1 Introduction

These are notes about a result of David Aldous from 1997 which relates the sizes of connected components of a critical random graph with something to do with a Brownian motion.

1.1 Random Graphs

We will consider the *Erdős-Rényi model* $G(n,p)$, which is a graph with n labeled vertices. The edges are chosen randomly, with each possible edge $\{u,v\}$ having probability p to be in the graph (all independently). This model is most interesting if $p = p_n$ is a function of n. Probably the most famous result here is that if $p_n \ll \frac{\log n}{n}$ then $G(n,p_n)$ is asymptotically almost surely disconnected, and if $p_n \gg \frac{\log n}{n}$ then $G(n,p_n)$ is asymptotically almost surely connected.

Now let's look a bit closer at the situation where $p_n = \frac{\lambda}{n}$, where λ is some fixed positive number. By the above paragraph, the graphs here are a.a.s. disconnected, but there is still some interesting behavior to observe. There is a *phase transition* of the behavior of $G(n,\frac{\lambda}{n})$ as λ changes, where a giant component emerges when $\lambda > 1$. Specifically, let $C_j(n)$ denote the size of the jth largest component in $G(n,\frac{\lambda}{n})$. Then

$$\begin{aligned} \text{for } \lambda < 1, \qquad & C_1(n) = \Theta(\log n), \\ \text{but for } \lambda > 1, \qquad & C_1(n) = \Theta(n) \text{ and } C_2(n) = \Theta(\log n). \end{aligned}$$

This was proved by Pál Erdős and Alfréd Rényi in 1960. They also examined the behavior at the critical point $\lambda = 1$ and found that in this regime we have

$$C_1(n) = \Theta(n^{2/3}) \text{ and } C_2(n) = \Theta(n^{2/3}).$$

In 1997, Aldous expanded on this result and showed that $C_j(n) = \Theta(n^{2/3})$ for any fixed j. Specifically, he proved that

$$\left(\frac{1}{n^{2/3}} C_j(n)\right)_{j\geq 1} \xrightarrow{d} (C_j)_{j\geq 1}$$

for some random sequence-valued process $(C_j)_{j\geq 1}$ with $0 < C_j < \infty$ almost surely. But what is $(C_j)_{j\geq 1}$?

1.2 Brownian Excursions

Let $W(s)$ denote a standard Brownian motion in $\mathbb{R}$ for $s \geq 0$, starting at 0. You don't need to know too much about Brownian motion to follow these notes, because I don't know too much about Brownian motion. One way to think about $W(s)$ is simply as a scaling limit of a simple random walk on $\mathbb{Z}$. Now let

$$\widetilde{W}(s) = W(s) - \frac{s^2}{2}, \qquad s \geq 0$$

which is a Brownian motion with a downward drift. Specifically, the drift at time s' is $-s'$. So, the $-\frac{s^2}{2}$ comes from integrating the drift at each time, up to time s. Now let

$$B(s) = \widetilde{W}(s) - \min_{0\leq s'\leq s} \widetilde{W}(s'), \qquad s \geq 0.$$

This is called a "reflecting Brownian motion" even though it's not really being reflected in the standard sense of the word. It is constrained to lie in $[0,\infty)$ however. Informally, when $\widetilde{W}(s)$ wants to go below 0, $B(s)$ simply stays at 0, and starts going back up exactly when $\widetilde{W}(s)$ starts going back up. Notice that the drift of $\widetilde{W}(s)$ means that $B(s)$ will spend more and more time very at 0 or very close as s increases, because the "drift" is *not* reflected—$B(s)$ is pulled down with the same intensity that $\widetilde{W}(s)$ is, but there is a "floor" in the way, preventing $B(s)$ from going below 0.

An *excursion* γ of $B(s)$ is a time interval $[l(\gamma), r(\gamma)]$ such that $B(l(\gamma)) = B(r(\gamma)) = 0$ and $B(s) > 0$ for $l(\gamma) < s < r(\gamma)$. Let C_j be the length of the jth largest excursion of $B(s)$. This definition makes sense because $B(s)$ is pulled down with greater intensity over time, so there will actually be a maximum excursion almost surely, and it will happen close to the start of $B(s)$. Of course, this needs to be proved, but we won't do that here. You can find the proof in Aldous's paper. By the way, these excursions can also be thought of as the excursions of $\widetilde{W}(s)$ above its minimum.

