vanishes as n tends to infinity. Therefore, applying the Borel–Cantelli lemma to the subsequence $\mathbf{P}_{2^m}^\alpha$, for $\alpha > 1/\pi$, and using interpolation for all n, we have $\mathbf{P}\left(\overline{\lim}\, T_n^\star \geq \alpha\pi(\log n)^2)\right) \to 0$. This gives an upper bound $1/\pi$.

3.2. **Lower bound.** We can try to adapt the proof from the upper bound and use the second moment method. Let $D(x,r)$ be the disk of center x and radius r and

$$Z_n = \sum_{x\in D(0,\sqrt{n})} 1_{\{T_n(x)\geq\beta(\log n)^2\}}.$$

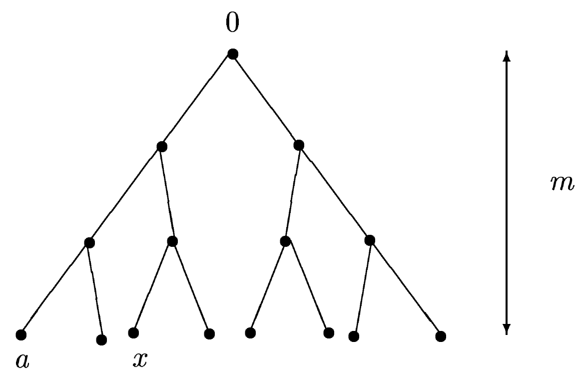
Adapting the proof of the upper bound (Equation (6)) gives $\mathbf{E}Z_n \approx n^{(1-\beta\pi)}$. Therefore,

$$\frac{\mathbf{E}Z_n^2}{(\mathbf{E}Z_n)^2} = \frac{1}{\mathbf{E}(Z_n)} + \frac{\Sigma_{x,y}}{\Sigma_{x,y}+\Sigma_x}, \qquad \text{where} \qquad \Sigma_x = \sum_{x\in D(0,\sqrt{n})} \Big(\mathbf{P}\big(T_n(x)\geq\beta(\log n)^2\big)\Big)^2$$

$$\text{and} \qquad \Sigma_{x,y} = \sum_{x\neq y\in D(0,\sqrt{n})} \mathbf{P}\big(T_n(x)\geq\beta(\log n)^2\big)\,\mathbf{P}\big(T_n(y)\geq\beta(\log n)^2\big).$$

A naive approach would say the following: the number of summand in $\Sigma_{x,y}$ is $O\big(n^{2(1-\beta\pi)}\big)$ while it is only $O\big(n^{(1-\beta\pi)}\big)$ in Σ_x. Therefore, for $\beta<1/\pi$, $\mathbf{E}Z_n^2/(\mathbf{E}Z_n)^2\to 1$ and $\mathbf{P}(T_n^\star\geq\frac{1}{\pi}(\log n)^2)=1$ almost surely. However, Erdős and Taylor [7] show that the correlation structure between points x such that $\mathbf{P}\big(T_n(x)\geq\beta(\log n)^2\big)$ is too strong to get this result. They obtain an upper limit $1/(4\pi)$. We move in the following section to a tree model to overcome this difficulty.

Modelling by a (toy) tree problem. We[1] consider a complete binary tree B_m of height m and a (nearest neighbor) random walk X starting from the left-most leaf a, with probability 1/3 of choosing any direction when being at an internal node. In this model, the starting point a and the root 0 respectively represent the origin (0,0) and the boundary of a "disk" of radius m on $\mathbb{Z}^2$. Let L_m be the set of leaves of B_m. We consider $T_m(x)$, the time spent at leaf x before hitting the root 0, and



$$T_m^\star = \max_{x\in L_m} T_m(x),$$

its maximum over all leaves.

Let us denote by $0,1,2,\ldots,a=m$ the nodes of the ray going from the root 0 to a and let $\mathbf{P}^y$ denote probability for walks starting from node y. We consider

$$H_y = H_y(u) = \sum_{u\geq 0} \mathbf{P}^y\,(X \text{ spends time } k \text{ at } a \text{ before hitting } 0)\ u^k.$$

For any node i of the ray $(0,a)$, and for any node y of the subtree rooted at the right child of i, the probability of k visits to a before hitting 0 of the walk starting from y is the same as if the walk starts from i; this implies $H_y=H_i$. This last result is true for all i from 1 to $m-1$.

