Open Problems in Mathematics, Volume **2** (2014)

RANDOM EULERIAN CIRCUITS

DAVID ALDOUS AND JIANGZHEN YU

Abstract. What can one say about a uniform random Eulerian tour on the (bi-directed) discrete torus $\mathbb{Z}_N^d$, or other graph families?

Original proposer(s) of the open problem: David Aldous
The year when the open problem was proposed: 1996 in conversation, 2003 on web site.
Sponsor of the submission: Persi Diaconis (Stanford University).

A finite connected balanced directed graph has at least one Eulerian circuit, so on such a graph one can consider a uniform random Eulerian circuit $\mathcal{C}$. There is a simple algorithm – see e.g. Kandel-Matias-Unger-Winkler [2] – for simulating $\mathcal{C}$. Use the random walk method for simulating a uniform spanning tree, and regard that tree as directed toward an arbitrary root. This specifies the final exit the circuit will make from each non-root vertex. Now start the circuit at the root, and at each step choose uniformly from the possible edges not previously used, saving the pre-specified exit for last.

The notion of *uniform random spanning tree* is central to a large body of theory (see e.g. Chapter 4 of Lyons-Peres [3] for an introduction). In complete contrast we know of no nontrivial theory whatsoever about properties of $\mathcal{C}$ on particular graphs, so our purpose is to draw attention to this potentially fertile research topic. Even though the algorithm is simple to code, it is not easy to deduce theoretical properties of $\mathcal{C}$.

As a natural family of graphs to study, fix $d \geq 3$ and consider the d-dimensional discrete torus $\mathbb{Z}_N^d$ and replace each undirected edge by two directed edges, giving $2dN^d$ directed edges. A uniform random Eulerian circuit will have $2d$ excursions from the origin: write the lengths of these excursions in decreasing order as $L_1^{(N)} \geq L_2^{(N)} \geq \ldots \geq L_{2d}^{(N)}$. What can we say about the distribution of $(L_1^{(N)}, L_2^{(N)}, \ldots, L_{2d}^{(N)})$, in the $N \to \infty$ limit?

To make a conjecture, first consider the length $X^{(N)}$ of an excursion of simple symmetric walk on $\mathbb{Z}_N^d$, and then consider the random vector defined by

2010 *Mathematics Subject Classification.* 60C05.

This paper was supported in part by N.S.F Grant DMS-0704159.

(i) Let $X_1^{(N)}, \ldots, X_{2d}^{(N)}$ be i.i.d. copies of $X^{(N)}$;
(ii) Let $(Y_1^{(N)}, \ldots, Y_{2d}^{(N)})$ have the conditional distribution of $(X_1^{(N)}, \ldots, X_{2d}^{(N)})$ given $\sum_{i=1}^{2d} X_i^{(N)} = 2dN^d$;
(iii) Let $(Z_1^{(N)}, \ldots, Z_{2d}^{(N)})$ be the decreasing reordering of $(Y_1^{(N)}, \ldots, Y_{2d}^{(N)})$.

Our conjecture, in vague language, is

(1) the asymptotic distribution of $(L_1^{(N)}, \ldots, L_{2d}^{(N)})$ is qualitatively similar to that of $(Z_1^{(N)}, \ldots, Z_{2d}^{(N)})$.

So, notwithstanding the intricate detailed constraints forcing the excursions to use each directed edge exactly once, we are conjecturing that on the large scale they behave like random walks that are unconstrained (except via total length).

To make a precise conjecture we need to describe the asymptotic behavior of $(Z_1^{(N)}, \ldots, Z_{2d}^{(N)})$, which is essentially well understood. Recall we have fixed $d \geq 3$. Simple random walk on the infinite lattice $\mathbb{Z}^d$ has a certain probability p of never returning to the origin, and there is a certain distribution W of excursion length conditional on return. The asymptotic distribution of $X^{(N)}$ is a mixture [1]:

> with probability $1 - p$ it is W; with probability p it is Exponential, mean $p^{-1}N^d$.

This enables us to describe the asymptotic distribution of $(Z_1^{(N)}, \ldots, Z_{2d}^{(N)})$ as follows.

(iv) Take Q with a certain distribution supported on $\{1, \ldots, 2d\}$.
(v) Given Q, let $S_1, \ldots, S_Q$ be the decreasing reordering of the lengths of the Q subintervals of $[0, 1]$ obtained by cutting at $Q - 1$ i.i.d. uniform random points, and let $W_{Q+1}, \ldots, W_{2d}$ be i.i.d. copies of W.
(vi) Then

$$(2dN^d)^{-1}(Z_1^{(N)}, \ldots, Z_{2d}^{(N)}) \ \to_d \ (S_1, \ldots, S_Q, 0, \ldots, 0) \tag{2}$$

$$(Z_1^{(N)}, \ldots, Z_{2d}^{(N)}) \ \to_d \ (\infty, \ldots, \infty, W_{Q+1}, \ldots, W_{2d}). \tag{3}$$

Here the distribution of Q can be described in terms of p, but the value of p (and similarly the distribution of W) depends on the detailed local behavior of simple random walk and cannot be expected to extend unchanged to the setting of Eulerian circuits. However, the qualitative behavior in (2,3) is that excursion lengths are either order N^d or order 1, and we conjecture that this does remain true for random Eulerian circuits.