This process $(C_j)_{j\geq 1}$ of excursions of $B(s)$ is the distributional limit of the sequence of component sizes of critical Erdős-Rényi graphs. You might think this isn't very interesting because we've just replaced a complicated object with an even more complicated one. But simulating Brownian motion is very easy, since it's just a limit of a simple random walk. On the other hand, finding the sizes of components of a large Erdős-Rényi graph can be quite slow. The Breadth-First Search algorithm runs in $O(n)$ time and space, which isn't too bad, but actually initializing the graph to run the search algorithm takes $O(n^2)$ time. Simulating a simple random walk for a very large time can be tough too, but it is only linear in the time. Of course, I am heavily glossing over the details here, since the order of approximation needs to be taken into account on both sides.

Anyway, I made a simulation of this "Brownian excursions" process, it's on my website at

vilas.us/miscmath/brownianexcursions

In practice, it seems that the top 10 excursions happen well before $s = \frac{1}{2}$; the simulation above shows the top 10 excursions and simulates to time $s = \frac{1}{2}$. I also simulated a lot of non-animated trials of the excursion process and generated the following sample means and sample covariance matrix for the top five excursions:

$$\mu = 10^{-2}\begin{bmatrix}4.250\\1.800\\1.212\\0.947\\0.781\end{bmatrix}, \qquad K = 10^{-4}\begin{bmatrix}6.152 & 0.059 & -0.243 & -0.243 & -0.219\\0.059 & 0.523 & 0.138 & 0.052 & 0.019\\-0.243 & 0.138 & 0.162 & 0.080 & 0.042\\-0.243 & 0.052 & 0.080 & 0.079 & 0.046\\-0.219 & 0.019 & 0.042 & 0.046 & 0.044\end{bmatrix}.$$

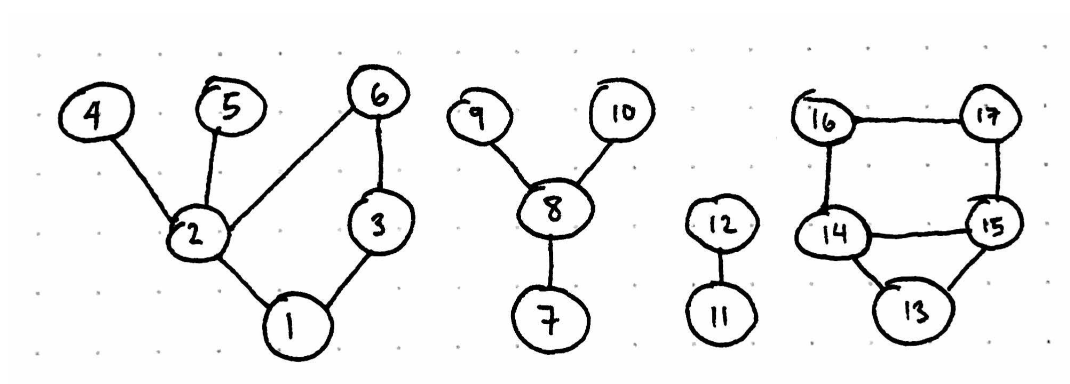
I did 1000 trials, and each time I simulated a random walk for 10000 steps to represent a Brownian motion between $s = 0$ and $s = 1$. In total, it took my computer about 3 minutes.

2 Ideas

The main idea of the proof that the component sizes in $G(n, \frac{1}{n})$ converge in distribution to the Brownian excursion process is the following construction, called the breadth-first walk. It is a walk in $\mathbb{Z}$ that is created based on a graph such that the excursions of this walk above its minimum are exactly the sizes of the components of the graph. Later, we will see that the proper scaling limit of this walk gives $B(s)$.

