Coalescing Random Walks on the N-Cycle

Iddo Ben-Ari[1] Lillian Makhoul[2] Bram Silbert[3]

[1]University of Connecticut

[2]Lehigh University

[3]Wesleyan University

Joint Mathematics Meetings
January 7, 2023

Overview

1. The Model
2. Coalescence Model on the Line
3. Coalescence Model on the Cycle
   - Time to Coalescence
   - Last Surviving Particle

The Model

The **N-cycle** is a graph with the vertex set $\{0, 1, ..., N-1\}$ and the edge set $\{\{i,j\} : (i-j) \equiv 1 \mod N \text{ or } (j-i) \equiv 1 \mod N\}$.

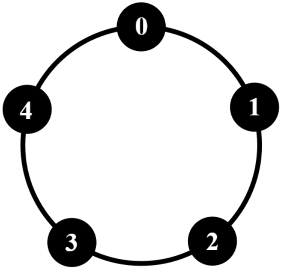


Figure: The 5-Cycle

The Model

We look at a particle system on the N-cycle under the following process:

1. At time $t = 0$, all nodes on the cycle are occupied by a particle

The Model

We look at a particle system on the N-cycle under the following process:

1. At time $t = 0$, all nodes on the cycle are occupied by a particle
2. At each time step, one particle is sampled uniformly and moved one node in the clockwise direction

The Model

We look at a particle system on the N-cycle under the following process:

1. At time $t = 0$, all nodes on the cycle are occupied by a particle
2. At each time step, one particle is sampled uniformly and moved one node in the clockwise direction
3. If the new vertex is occupied, the particle absorbs the previous occupant

The Model

We look at a particle system on the N-cycle under the following process:

1. At time $t = 0$, all nodes on the cycle are occupied by a particle
2. At each time step, one particle is sampled uniformly and moved one node in the clockwise direction
3. If the new vertex is occupied, the particle absorbs the previous occupant
4. Repeat until only one particle remains

The Model

We look at a particle system on the N-cycle under the following process:

1. At time $t = 0$, all nodes on the cycle are occupied by a particle
2. At each time step, one particle is sampled uniformly and moved one node in the clockwise direction
3. If the new vertex is occupied, the particle absorbs the previous occupant
4. Repeat until only one particle remains

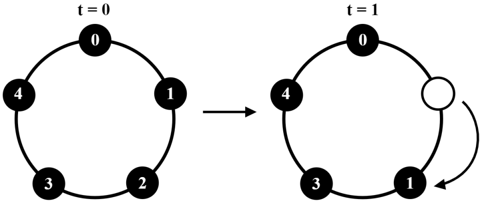


Figure: One step of the process on the 5-Cycle, in which particle 1 is sampled

Problems

There are two central questions we seek to answer:

1. What can we say about the distribution of the time before all particles coalesce in a single node?
2. What can we say about the distribution of the number of steps taken by the last surviving particle?

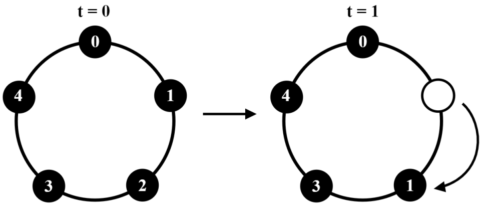


Figure: One step of the process on the 5-Cycle, in which particle 1 is sampled

Coalescence Model on the Line

A closely related model was studied by Russell Lyons and Michael Larsen in 1999. They focused on a coalescing random walk model on the line, where each particle moves to the left when sampled until eventually reaching an absorbing state at 0.

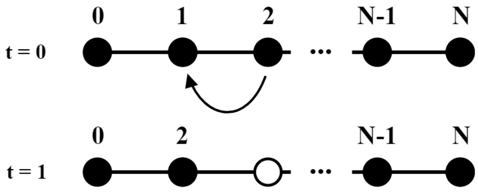


Figure: Coalescing random walk on the line

This linear model is essentially our model on the cycle, but with a single edge removed.

