The asymptotic distribution of the diameter of a random mapping *

David Aldous and Jim Pitman

Technical Report No. 606

Department of Statistics, University of California,
367 Evans Hall # 3860, Berkeley, CA 94720-3860

February 14, 2002

Abstract: The asymptotic distribution of the diameter of the digraph of a uniformly distributed random mapping of an n-element set to itself is represented as the distribution of a functional of a reflecting Brownian bridge. This yields a formula for the Mellin transform of the asymptotic distribution, generalizing the evaluation of its mean by Flajolet and Odlyzko (1990).

Titre: La loi limite du diamètre d'une application aléatoire.

Résumé: On exprime la loi limite du diamètre du digraphe d'une application aléatoire, choisie uniformément parmi les applications d'un ensemble à n éléments dans lui-même, comme la loi d'une fonctionnelle du pont brownien réfléchi. Ceci donne une formule pour la transformée de Mellin de cette loi limite, généralisant la formule pour sa moyenne due a Flajolet et Odlyzko (1990).

*Research supported in part by N.S.F. Grants DMS-9970901 and DMS-0071448

1 Introduction

Let F_n be a uniformly distributed random mapping from the set $[n] := \{1, 2, \ldots, n\}$ to itself, as studied in [10, 8, 1] and papers cited there. The *diameter* of F_n is the random variable

$$\Delta_n := \max_{i \in [n]} T_n(i)$$

where $T_n(i)$ is the number of iterations of F_n starting from i until some value is repeated:

$$T_n(i) := \min\{j \geq 1 : F_n^j(i) = F_n^k(i) \text{ for some } 0 \leq k < j\}$$

where $F_n^0(i) = i$ and $F_n^j(i) := F_n(F_n^{j-1}(i))$ is the image of i under j-fold iteration of F_n for $j \geq 1$. Flajolet-Odlyzko [8, Theorem 7] showed by singularity analysis of generating functions that

$$\lim_{n \to \infty} E(\Delta_n / \sqrt{n}) = \sqrt{\frac{\pi}{2}} \int_0^\infty (1 - e^{-E_1(v) - I(v)}) dv \qquad (1)$$

where

$$\begin{aligned} E_1(v) &:= \int_v^\infty u^{-1} e^{-u} du \\ I(v) &:= \int_0^v u^{-1} e^{-u} \left[1 - \exp\left(\frac{-2u}{e^{v-u} - 1}\right)\right] du. \end{aligned}$$

According to our analysis [1] of the asymptotic distributions of various functionals of random mappings, there is the convergence in distribution

$$\lim_{n \to \infty} P(\Delta_n / \sqrt{n} \leq x) = P(\Delta \leq x) \qquad (2)$$

for a limiting random variable Δ which can be constructed as a function of a standard Brownian bridge and a sequence of independent uniform$[0, 1]$ random variables, as indicated in Section 2. The main purpose of this note is to present the following more explicit description of the law of Δ, which gives probabilistic meaning to the function $e^{-E_1(v) - I(v)}$ appearing in (1).

Theorem 1 *The distribution of Δ is characterized by the formula*

$$P(|B_1| \Delta \leq v) = e^{-E_1(v) - I(v)} \qquad (v \geq 0) \qquad (3)$$

where B_1 is a standard Gaussian variable independent of Δ.

Corollary 2 *For each $p > 0$*

$$\lim_{n\to\infty} E\left[\left(\frac{\Delta_n}{\sqrt{n}}\right)^p\right] = E(\Delta^p) = \frac{p}{E(|B_1|^p)}\int_0^\infty v^{p-1}(1-e^{-E_1(v)-I(v)})dv. \quad (4)$$

Here $E(|B_1|^p) = 2^{p/2}\Gamma((p+1)/2)/\sqrt{\pi}$, so (4) for $p = 1$ reduces to (1). Formula (3) yields the second equality in (4), which characterizes the distribution of Δ by its Mellin transform. To justify the first equality in (4) we need uniform boundedness of each moment of $\Delta_n/\sqrt{n}$. But a bijection of Joyal [9] bounds Δ_n by twice the height of a uniform random tree labeled by $[n]$, and the corresponding uniform boundedness for this height follows from estimates of Łuczak [13]. See also [23, 12, 7] for related asymptotic studies of the diameter of undirected random trees and graphs.