2.1 The Breadth-First Walk

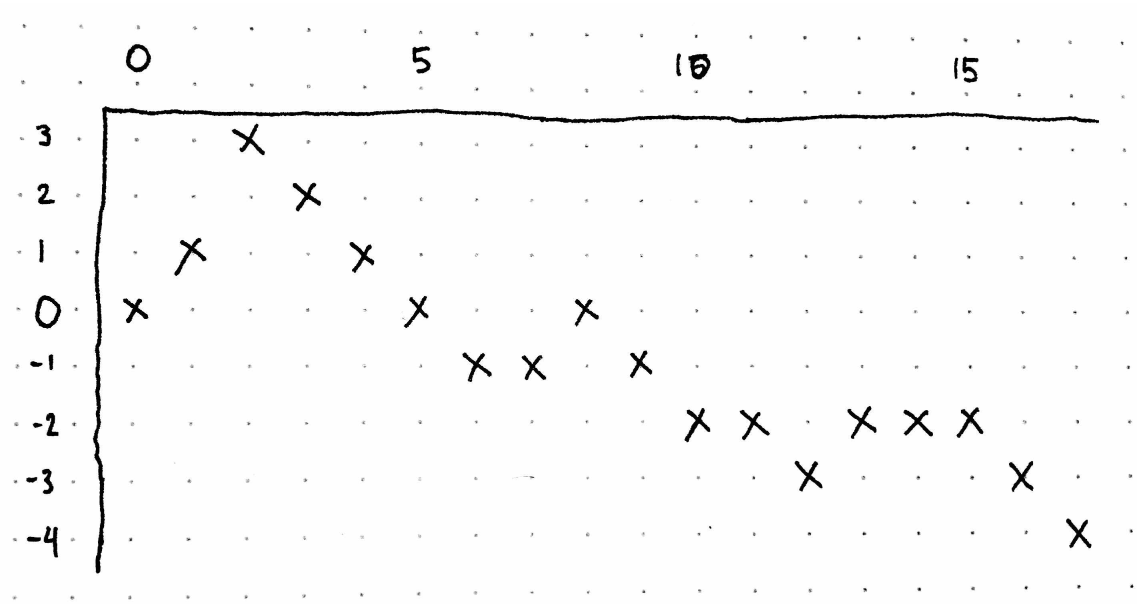
Given any graph labeled graph G, order the vertices $v(1), \ldots, v(n)$ in BFS order. Here's an example of BFS ordering just to refresh your memory:



For posterity, we'll use the labels of G to choose the next vertex in case there are no children. Let $c(i)$ denote the number of "children" of $v(i)$ in the BFS tree. Now define the Breadth-First Walk z by

$$z(0) = 0, \qquad \text{and} \qquad z(i) = z(i-1) + c(i) - 1 \quad \text{for} \quad i = 1, \ldots, n.$$

Let's see what the breadth-first walk looks like on our previous example:



Notice that the walk stays $\geq$ its minimum exactly as long as there are more nodes to explore in the "current" component. So the lengths of the excursions of $z(i)$ above its minimum are exactly the sizes of the components. To prove this rigorously, we need a bit more notation.

Let $\mathscr{N}_i$ denote the set of neighbors of vertices in $\{v(1), \ldots, v(i)\}$ which are not elements of $\{v(1), \ldots, v(i)\}$ themselves. Notice that $c(i) = |\mathscr{N}_i \setminus \mathscr{N}_{i-1}|$. Also let $\mathscr{C}_j$ denote the jth component in the BFS order (this is *not* the same as the jth largest component in general). Let

$$\begin{aligned} \zeta(j) &= |\mathscr{C}_1| + \cdots + |\mathscr{C}_j|, \\ \zeta^{-1}(i) &= \min\{j : \zeta(j) \geq i\} = \text{the index of the component containing } v(i). \end{aligned}$$

We claim that

$$z(i) = |\mathscr{N}_i| - \zeta^{-1}(i). \tag{$*$}$$

Since $|\mathscr{N}_i| = 0$ exactly when $v(i)$ is the last vertex in its component, this implies

$$z(\zeta(j)) = -j, \qquad \text{and} \qquad z(i) \geq -j \quad \text{for} \quad \zeta(j) < i < \zeta(j+1).$$

Therefore we have

$$\begin{aligned} \zeta(j) &= \min\{i : z(i) = -j\}, \\ |\mathscr{C}_j| &= \zeta(j) - \zeta(j-1), \\ \zeta^{-1}(i) &= 1 - \min_{i' \leq i-1} z(i'). \end{aligned}$$

Now, to prove ($*$) we go by induction. We need to show that

$$|\mathscr{N}_i| - |\mathscr{N}_{i-1}| = c(i) - 1 + \zeta^{-1}(i) - \zeta^{-1}(i-1) \qquad \text{for} \quad i = 2, \ldots, n.$$

Suppose $v(i-1)$ is *not* the last vertex in its component. Then $\zeta^{-1}(i) = \zeta^{-1}(i-1)$. Also, because $v(i) \in \mathscr{N}_{i-1}$, we have $|\mathscr{N}_i| - |\mathscr{N}_{i-1}| = c(i) - 1$. On the other hand, if $v(i-1)$ *is* the last vertex in its component, then $\zeta^{-1}(i) = 1 + \zeta^{-1}(i-1)$. Also, $|\mathscr{N}_i| = c(i)$ and $|\mathscr{N}_{i-1}| = 0$, so everything works out.