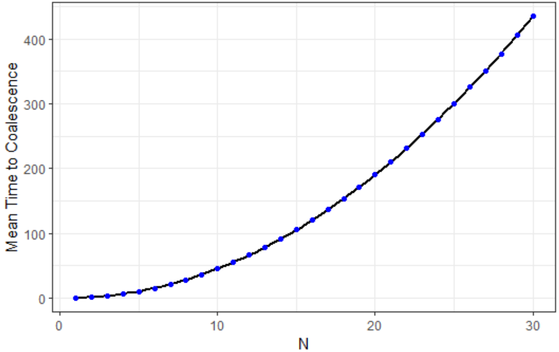
Contrasting the Two Models

- On the line:
  - The formula for $\mathbb{E}[T_N]$ and the upper bound for $Var[T_N]$ can be derived using a single generating function
  - The last surviving particle is always the particle starting at node N
  - The last surviving particle always takes N steps until absorption
- On the cycle:
  - The formula for $\mathbb{E}[T_N]$ is known, while the second moment and variance remain elusive
  - The identity of the last surviving particle is uniformly distributed among all particles in the system
  - The number of steps taken by the last surviving particle is random

Numerics: Expectation of Time to Coalescence

Let T_N be the time until all particles have coalesced on the N-cycle. Then, the simulated data suggests that $\mathbb{E}[T_N]$ is simply the $(N-1)$th triangular number, $N(N-1)/2$.

N	Simulated $\mathbb{E}[T_N]$
1	0.0000
2	1.0000
3	2.9995
4	5.9963
5	9.9763
6	15.0168
7	20.9875
8	28.0297
9	35.9710
10	44.9907



Best fit $f(N) = 0.4985N^2 - 0.4579N - 0.1658$

Theorem: Expectation of Time to Coalescence

Theorem

Let T_N be the time at which all particles on the N-cycle coalesce. Then

$$\mathbb{E}[T_N] = \frac{N(N-1)}{2}$$

Proof: Expectation of Time to Coalescence

Proof

Let S_i be the number of steps taken by particle i on the N-cycle before either i absorbs $i+1$ or $i+1$ absorbs i.

Then, the sum of all the S_i's will count the steps taken by all particles in the system before coalescence, with no double-counting. Thus, since exactly one particle moves at every time step,

$$T_N = \sum_{i=0}^{N-1} S_i \tag{1}$$

By the linearity of expectation,

$$\mathbb{E}[T_N] = \sum_{i=0}^{N-1} \mathbb{E}[S_i] \tag{2}$$

Proof: Expectation of Time to Coalescence

To find $\mathbb{E}[S_i]$, we will think of the two-particle system with i and $i+1$ as a simple random walk of a particle p on [0, N], starting at $N-1$, and with absorbing states at 0 and N.

- If particle i moves, then p moves right
- If particle $i+1$ moves, then p moves left
- Absorption corresponds to p reaching either 0 or N

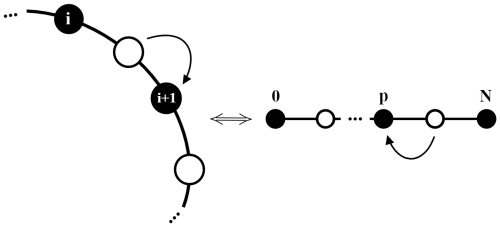


Figure: Walk on the cycle vs. walk on the line

Proof: Expectation of Time to Coalescence

Definition

A sequence of random variables $\{X_n\}$ is a **Martingale** if it satisfies $\mathbb{E}[X_{n+1}|X_1, ..., X_n] = X_n$.

Doob's Optional Stopping Theorem

Let $\{X_t\}$ be a discrete time Martingale with stopping time τ. Suppose that one of the following conditions is true:

1. $\mathbb{P}(\tau < \infty) = 1$
2. $\mathbb{E}[\tau] < \infty$ and $\mathbb{E}[|X_{t+1} - X_t|] \leq c$ for some constant c
3. $|X_{t\wedge\tau}| \leq c$ for some constant c

Then, $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

Proof: Expectation of Time to Coalescence

The essential idea that allows us to complete the proof is that we can define $Y_t = X_t^2 - t$, where X_t is the position of the particle p at time t, and Y_t will be a Martingale.