We can therefore consider only the nodes of the ray $(0,a)$, which provide the set of equations

$$H_1 = \frac{H_2}{3}+\frac{H_1}{3}+\frac{1}{3}, \quad H_k = \frac{H_{k-1}}{3}+\frac{H_k}{3}+\frac{H_{k+1}}{3}\ (2\leq k\leq m-2), \quad H_{m-1} = \frac{H_{m-2}}{3}+\frac{(1+u)H_{m-1}}{3}.$$

[1]The elementary proof leading to Equation (7) was not presented by the speaker and is due to the authors of the summary.

Solving yields

(7) $$H_a(u) = H_m = \frac{1}{m} \times \frac{1}{1-\left(1-\frac{1}{m}\right)u}, \quad \text{and} \quad H_1(u) = \frac{m-1-(m-2)u}{m-(m-1)u}.$$

The random variable $T_m(a)$ therefore has a geometric distribution with mean $m-1$, which induces (for large m)

$$\mathbf{P}\big(T_m(a) > \alpha m^2\big) = \left(\left(1-\frac{1}{m}\right)^m\right)^{\alpha m} \simeq e^{-\alpha m} \quad \text{and} \quad \mathbf{P}(T_m^\star > \alpha m^2) \le e^{-\alpha m}2^m = e^{-(\alpha - \log 2)m}.$$

This implies the same upper bound as precedently (up to the change of model).

We now consider a variation of the second moment method. We fix some K large. We denote by x-ray the ray from the root 0 to a leaf x and $N_i(x)$ counts the number of excursions from level i to level $i+1$ on the ray x. We define the x-ray as α-successfull if

$$N_i(x) \simeq \alpha i^2, \quad \text{for} \quad i = 0, K, 2K, \dots, K\left\lfloor \frac{m}{K} \right\rfloor.$$

We have

$$\mathbf{P}\big(\, N_{i+K}(x) \simeq \alpha(i+K)^2 \;\big|\; N_i(x) \simeq \alpha i^2\big) \simeq e^{-\alpha K} \quad \Rightarrow \quad \mathbf{P}(x\text{-ray is } \alpha\text{-successfull}) \simeq e^{-\alpha m}.$$

We now have

$$\mathbf{P}(x\text{-ray and } y\text{-ray are } \alpha\text{-successfull}) \simeq e^{-2\alpha m} e^{\alpha r(x,y)},$$

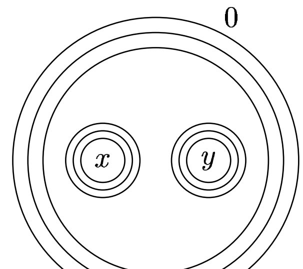
where $r(x,y)$ is the depth of the first common ancestor of x and y. This induces a reduction of variance. Considering now the random variable Z_m defined by

$$Z_m = \sum_{x \in L_m} 1_{\{x\text{-ray} \quad \alpha\text{-successfull}\}},$$

we have

$$\frac{\mathbf{E}Z_m^2}{(\mathbf{E}Z_m)^2} \simeq \sum_{s=1}^{m/K} e^{(\alpha - \log 2)Ks} \to 1 \quad \text{for} \quad \alpha < \log 2,$$

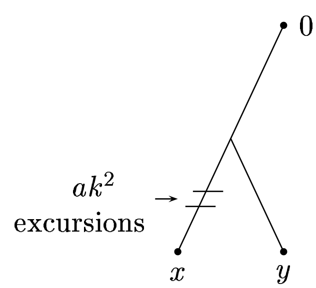
when first m and then K tend to infinity. There is no obvious way to adapt this result to the standard random walk, but it is possible to adapt it to the planar Brownian motion that we denote $w = (w_t)$.