2.2 Main Results

First of all, the breadth-first walk, appropriately scaled, converges to $\tilde{W}(s)$ in distribution. Specifically

Theorem 1. *Let $(Z_n(i), 0 \le i \le n)$ be the breadth-first walk associated to $G(n, \frac{1}{n})$. Then*

$$\frac{1}{n^{1/3}} Z_n\left(\left\lfloor n^{2/3} s \right\rfloor\right) \xrightarrow{d} \widetilde{W}(s),$$

where $\widetilde{W}(s) = W(s) - \frac{s^2}{2}$ (and $W(s)$ is standard Brownian motion in $\mathbb{R}$ started at 0). The convergence is "uniform on finite intervals" rather than on the whole interval $[0, \infty)$.

A bit more needs to be said in order to prove the real main result of this paper:

Corollary 2. *Let $C_j(n)$ denote the jth largest component of $G(n, \frac{1}{n})$. Then*

$$\left(\frac{1}{n^{2/3}} C_j(n)\right)_{j \ge 1} \xrightarrow{d} (C_j)_{j \ge 1},$$

where $(C_j)_{j\ge1}$ is the sequence of excursions of $B(s) = \widetilde{W}(s) - \min_{s' \le s} \widetilde{W}(s')$ above 0, ordered in decreasing order. The convergence is with respect to the product topology—namely, convergence of initial segments of arbitrary fixed length.

To prove this, we need to show two more things. First, we need to check that the excursions of the limit process are really the limits of the excursions of the breadth-first walk. We also need to check that no large components are overlooked by virtue of their positions in the breadth-first walk going off to infinity. Both Theorem 1 and Corollary 2 have rather technical proofs that may or may not be covered in these notes.

2.3 Extras

Actually, a bit more can be said and proved using essentially the same techniques and technology as we have already developed. First, we can examine the behavior not just of $G(n, \frac{1}{n})$, but of the more general

$$G\left(n, \frac{1}{n} + \frac{t}{n^{4/3}}\right) \qquad \text{for arbitrary fixed } t \in \mathbb{R}.$$

Then everything we have said goes through but with $\widetilde{W}(s)$ replaced by

$$\widetilde{W}^t(s) = W(s) + ts - \frac{s^2}{2}.$$

This brownian motion has drift $t - s$ at time s, so if t is positive it starts out drifting up and then comes back down. Aldous says that it is "well known" that the $n^{-4/3}$ scaling is "correct" for the emergence of the giant component, in the sense that

$$C_1^t \xrightarrow{d} 0 \qquad\qquad \text{as } t \to -\infty,$$
$$C_1^t \xrightarrow{d} \infty \quad \text{but} \quad C_2^t \xrightarrow{d} 0 \qquad \text{as } t \to +\infty,$$

where C_j^t is the length of the jth longest excursion of $B^t(s) = \widetilde{W}^t(s) - \min_{s' \le s} \widetilde{W}^t(s')$ above zero.

Another "extra", we can get is a better understanding of the number of *edges* in each component. Specifically, define the number of surplus edges in a component to be the number of edges that component has *beyond* the ones it needs to be connected, i.e.

$$\text{surplus} = (\text{number of edges}) - (\text{number of vertices} - 1) \ge 0.$$

If $\sigma_j^t(n)$ is the number of surplus edges in the jth component of $G(n, \frac{1}{n} + \frac{t}{n^{3/4}})$, then

$$\left(n^{-2/3} \sigma_j^t(n)\right)_{j \ge 1} \xrightarrow{d} \left(\sigma_j^t\right)_{j \ge 1}$$

for some process $(\sigma_j^t)_{j\geq 1}$ also derived from our reflecting brownian motion. In particular, σ_j^t is the number of *marks* in the excursion of length C_j^t under a point process on $\mathbb{R}$ of intensity $B^t(s)$. Informally, this means that

$$\mathbb{P}(\text{some mark during } [s, s+ds] | B^t(u), u \leq s) = B^t(s)\, ds.$$

The idea here is that in the breadth-first search, whenever we see an edge that leads back to the already-explored vertices, we should add a "mark" to $\mathbb{Z}$ (note that there may be multiple marks at one time step). If the breadth-first walk is currently higher above its previous minimum, this means that there are a lot of already-explored vertices in the current component, and so we will be more likely to see an extra edge and write a mark. And of course, this point process on $\mathbb{Z}$, appropriately scaled, should converge to the point process described above on $\mathbb{R}$.