From there, we can apply the Optional Stopping Theorem in order to find

$$\mathbb{E}[\tau] = N - 1 \tag{3}$$

where τ is the time until absorption.

We can then see that $S_i = R$, $X_\tau = X_0 + R - L$, and $\tau = R + L$, where R and L are the number of right and left steps taken by p, respectively.

This allows us to deduce that

$$\mathbb{E}[S_i] = \frac{\mathbb{E}[\tau]}{2} = \frac{N-1}{2} \tag{4}$$

Proof: Expectation of Time to Coalescence

Note that this process is the same for all particles in the system, so $\mathbb{E}[S_0] = ... = \mathbb{E}[S_{N-1}]$. Therefore, equation (2) implies that

$$\mathbb{E}[T_N] = N\mathbb{E}[S_i] \tag{5}$$

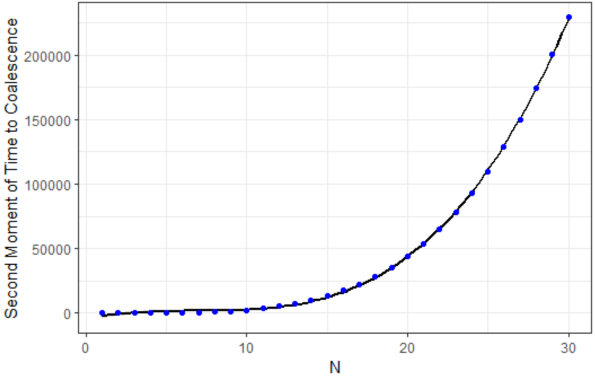
Therefore, from (4) and (5) we get our final result:

$$\mathbb{E}[T_N] = \frac{N(N-1)}{2} \tag{6}$$

$\square$

Numerics: Second Moment of Time to Coalescence

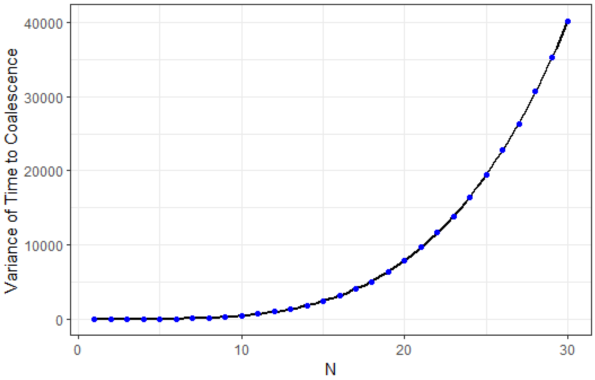
N	Simulated $\mathbb{E}[T_N^2]$
1	0.0000
2	1.0000
3	10.9876
4	44.9945
5	124.6430
6	281.4022
7	548.0339
8	976.2537
9	1601.3991
10	2493.4380



Best fit $f(N) = 0.28991N^4 - 0.05823N^3 - 5.91343N^2 + 27.67320N - 33.18489$

Numerics: Variance of Time to Coalescence

N	Simulated $Var[T_N]$
1	0.0000
2	0.0000
3	1.9904
4	9.0386
5	25.1159
6	55.8987
7	107.5602
8	190.5904
9	307.4886
10	469.2778



$$\text{Best fit } f(N) = 0.04663N^4 + 0.13320N^3 - 1.84915N^2 + 6.55435N - 6.78751$$

Numerics: Second Moment and Variance of Time to Coalescence

N	Simulated $\mathbb{E}[T_N^2]$
1	0.0000
2	1.0000
3	10.99976485
4	45.00229318
5	125.20625498