2 A Brownian bridge representation of Δ

In [1] we showed how various features of the uniformly distributed random mapping F_n could be encoded as functionals of a particular non-Markovian random walk on the non-negative integers. This *mapping-walk*, with $2n$ equally spaced steps of size ± 1 starting and ending at 0, is constructed in such a way that as $n \to \infty$ the corresponding *scaled mapping-walk* $(F_u^{[n]}, 0 \leq u \leq 1)$, with $2n$ steps of $\pm 1/\sqrt{n}$ per unit time, and linear interpolation between steps, converges in distribution in $C[0,1]$ to a reflecting Brownian bridge:

$$(F_u^{[n]}, 0 \leq u \leq 1) \xrightarrow{d} (2|B^{\rm br}|, 0 \leq u \leq 1) \quad (5)$$

where $B^{\rm br}$ is the standard Brownian bridge derived from a one-dimensional Brownian motion $(B_t, t \geq 0)$ with $E(B_t) = 0$ and $E(B_t^2) = t$ as

$$B_u^{\rm br} := B_u - uB_1 \quad (0 \leq u \leq 1). \quad (6)$$

Let $|\mathcal{C}_n|$ denote the size of the random set of *cyclic points* of the mapping, that is

$$\mathcal{C}_n := \{i \in [n] : F_n^k(i) = i \text{ for some } k \geq 1\}.$$

As shown in [1], the convergence (5) holds jointly with

$$\frac{|\mathcal{C}_n|}{\sqrt{n}} \xrightarrow{d} L_1^{\rm br} \text{ where } P(L_1^{\rm br} \in d\ell) = \ell e^{-\frac{1}{2}\ell^2} d\ell \quad (\ell > 0) \quad (7)$$

and $L^{\rm br}_u, 0 \le u \le 1$ is the continuous increasing process of local time at 0 for $B^{\rm br}$. The equality in (7) is due to Lévy [11].

Still following [1], let the basins of attraction of F_n (i.e. the connected components of the usual digraph associated with F_n) be put in increasing order of their least elements. For $j = 1, 2, \ldots$

- let $N_{j,n}$ be the number of elements of $[n]$ in the jth basin of F_n,
- $C_{j,n}$ the length of the unique cycle in the jth basin of F_n,
- let $H_{j,n}$ the height above this cycle of the tallest tree in the jth basin of F_n.

According to [1, Theorem 8], the convergences in distribution (5) and (7) hold jointly with the convergence of finite-dimensional distributions

$$\left(\frac{N_{j,n}}{n}, \frac{C_{j,n}}{\sqrt{n}}, \frac{H_{j,n}}{\sqrt{n}}\right)_{j=1,2,\ldots} \stackrel{d}{\to} (\lambda_j, L_j, 2M_j)_{j=1,2,\ldots} \tag{8}$$

where the elements in the limit can be constructed as follows from the Brownian bridge $B^{\rm br}$ and a sequence of independent uniform $[0,1]$ variables $U_1, U_2, \ldots$ assumed independent of $B^{\rm br}$. For $0 \le v < 1$ let

$$D_v := \inf\{t > v : B^{\rm br}_t = 0\},$$

and note that that $D_0 = 0$ almost surely. Let $V(0) = 0$ and let random times $V(j)$ be defined inductively as follows for $j = 1, 2, \ldots$: given that $V(i)$ has been defined for $0 \le i < j$,

$$V(j) := D_{V(j-1)} + U_j(1 - D_{V(j-1)}),$$

so that $V(j)$ is uniform on $[D_{V(j-1)}, 1]$ given $B^{\rm br}$ and $V(i)$ for $0 \le i < j$. Then the sequence $(\lambda_j, L_j, 2M_j)_{j=1,2,\ldots}$ in (8) can be constructed from

$$\lambda_j := D_{V(j)} - D_{V(j-1)};\ L_j := L^{\rm br}_{D_{V(j)}} - L^{\rm br}_{D_{V(j-1)}};\ M_j := \max_{D_{V(j-1)} \le u \le D_{V(j)}} |B^{\rm br}_u|. \tag{9}$$

Since by definition

$$\Delta_n = \max_j (C_{j,n} + H_{j,n})$$

it is to be expected from (8) that the asymptotic distribution of $\Delta_n/\sqrt{n}$ is the distribution of

$$\Delta := \max_j (L_j + 2M_j) \tag{10}$$

and this is confirmed by further application of [1, Theorem 8].