3 Proofs

The full details of the proofs will not be spelled out here. Also, we will only discuss the proofs of Theorem 1 and Corollary 2, none of the extras will be discussed. The first extra (with the parameter t) can be proved with essentially no change to the idea we present here. The second extra (about the surplus edges) is similar.

3.1 Convergence of the Breadth-First Walk

We will need to interpolate between integer values of i to define $Z_n(s)$ for noninteger s. One strategy is to set $Z_n(s) = Z_n(\lfloor s \rfloor)$. This imagines that the children of $v(i)$ are all added at once at step i. But another way to interpolate, that turns out to be easier to analyze, is as follows.

Let $(U_{i,j} : 1 \leq i \leq n, 1 \leq j \leq c(i))$ be independent random variables uniform in $[0, 1]$ (also independent from $G(n, \frac{1}{n})$). Then for $0 \leq u \leq 1$ define

$$Z_n(i - 1 + u) = Z_n(i - 1) - u + \sum_{j=1}^{c(i)} \mathbf{1}_{\{U_{i,j} \leq u\}}.$$

In other words, the children of $v(i)$ are added at uniformly random times between $i - 1$ and i. Now let

$$\bar{Z}_n(s) = \frac{1}{n^{1/3}} Z_n(n^{2/3} s).$$

Here is the outline of the proof that $\bar{Z}_n(s) \xrightarrow{d} \widetilde{W}(s)$ uniformly on finite intervals. Write

$$Z_n = M_n + A_n,$$

where $M_n(s)$ is a (continuous-time) martingale in the variable s, and $A_n(s)$ is a continuous, bounded variation process. Then write

$$M_n^2 = Q_n + B_n,$$

where $Q_n(s)$ is a martingale and $B_n(s)$ is a continuous increasing process. By the way, we can ensure that

$$Z_n(0) = M_n(0) = A_n(0) = Q_n(0) = B_n(0) = 0.$$

We will show that as $n \to \infty$ with s_0 fixed, we have

$$\frac{1}{n^{1/3}} \sup_{s \leq n^{2/3} s_0} \left| A_n(s) + \frac{s^2}{2n} \right| \xrightarrow{p} 0, \tag{1}$$

$$\frac{1}{n^{2/3}} B_n\left(n^{2/3} s_0\right) \xrightarrow{p} s_0, \tag{2}$$

$$\frac{1}{n^{2/3}} \mathbb{E}\left[\sup_{s \leq n^{2/3} s_0} |M_n(s) - M_n(s-)|^2 \right] \longrightarrow 0. \tag{3}$$

If we rescale the same way that we did for $\bar{Z}_n$ to define $\bar{A}_n$, $\bar{M}_n$, and $\bar{B}_n$, the above three assertions become

$$\sup_{s \le s_0} \left| \bar{A}_n(s) + \frac{s^2}{2} \right| \xrightarrow{p} 0, \tag{1'}$$

$$\bar{B}_n(s_0) \xrightarrow{p} s_0, \tag{2'}$$

$$\mathbb{E}\left[\sup_{s \le s_0} \left| \bar{M}_n(s) - \bar{M}_n(s-) \right|^2 \right] \longrightarrow 0. \tag{3'}$$

It turns out that (2′) and (3′) are exactly the hypothesis of the "functional central limit theorem for continuous-time martingales" whose conclusion is simply that $\bar{M}_n \xrightarrow{d} W$, the standard Brownian motion. And together with (1′), this implies that

$$\bar{Z}_n(s) = \bar{M}_n(s) + \bar{A}_n(s) \xrightarrow{d} W(s) - \frac{s^2}{2} = \widetilde{W}(s),$$

by Slutsky's theorem. So it just remains to prove (1), (2), and (3). For (3), notice that the jumps of $Z_n(s)$ have size exactly 1, and since $A_n(s)$ is continuous the jumps of $M_n(s)$ are the same as the jumps of $Z_n(s)$. Therefore (3) is clear, and we just need to prove (1) and (2). For this we need a lemma, for which we define

$$\zeta_n^{-1}(i) = 1 - \min_{s \le i-1} Z_n(s). \tag{$\dagger$}$$

Lemma 3. *We have*

$$A_n(s) = \int_0^s (a_n(u) - 1)\, du,$$
$$B_n(s) = \int_0^s a_n(u)\, du,$$

where

$$a_n(u) = \frac{n - u - \zeta^{-1}(\lceil u \rceil) - Z_n(u)}{n - (u - \lfloor u \rfloor)}.$$