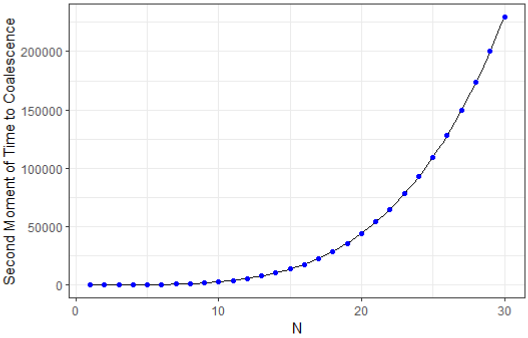
Best fit polynomial: $f(N) = 0.299819640000005N^4 - 0.497696628333394N^3 - 0.00942880499997549N^2 - 0.0148682133329450 5N + 0.192437580000001984$

From these results we can conjecture formulas for the second moment and variance:

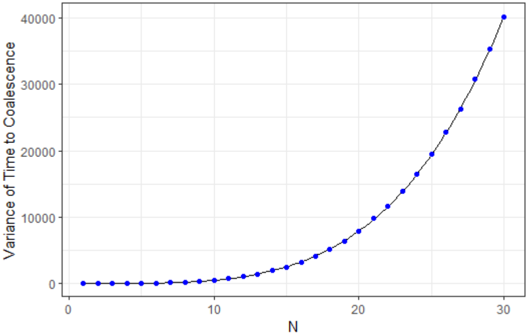
$$\mathbb{E}[T_N^2] = \frac{3N^4}{10} - \frac{N^3}{2} + \frac{1}{5}$$

$$Var[T_N] = \frac{(N^2-1)(N^2-4)}{20}$$

Numerics: Second Moment and Variance of Time to Coalescence



Simulated data plotted against
$$f(N) = \frac{3N^4}{10} - \frac{N^3}{2} + \frac{1}{5}$$



Simulated data plotted against
$$f(N) = \frac{(N^2-1)(N^2-4)}{20}$$

Numerics: Second Moment and Variance of Time to Coalescence

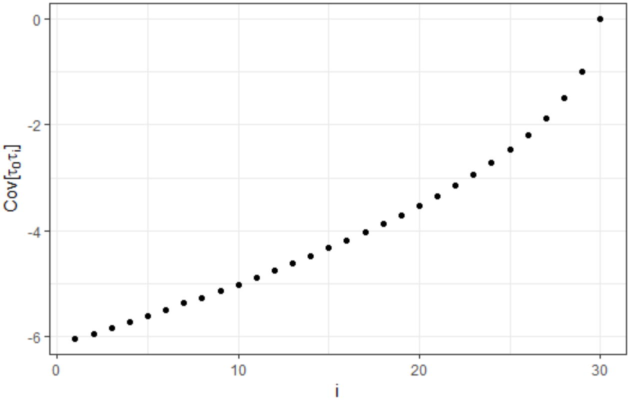
Unfortunately, we have been unable to prove an explicit formula for the second moment or variance of T_N, or even find a bound for the variance as Lyons and Larsen did for the coalescence model on the line.

They identified a generating function that allowed them to get explicit formulas for every $\mathbb{E}[\tau_i]$ and $\mathbb{E}[\tau_i^2]$, where τ_i is the number of steps particle $i+1$ takes before absorbing particle i. Thus, noting that $T_N = \sum_{i=0}^{N-1} \tau_i$, they arrived at

$$\begin{aligned} Var[T_N] &= \mathbb{E}[T_N^2] - \mathbb{E}[T_N]^2 \\ &= (\sum_{i=0}^{N-1} \mathbb{E}[\tau_i^2] + \sum_{i \neq j} \mathbb{E}[\tau_i \tau_j]) - (\sum_{i=0}^{N-1} \mathbb{E}[\tau_i]^2 + \sum_{i \neq j} \mathbb{E}[\tau_i]\mathbb{E}[\tau_j]) \\ &= (\sum_{i=0}^{N-1} \mathbb{E}[\tau_i^2] - \mathbb{E}[\tau_i]^2) + (\sum_{i \neq j} \mathbb{E}[\tau_i \tau_j] - \mathbb{E}[\tau_i]\mathbb{E}[\tau_j]) \end{aligned} \tag{7}$$