It follows easily from the construction (9), the strong Markov property of B^{br} at the times $D_{V(j)}$, and Brownian scaling, that

$$\lambda_j = W_j \prod_{i=1}^{j-1} (1 - W_i) \tag{11}$$

for a sequence of independent random variables W_j with the beta$(1, \frac{1}{2})$ distribution $P(W_j > x) = \sqrt{1-x}$, $0 \le x \le 1$, and that

$$(L_j, M_j) = \sqrt{\lambda_j}(\tilde{L}_j, \tilde{M}_j) \tag{12}$$

for a sequence of independent and identically distributed random pairs $(\tilde{L}_j, \tilde{M}_j)$, independent of (λ_j). The common distribution of $(\tilde{L}_j, \tilde{M}_j)$ is that of

$$(\tilde{L}_1, \tilde{M}_1) := \left(\frac{L^{\mathrm{br}}_{D_1}}{\sqrt{D_U}}, \frac{M^{\mathrm{br}}_{D_1}}{\sqrt{D_U}} \right) \tag{13}$$

where D_U is the time of the first zero of B^{br} after a uniform$[0,1]$ random time U which is independent of B^{br}, and $M^{\mathrm{br}}_t := \max_{0 \le u \le t} |B^{\mathrm{br}}_u|$ for $0 \le t \le 1$. It follows from [18, Theorem 1.3] and [1, Proposition 2] that the process $B^{\mathrm{br}}_*[0, D_U]$, obtained by rescaling the path of B^{br} on $[0, D_U]$ to have length 1 by Brownian scaling, has the same distribution as a rearrangement of the path of the pseudo-bridge $\tilde{B}^{\mathrm{br}} := B_*[0, \tau_1]$ where τ_1 is an inverse local time at 0 for the unconditioned Brownian motion B. Neither the maximum nor the local time at 0 are affected by such a rearrangement, so there is the equality in distribution

$$(\tilde{L}_1, \tilde{M}_1) \stackrel{d}{=} (\tilde{L}^{\mathrm{br}}_1, \tilde{M}^{\mathrm{br}}_1) \tag{14}$$

where $\tilde{L}^{\mathrm{br}}_1$ is the local time of the pseudo-bridge $\tilde{B}^{\mathrm{br}}$ at 0 up to time 1, and $\tilde{M}^{\mathrm{br}}_1 := \max_{0 \le u \le 1} |\tilde{B}^{\mathrm{br}}_u|$. According to the absolute continuity relation between the laws of $\tilde{B}^{\mathrm{br}}$ and B^{br} found in [5], for each non-negative measurable function g on $C[0,1]$, $E[g(\tilde{B}^{\mathrm{br}})] = \sqrt{\frac{2}{\pi}} E[g(B^{\mathrm{br}})/L^{\mathrm{br}}_1]$. So (14) yields the formula

$$P(\sqrt{t}\tilde{L}_1 \in d\ell, \sqrt{t}\tilde{M}_1 \le y) = \sqrt{\frac{2}{\pi}} \frac{\sqrt{t}}{\ell} P(\sqrt{t} L^{\mathrm{br}}_1 \in d\ell, \sqrt{t} M^{\mathrm{br}}_1 \le y), \tag{15}$$

for $t, \ell, y > 0$, where the joint law of $L_1^{\rm br}$ and $M_1^{\rm br}$ is characterized by the following identity: for all $\ell > 0$ and $y > 0$

$$\int_0^\infty \frac{e^{-t/2}}{\sqrt{2\pi t}}\, dt\, P(\sqrt{t} L_1^{\rm br} \in d\ell, \sqrt{t} M_1^{\rm br} \le y) = e^{-\ell} d\ell \,\exp\left(\frac{-2\ell}{e^{2y}-1}\right). \qquad (16)$$

This can be read from [21, Theorem 3, Lemma 4 and (36)], with the following interpretation. Let $(L_t, t \ge 0)$ be the continuous increasing process of local time of the Brownian motion B at 0, let T be an exponential random variable with mean 2 independent of B, and let G_T be the time of the last 0 of B before time T. Then (16) provides two expressions for

$$P\left(L_T \in d\ell, \sup_{0 \le u \le G_T} |B_u| \le y \right),$$

on the left side by conditioning on G_T, and on the right side by conditioning on L_T. See also [22, Exercise (4.24)].