Proof. A_n should capture the "unfairness" of Z_n. More precisely (and informally), we should have

$$A_n(s + ds) - A_n(s) = \mathbb{E}[Z_n(s + ds) - Z_n(s) | Z_n(s'), s' \le s],$$

i.e. the "increments" of A_n should be the expected increments of Z_n, so that the expected increments of M_n are zero, to ensure that M_n is a martingale. Since Z_n is a process with drift -1 and jumps $+1$ whenever a new edge appears, the first formula should hold with a_n defined by

$$a_n(u)\, du = \mathbb{P}(\text{some new edge appears during } [u, u + du] | Z_n(u'), u' \le u).$$

As for B_n, observe that (again, informally)

$$\mathbb{E}[M_n^2(s)] = \int \mathbb{E}[(M_n(u + du) - M_n(u))^2 | M_n(u'), u' \le u],$$

by the non-correlation of increments of a martingale (all of the "cross terms" vanish). Therefore, we should have

$$B_n(s + ds) - B_n(s) = \text{variance of } M_n(s + ds), \text{ conditioned on } M_n(u'), u' \le u.$$

Recall that, conditional on $M_n(s')$ for $s' \le s$, the increment $M_n(s + ds) - M_n(s)$ is 1 with probability $a_n(s)\, ds$, and 0 with probability $1 - a_n(s)\, ds$. This is a Bernoulli random variable with success probability $a_n(s)\, ds$, so its variance is $a_n(s)\, ds - (a_n(s)\, ds)^2$. Since $ds^2 = 0$, we can ignore the second term, and obtain that $B_n(s + ds) - B_n(s) = a_n(s)\, ds$, which shows that the second formula in the statement holds with the same function $a_n(s)$ as defined above.

Now it's time to actually compute what $a_n(s)$ is. An apparently elementary calculation shows that if an event occurs with probability q and, conditionally on independence, it occurs at a random time uniform in $[0,1]$, then

$$\mathbb{P}(\text{it occurs during } [u, u+du] | \text{it does not occur before } u) = \frac{q}{1-uq}\, du.$$

So, by construction of the breadth-first walk, using $q = \frac{1}{n}$ which is the probability that a specific vertex is in the child set of $v(\lceil u \rceil)$, we have

$$a_n(u) = (n - \nu_n(u)) \frac{1/n}{1-(u-\lfloor u \rfloor)/n},$$

where $\nu_n(u)$ is the number of vertices at time s which are ineligible to be children of $v(\lceil u \rceil)$. When we start looking for children of $v(i)$ at time $i-1$, the number of ineligible vertices is

$$\nu_n(i-1) = i - 1 + |\mathscr{N}_{i-1}| + (\zeta^{-1}(i) - \zeta^{-1}(i-1)),$$

where $v(i)$ itself is taken care of by the final term. Now by $(*)$, this is the same as

$$\nu_n(i-1) = i - 1 + \zeta_n^{-1}(i) + Z_n(i-1).$$

So, at time $i-1+w$ (for $0 < w < 1$) the number ineligible is

$$\begin{aligned}\nu_n(i-1+w) &= i - 1 + \zeta_n^{-1}(i) + Z_n(i-1) + \sum_j \mathbf{1}_{\{U_{i,j} \le w\}} \\ &= (i-1+w) + \zeta_n^{-1}(i) + Z_n(i-1+w),\end{aligned}$$

by our interpolation convention. In other words, $\nu_n(s) = s + \zeta_n^{-1}(\lceil s \rceil) + Z_n(s)$ and the lemma is proved. ■