Numerics: Second Moment and Variance of Time to Coalescence

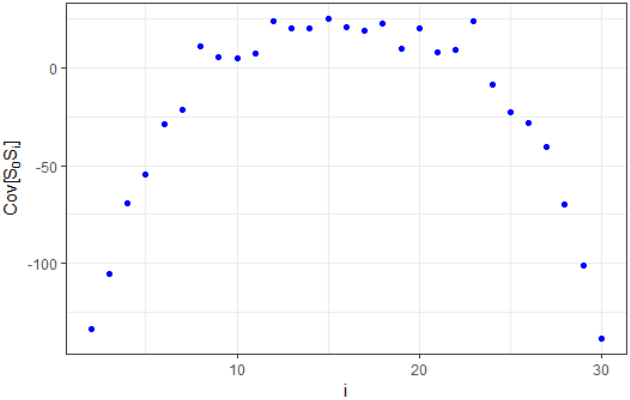
To get a valid upper bound for the variance, it is sufficient to show that all the τ_i's are negatively correlated, which would make the second term of (7) negative.



Numerics: Second Moment and Variance of Time to Coalescence

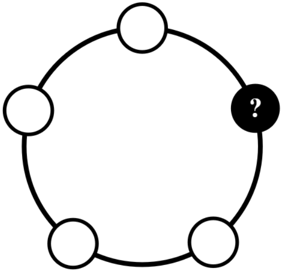
One might expect that the same approach could be applied to the model on the cycle, with the S_i's from the proof of Theorem 1 replacing the τ_i's.

Unfortunately, the numerical data shows that the S_i's are not necessarily negatively correlated, making Lyons and Larsen's method inapplicable in our case.



The Last Surviving Particle

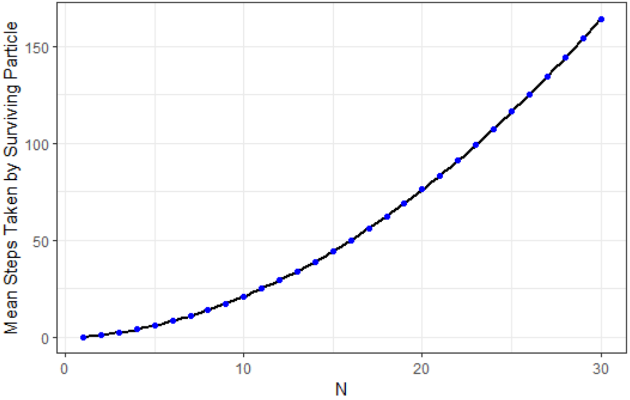
We now turn to the question of the last surviving particle. At first glance, we can tell that it must take at least $N-1$ steps before the process concludes, as it must necessarily absorb its counterclockwise neighbor as its final action.



However, beyond this basic fact, it is not necessarily clear what we should expect.

Numerics: The Last Surviving Particle

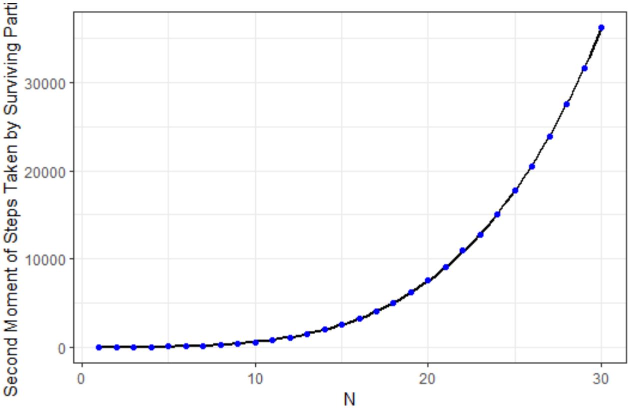
N	Simulated $\mathbb{E}[\mathcal{L}]$
1	0.000
2	1.000
3	2.3312
4	3.9939
5	6.000
6	8.3402
7	11.0000
8	13.9607
9	17.3364
10	20.9714