3 A Poisson representation of Δ

It is known [16] that for (λ_j) as in (11), assumed independent of B_1, the $B_1^2 \lambda_j$ are the points (in size-biased random order) of a Poisson process on $\mathbb{R}_{>0}$ with intensity measure $\frac{1}{2} t^{-1} e^{-t/2} dt$ which is the Lévy measure of the infinitely divisible gamma$(\frac{1}{2}, \frac{1}{2})$ distribution of B_1^2. Together with standard properties of Poisson processes, this observation and the previous formulae (10) to (13) yield the following lemma.

Lemma 3 *If B_1 is a standard Gaussian variable independent of the sequence of triples $(\lambda_j, L_j, M_j)_{j=1,2,\ldots}$ featured in (8) and (9), then the random vectors $(B_1^2 \lambda_j, |B_1| L_j, |B_1| M_j)$ are the points of a Poisson point process on $\mathbb{R}^3_{>0}$ with intensity measure μ defined by*

$$\mu(dt\, d\ell\, dm) = \frac{e^{-t/2}\, dt}{2t} P(\sqrt{t} \tilde{L}_1 \in d\ell, \sqrt{t} \tilde{M}_1 \in dm) \qquad (17)$$

for $t, \ell, m > 0$, where $(\tilde{L}_1, \tilde{M}_1)$ is the pair of random variables derived from a Brownian bridge by (13), and the distribution of Δ defined by either (2) or (10) is characterized by the formula

$$|B_1| \Delta = \max_j (|B_1| L_j + 2 |B_1| M_j). \qquad (18)$$

Using (17), (15) and (16), we deduce that the expected number of points $(|B_1|L_j, |B_1|M_j)$ with $|B_1|L_j \in d\ell$ and $|B_1|M_j \leq y$ is

$$\int_0^\infty \frac{e^{-t/2}\,dt}{2t} P(\sqrt{t}\tilde{L}_1 \in d\ell, \sqrt{t}\tilde{M}_1 \leq y) = \ell^{-1}e^{-\ell}d\ell \exp\left(\frac{-2\ell}{e^{2y}-1}\right). \tag{19}$$

The functions $E_1(v)$ and $I_1(v)$ featured in Theorem 1 can now be interpreted as follows: $E_1(v)$ is the expected number of j with $|B_1|L_j \geq v$, while $I(v)$ is the expected number of j with $|B_1|L_j < v$ and $|B_1|L_j + 2|B_1|M_j > v$. The probability of the event $|B_1|\Delta \leq v$, that there is no j with $|B_1|L_j + 2|B_1|M_j > v$, is therefore $e^{-E_1(v)-I(v)}$. The conclusion of Theorem 1 is now evident.

4 Related Results

As indicated in [8] there are companions to (1) for other functionals of F_n besides the diameter, in particular the total length of cycles and the maximum tree height. One advantage of the present approach is that all these results can be understood in terms of the Poisson representation of Lemma 3. For instance, the limit distribution of maximum tree height, with normalization by $\sqrt{n}$, is known [1] to be that of $2M_1^{\mathrm{br}}$. The analog of (3), which follows from Lemma 3, (19), and the Lévy-Khintchine formula for the exponential distribution, is the known result that for B_1 standard Gaussian independent of M_1^{br}, and $y > 0$

$$P(|B_1|M_1^{\mathrm{br}} \leq y) = \tanh y = \frac{1}{1 + 2/(e^{2y}-1)}. \tag{20}$$

As observed in [6], formula (20) allows the Mellin transform of M_1^{br} to be expressed in terms of the Riemann zeta function. See also [15, 19, 20] for closely related Mellin transforms obtained by the technique of multiplication by a suitable independent random factor to introduce Poisson or Markovian structure.