Define θ as the first time where the Brownian motion w hits the circle of radius 1 and $\mu_\theta^w(A)$ as the occupation time of a subset A of the disc $D(0,1)$ until this time. We have

$$\theta = \min\big\{\, t \;\big|\; |w_t| = 1 \,\big\}$$

$$\text{and} \quad \mu_\theta^w(A) = \int_0^\theta 1_A(w_t)dt.$$



The Perkins–Taylor conjecture states for the Brownian motion that

(8) $$\limsup_{\epsilon \to 0} \sup_{|x|<1} \frac{\mu_\theta^w(D(x,\epsilon))}{\epsilon^2 (\log \epsilon)^2} = 2.$$

We shall in a first time sketch a proof of this conjecture and apply then the KMT approximation theorem of the Brownian motion by the standard random walk.

Sketch of proof for the Perkins–Taylor conjecture. In the following, let $\partial D(x,r)$ be the boundary of the disk $D(x,r)$.

The proof of the upper bound of the conjecture follows the same line as for the standard random walk. When considering the lower bound, the difficulty relies again in the correlation structure.

Let $\epsilon_k = e^{-k}$ and define a point x of $D(0,1)$ as k-successful if the number of excursions of the Brownian motion between $\partial D(x,\epsilon_k)$ and $\partial D(x,\epsilon_{k+1})$ is ak^2 for fixed a. We remark that if x is successful, the time spent at the ball $D(x,\epsilon_{k+1})$ is $ak^2\epsilon^2 \simeq a\epsilon^2(\log\epsilon)^2$, where $\epsilon = \epsilon_{k+1}$, with probability close to 1.

KMT approximation theorem. The Komlós–Major–Tusnády (KMT) approximation theorem [9] states that for each n it is possible to construct a random walk $\{X_k\}_{k=1}^n$ and the Brownian motion $\{w_t\}_{0\le t\le 1}$ on the same probability space so that for any $\delta > 0$ and any $\eta > 0$

$$\lim_{n\to\infty} \mathbf{P}\left(\max_{k=1,\dots,n}\left|w_{k/n} - \frac{\sqrt{2}}{\sqrt{n}}S_k\right| > \delta n^{\eta-1/2}\right) = 0. \tag{9}$$

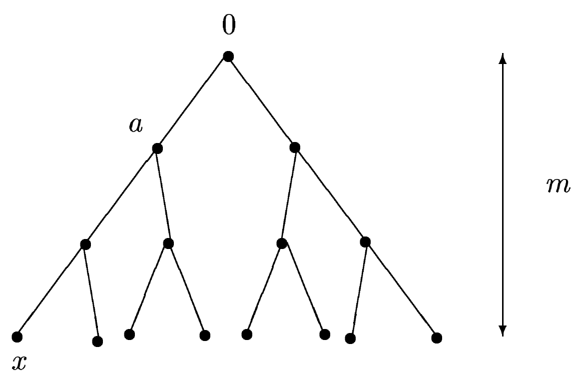
(The original one-dimension KMT approximation has been extended to the multivariate case by Einmahl [6]).

Note that the Brownian motion between two successful points x and y before reaching the boundary may again be modelized by a tree structure, and that the same technique as for trees works once more (with many technical issues).

Application of the KMT approximation theorem. The proof follows by considering the lattice points inside the circle $\{z : |\sqrt{2}z - y| < \sqrt{n}(1+2\delta)\epsilon_n\}$ whose number is less than

$$\frac{\pi}{2}n(1+2\delta)^3\epsilon_n^2.$$

4. Covering Time of the Torus



First, we once again consider the "toy" problem of the covering time of the binary tree B_m. Let $X = (X_n)$ be the first neighbor random walk starting from the left son a of the root, and consider hits to x, the leftmost leaf. $\mathbf{P}^x$ again refers to walks starting at point x.