Lemma 3 allows us to rewrite (2) as

$$\frac{1}{n^{2/3}} A_n\left(n^{2/3} s_0\right) \xrightarrow{p} 0,$$

which is implied by (1). So we will just prove (1). Instead of a_n, we will consider a slightly modified integrand

$$a'_n(u) = \frac{n - u - \zeta^{-1}(\lceil u \rceil) - Z_n(u)}{n}.$$

Aldous says it's straightforward to see that $|a'_n(u) - a_n(u)| = O(1/n)$ uniformly in u. Now we have

$$a'_n(u) - 1 = -\frac{u + \zeta^{-1}(\lceil u \rceil) + Z_n(u)}{n},$$

and so

$$\left| a'_n(u) - 1 + \frac{u}{n} \right| \le 2 \frac{\zeta_n^{-1}(\lceil u \rceil) + |Z_n(u)|}{n}. \qquad (\#)$$

Integrating u from 0 to s and using (†), we obtain

$$\left| A_n(s) + \frac{s^2}{2n} \right| \le \frac{4s \max_{u \le s} |Z_n(u)|}{n} + O\left(\frac{s}{n}\right).$$

So the proof of (1) reduces to proving that

$$\frac{1}{n^{2/3}} \sup_{s \le n^{2/3} s_0} |Z_n(s)| \xrightarrow{p} 0.$$

We will prove the stronger result that

$$\frac{1}{n^{1/3}} \sup_{s \le n^{2/3} s_0} |Z_n(s)| \qquad \text{is stochastically bounded (tight) as } n \to \infty.$$

This uses the martingale optional stopping theorem. Fix a large constant K and define

$$\begin{aligned} T_n^* &= \min\{s : |Z_n(s)| > Kn^{1/3}\}, \\ T_n &= \min\{T_n^*, s_0 n^{2/3}\}. \end{aligned}$$

Then T_n is a stopping time, and so

$$\begin{aligned} \mathbb{E}\left[M_n^2(T_n)\right] &= \mathbb{E}\left[B_n(T_n)\right] \\ &= \mathbb{E}\left[\int_0^{T_n} a_n(u)\,du\right] \\ &\le \int_0^{s_0 n^{2/3}} \frac{n}{n-(u-\lfloor u\rfloor)}\,du \\ &\le 2s_0 n^{2/3}, \end{aligned}$$

the final inequality holding only for n at least 2. Thus

$$\begin{aligned} \mathbb{E}[|Z_n(T_n)|] &\le \mathbb{E}[|M_n(T_n)|] + \mathbb{E}[|A_n(T_n)|] \\ &\le (2s_0)^{1/2} n^{1/3} + \mathbb{E}\left[\int_0^{T_n} |a_n(u)-1|\,du\right]. \end{aligned}$$

Using (†) and (#), we have

$$\mathbb{E}\left[\int_0^{T_n} |a_n(u)-1|\,ds\right] \le \mathbb{E}\left[\int_0^{s_0 n^{2/3}} |a_n'(u)-a_n(u)|\,du\right] + \int_0^{s_0 n^{2/3}} \frac{u}{n}\,du + (s_0 n^{2/3})4\frac{Kn^{1/3}}{n}.$$

Putting it all together, we get a bound for large n:

$$\mathbb{E}[|Z_n(T_n)|] \le \alpha n^{1/3} + 4s_0 K,$$

where α depends on s_0 but not on n or K. Therefore

$$\mathbb{P}\left(\frac{1}{n^{1/3}} \sup_{s \le s_0 n^{2/3}} |Z_n(s)| > K\right) = \mathbb{P}\left(|Z_n(T_n)| > Kn^{1/3}\right) \le \frac{\alpha}{K} + \frac{4s_0}{n^{1/3}}$$

by Markov's inequality. This finishes the proof of Theorem 1.

3.2 Convergence of the Excursions

The first thing we need to show is that the excursions of $\bar{Z}_n(s)$ actually converge to the excursions of $B(s)$. Since we have $\bar{Z}_n(s) \xrightarrow{d} \widetilde{W}(s)$ uniformly on finite intervals, we can use Skorohod's representation theorem to upgrade this to almost sure convergence (uniformly on finite intervals). Then we apply the following deterministic lemma (we write an excursion $\gamma = (\ell(\gamma), r(\gamma))$, and $\ell(\gamma)$ denotes the length of an excursion).