In particular, take $\omega_N \uparrow \infty$ very slowly, and define $(b^{(N)}, t^{(N)}, m^{(N)})$ (for *big, tiny, medium*) to be the number of excursions of $\mathcal{C}$ that are

$$(\text{longer than } N^d/\omega_N, \text{ shorter than } \omega_N, \text{ between } \omega_N \text{ and } N^d/\omega_N)$$

Conjecture 0.1. *For $d \geq 3$,*

$$(b^{(N)}, t^{(N)}, m^{(N)}) \to_d (S^*, 2d - S^*, 0) \text{ for some } S^* \text{ supported on } \{1, \ldots, 2d\}.$$

Data from simulations of excursion lengths (Figure 1) in $d = 3$ are consistent with this conjecture.

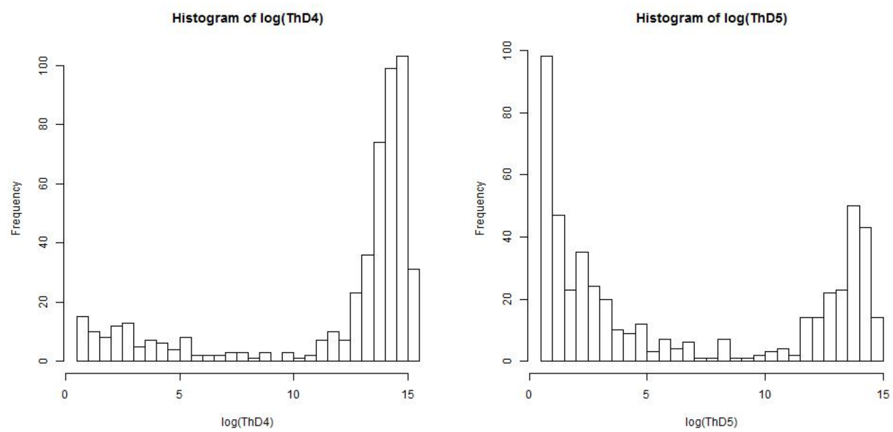

Figure 1. Simulated distributions of $\log L_4^{(N)}$ and $\log L_5^{(N)}$ for $N = 150$, $d = 3$.

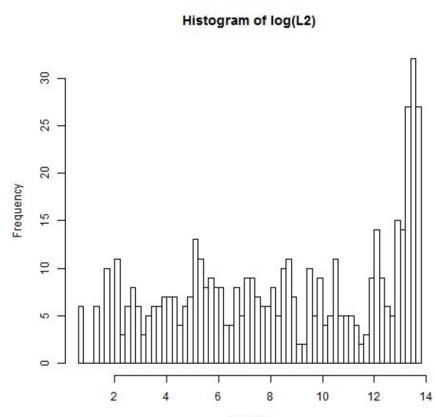

Figure 2. Simulated distribution of $\log L_2^{(N)}$ for $N = 700$, $d = 2$.

In the case $d = 2$ the vague conjecture (1) is not plausible because paths of the (now recurrent) random walk on $\mathbb{Z}^2$ must be rather different from those of a circuit. But thinking instead about the maximum distance from the origin reached in a walk excursion suggests the alternate conjecture that the distribution of $L_2^{(N)}$, the length of the second-longest excursion, is spread over the range $[N^{o(1)}, N^{2-o(1)}]$, more precisely that

$$\frac{\log L_2^{(N)}}{\log N} \to_d \xi \text{ for some limit } \xi \text{ with support } [0,2]. \tag{4}$$

Again this is consistent with simulations (Figure 2). Simulations of the circuits themselves can be seen at `http://www.stat.berkeley.edu/∼aldous/Research/Eulerian/`.

To briefly mention two other examples (again with undirected edges replaced by two directed edges):

(a) On the complete n-vertex graph, the random Eulerian circuit has n excursions, each of mean length $n-1$, and the natural conjecture is that the expected number of length-i excursions is asymptotic to $e^{-i/n}$.

(b) On the Hamming cube $\{0,1\}^d$ there are d excursions of mean length 2^d, and perhaps the most interesting open problem is to study the length $L_d^{(d)}$ of the shortest one.

Motivation. As mentioned initially, our original (1996) motivation was pure curiosity – is there some theory surrounding uniform random Eulerian circuits, parallel to the theory surrounding uniform random spanning trees? Also at that time there was no tractable model for random space-filling curves, and continuum limits of Eulerian circuits might possible serve as such. In recent years SLE with parameter $\kappa > 8$ has become a well-known such model [4]; is there some relation between that model and the presumed continuum limit of Eulerian circuits on $\mathbb{Z}_N^2$?

References

[1] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^d$. *Ann. Probab.*, 17(4):1333–1366, 1989.

[2] D. Kandel, Y. Matias, R. Unger, and P. Winkler. Shuffling biological sequences. *Discrete Appl. Math.*, 71(1-3):171–185, 1996.

[3] R. Lyons, with Y. Peres. (2014). *Probability on Trees and Networks*. Cambridge University Press. In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/`.

[4] S Rohde and O. Schramm. Basic properties of SLE. *Ann. of Math. (2)*, 161(2):883–924, 2005.

Department of Statistics, 367 Evans Hall # 3860, U.C. Berkeley CA 94720;

E-mail address: `aldous@stat.berkeley.edu`