In [2] and [4] we show that Brownian bridge asymptotics apply to models of random mappings more general than the uniform model, in particular for the *p-mapping* model of [14, 17], and that proofs can be simplified by use of Joyal's bijection between mappings and trees. In [3] we develop a variety of distributional results dealing with the decomposition of Brownian bridge at the times $D_{V(j)}$ and with an alternate decomposition motivated by an alternate encoding of mappings as walks; these results involve a Poisson point process representation of path fragments which extends Lemma 3.

References

[1] D. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5:487–512, 1994.

[2] D. Aldous and J. Pitman. Invariance principles for non-uniform random mappings and trees. Technical Report 594, Dept. Statistics, U.C. Berkeley, 2001. To appear in *Asymptotic Combinatorics and Mathematical Physics*, Proceedings of NATO Advanced Study Institute, St Petersburg 2001, edited by V. Malyshev and A. Vershik, 2002.

[3] D. Aldous and J. Pitman. Two recursive decompositions of Brownian bridge related to the asymptotics of random mappings. Technical Report 595, Dept. Statistics, U.C. Berkeley, 2002.

[4] D.J. Aldous, G. Miermont, and J. Pitman. Limit distributions for non-uniform random mappings, involving Brownian bridge and inhomogeneous continuum random trees. In preparation, 2002.

[5] P. Biane, J. F. Le Gall, and M. Yor. Un processus qui ressemble au pont brownien. In *Séminaire de Probabilités XXI*, pages 270–275. Springer, 1987. Lecture Notes in Math. 1247.

[6] P. Biane, J. Pitman, and M. Yor. Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc.*, 38:435–465, 2001.

[7] F. Chung and L. Lu. The diameter of sparse random graphs. *Adv. in Appl. Math.*, 26:257–279, 2001.

[8] P. Flajolet and A. Odlyzko. Random mapping statistics. In J.-J. Quisquater and J. Vandewalle, editors, *Advances in Cryptology – EUROCRYPT '89*, pages 329–354. Springer-Verlag, 1990. Lecture Notes in C.S. 434.

[9] A. Joyal. Une théorie combinatoire des séries formelles. *Adv. in Math.*, 42:1–82, 1981.

[10] V.F. Kolchin. *Random Mappings.* Optimization Software, New York, 1986. (Translation of Russian original).

[11] P. Lévy. Sur certains processus stochastiques homogènes. *Compositio Math.*, 7:283–339, 1939.

[12] T. Łuckzak. Random trees and random graphs. *Random Structures Algorithms*, 13:485–500, 1998.

[13] T. Łuczak. The number of trees with large diameter. *J. Austral. Math. Soc.*, 58:298–311, 1995.

[14] C. A. O'Cinneide and A. V. Pokrovskii. Nonuniform random transformations. *Ann. Appl. Probab.*, 10(4):1151–1181, 2000.

[15] M. Perman. Order statistics for jumps of normalized subordinators. *Stoch. Proc. Appl.*, 46:267–281, 1993.

[16] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Th. Rel. Fields*, 92:21–39, 1992.

[17] J. Pitman. Random mappings, forests and subsets associated with Abel-Cayley-Hurwitz multinomial expansions. *Séminaire Lotharingien de Combinatoire*, Issue 46:45 pp., 2001. www.mat.univie.ac.at/˜slc/.

[18] J. Pitman and M. Yor. Arcsine laws and interval partitions derived from a stable subordinator. *Proc. London Math. Soc. (3)*, 65:326–356, 1992.

[19] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25:855–900, 1997.

[20] J. Pitman and M. Yor. The law of the maximum of a Bessel bridge. *Electron. J. Probab.*, 4:Paper 15, 1–35, 1999.

[21] J. Pitman and M. Yor. On the distribution of ranked heights of excursions of a Brownian bridge. *Ann. Probab.*, 29:362–384, 2001.

[22] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer, Berlin-Heidelberg, 1999. 3rd edition.

[23] G. Szekeres. Distribution of labelled trees by diameter. In *Combinatorial Mathematics X*, pages 392–397. Springer-Verlag, 1982. Lecture Notes in Math. 1036.