Lemma 4. *Suppose the union of the excursions of a function $f : [0,\infty) \to \mathbb{R}$ above its previous minimum has full Lebesgue measure. Let*

$$\Xi = \{(l(\gamma), \ell(\gamma)) : \gamma \text{ an excursion of } f \text{ above its previous minimum}\}.$$

Suppose now that $f_n \to f$ and let $(t_{n,i} : i \ge 1)$ satisfy

(*i*) $0 = t_{n,1} < t_{n,2} < \cdots,$ *and* $\lim_{i\to\infty} t_{n,i} = \infty;$

(*ii*) $f_n(t_{n,i}) = \min_{u \le t_{n,i}} f_n(u);$

(*iii*) $\max_{i : t_{n,i} \le s_0} (f_n(t_{n,i}) - f_n(t_{n,i+1})) \to 0$ *as* $n \to \infty,$ *for each* $s_0 < \infty.$

Define

$$\Xi_n = \{(t_{n,i}, t_{n,i+1} - t_{n,i}) : i \geq 1\}.$$

Then $\Xi_n \to \Xi$*, where the convergence is interpreted as convergence of point processes on* $[0,\infty) \times (0,\infty)$*, i.e. as vague convergence of counting measures on* $[0,\infty) \times (0,\infty)$*.*

Aldous leaves the proof of this lemma as an exercise. Since the convergence obtained in the lemma could, a priori, have some mass escaping off to infinity, we need to prove that this doesn't happen in our scenario. In other words, we need to ensure that no components of size $\Omega(n^{2/3})$ are overlooked by appearing later and later in the breadth-first walk so that their position goes off to infinity. Consider

$$T(y) = \min\{s : \tilde{W}(s) = -y\},$$
$$T_n(y) = \min\left\{i : Z_n(i) = -\left\lfloor yn^{1/3}\right\rfloor\right\}.$$

Notice that by step $T_n(y)$, the breadth-first walk has encountered all vertices labeled $\left\{1, 2, \ldots, \left\lfloor yn^{1/3}\right\rfloor\right\}$ in the *original* labeling. The weak convergence implies that $n^{-2/3}T_n(y) \xrightarrow{d} T(y)$. So if we fix y_0 and only consider excursions of $B(s)$ starting before $T(y_0)$ and components of the graph whose minimal original vertex labels are $\leq y_0 n^{1/3}$, then we obtain the convergence we want. The next lemma shows that this is enough:

Lemma 5. *Let* $p(n, y, \delta)$ *be the probability that* $G(n, \frac{1}{n})$ *contains a component of size* $\geq \delta n^{2/3}$ *which does not contain any vertex* i *with* $1 \leq i \leq yn^{1/3}$*. Then*

$$\lim_{y\to\infty} \limsup_{n\to\infty} p(n, y, \delta) = 0 \qquad \textit{for all } \delta > 0.$$

Proof. To prove this, we punt to a result from the random graphs literature. For any fixed $\delta > 0$, define $q_\delta(n, I)$ to be the average number of components of size at least $\delta n^{2/3}$ whose minimal vertex label is in $n^{1/3}I$. Notice that

$$p(n, y, \delta) \leq q_\delta(n, [y, \infty)).$$

Now, Conditional on component sizes, the labels $\{1, \ldots, n\}$ of the vertices of $G(n, \frac{1}{n})$ are in random order. For a component having size $vn^{2/3}$, write $\chi_n = n^{-1/3} \times$ (label of minimal vertex). Apparently, $\chi_n \xrightarrow{d}$ exponential of rate v, which implies that

$$\mathbb{P}(\chi_n > y) \sim \frac{e^{-vy}}{1 - e^{-v}} \mathbb{P}(\chi_n \leq 1).$$

By summing over components, this means that

$$\limsup_{n\to\infty} \frac{q_\delta(n, [y, \infty))}{q_\delta(n, [0, 1])} \leq \sup_{v \geq \delta} \frac{e^{-vy}}{1 - e^{-v}} = \frac{e^{-\delta y}}{1 - e^{-\delta}}.$$

Therefore, it suffices to prove that

$$\sup_n q_\delta(n, [0, 1]) < \delta.$$

The result "from the literature" (personal communication from Boris Pittel to David Aldous, who says it follows from bounds on the numbers of tree components, unicyclic components, and complex components of $G(n, \frac{1}{n})$, which are given in a 1994 paper of Luczak, Pittel, and Wierman) is

$$\sup_n q_\delta(n, [0, \infty)) < \infty.$$

This implies the result. ∎

And that's the end of the proof of Corollary 2, and the end of these notes. Thanks for reading!