4.1. **Upper bound.** From Section 3.2 we get

$$\mathbf{P}^a(X \text{ hits } x \text{ before } 0) = 1 - H_1(0) = \frac{1}{m}.$$

This implies that

$$\mathbf{P}^0(X \text{ does not cover } x \text{ during first } N \text{ visits to } 0) \simeq \left(1 - \frac{1}{2m}\right)^N.$$

Let Π^0 be the probability that the random walk starting at zero does not cover the binary tree B_m during N visits to 0. We have

$$\Pi^0 \le 2^m \left(1 - \frac{1}{2m}\right)^N \qquad \text{so that} \qquad \Pi^0 \to 0 \qquad \text{for} \quad N = 2(1+\delta)m^2 \log 2.$$

The time needed for N visits to the root is $2^{m+1}N$; this implies that

$$\mathbf{P}^0\big(X \text{ does not cover } B_m \text{ before time } 2(1+\delta)\log 2 \times m^2 2^{m+1}\big) \to 0.$$

4.2. **Lower bound.** A ray x is called *successful* if the number of excursions from level i to level $i+1$ in the ray is $a(m-i)^2$. Dembo *et al.* apply a second moment analysis to the successful rays to show that, with probability one, before $2(1-\delta)m^2 \log 2$ visits to the root, there are points which are not covered. Then, the time needed to visit the root that many times is about $2(1-\delta)m^2(\log 2)2^{m+1}$.

To solve the standard random walk problem on $\mathbb{Z}^2$, Dembo *et al.* first solve the equivalent problem for the Brownian motion on the torus $\mathbb{T}^2$, where $\mathbb{T}^2$ is identified with the set $(-1/2, 1/2]^2$.

Let $\mathcal{T}(x,\epsilon)$ denote the time needed by the Brownian motion to enter the ball $D(x,\epsilon)$,

$$\mathcal{T}(x,\epsilon) = \inf\big\{\, t > 0 \;\big|\; w_t \in D(x,\epsilon) \,\big\}, \quad \text{and} \quad C_\epsilon = \sup_{x \in \mathbb{T}^2} \mathcal{T}(x,\epsilon).$$

Therefore, C_ϵ is the minimum time needed for the Brownian motion W_t to come within ϵ of each point of $\mathbb{T}^2$. Equivalently, C_ϵ is the amount of time needed for the Wiener sausage of radius ϵ to completely cover $\mathbb{T}^2$. Dembo *et al.* [4] prove that

$$\lim_{\epsilon \to 0} \frac{C_\epsilon}{(\log \epsilon)^2} = \frac{2}{\pi}, \qquad \text{almost surely.}$$

Using the KMP strong approximation theorem again provides the result for the standard random walk on $\mathbb{T}^2$.

Bibliography

[1] Aldous (D.) and Fill (J.). – Reversible Markov chains and random walks on graphs. – Monograph in preparation.
[2] Aldous (David) and Pemantle (Robin) (editors). – *Random discrete structures.* – Springer-Verlag, New York, 1996, xviii+225p. Papers from the workshop held in Minneapolis, Minnesota, November 15–19, 1993.
[3] Bingham (N. H.), Goldie (C. M.), and Teugels (J. L.). – *Regular variation.* – Cambridge University Press, Cambridge, 1987, xx+491p.
[4] Dembo (Amir), Peres (Yuval), Rosen (Jay), and Zeitouni (Ofer). – Cover times for Brownian motion and random walks in two dimensions. – To appear.
[5] Dembo (Amir), Peres (Yuval), Rosen (Jay), and Zeitouni (Ofer). – Thick points for planar Brownian motion and the Erdős-Taylor conjecture on random walk. *Acta Mathematica*, vol. 186, n° 2, 2001, pp. 239–270.
[6] Einmahl (Uwe). – Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *Journal of Multivariate Analysis*, vol. 28, n° 1, 1989, pp. 20–68.
[7] Erdős (P.) and Taylor (S. J.). – Some problems concerning the structure of random walk paths. *Acta Mathematica Academiae Scientiarum Hungaricae*, vol. 11, 1960, pp. 137–162. (Unbound insert).
[8] Feller (W.). – *An introduction to probability theory and its applications.* – Wiley international edition, 1970. Volume 1.
[9] Komlós (J.), Major (P.), and Tusnády (G.). – An approximation of partial sums of independent rv's and the sample df. I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 32, 1975, pp. 111–131.