Best fit $f(N) = 0.1666N^2 + 0.5058N - 0.7037$

Numerics: The Last Surviving Particle

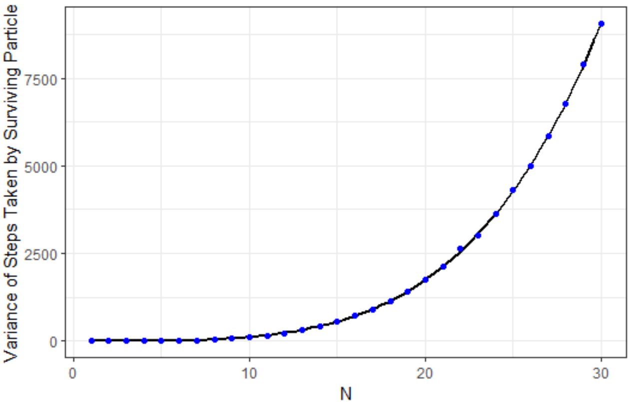
N	Simulated $\mathbb{E}[\mathcal{L}^2]$
1	0.0000
2	1.0000
3	5.8881
4	17.9447
5	41.5349
6	81.7824
7	145.5295
8	238.1156
9	368.9168
10	547.6482



Best fit $f(N) = 0.0462N^4 - 0.2260N^3 + 6.6775N^2 - 40.4179N + 58.5697$

Numerics: The Last Surviving Particle

N	Simulated Var[$\mathcal{L}$]
1	0.0000
2	0.0000
3	0.4474
4	1.9901
5	5.5858
6	12.4767
7	24.3715
8	42.1143
9	68.5747
10	106.3594



Best fit $f(N) = 0.0149N^4 - 0.2059N^3 + 3.4974N^2 - 21.3330N + 31.6806$

Proof: The Last Surviving Particle

Theorem

Let $\mathcal{L}$ denote the number of steps taken by the last surviving particle. Then,

$$\mathbb{E}[\mathcal{L}] = \frac{N^2}{6} + \frac{N}{2} - \frac{2}{3} \tag{8}$$

$$\mathbb{E}[\mathcal{L}^2] = \frac{7N^4}{180} + \frac{N^3}{6} - \frac{N^2}{36} - \frac{2N}{3} + \frac{22}{45} \tag{9}$$

$$Var[\mathcal{L}] = \frac{(N^2-1)(N^2-4)}{90} \tag{10}$$

Proof: The Last Surviving Particle

Outline of proof:

- Note that $\mathcal{L}$ is equal to the number of steps taken by $i+1$ in the two-particle system described earlier, if we condition on $i+1$ absorbing i
- Think of this two-particle system as a single particle random walk on the line
- Find a moment generating function for this conditioned random walk

Lemma: The Last Surviving Particle

Lemma

Let a and b be two consecutive particles on an N-cycle, and let T_N be the time before b absorbs a, conditioned on b absorbing a. Then,

$$\mathbb{E}[e^{-\lambda T_N}] = N \frac{\sinh\left(\frac{1}{2}\log(x)\right)}{\sinh\left(\frac{N}{2}\log(x)\right)}$$

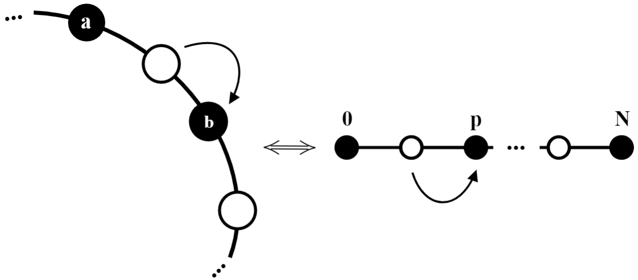
where

$$x = \frac{1-\sqrt{1-e^{-2\lambda}}}{1+\sqrt{1-e^{-2\lambda}}}$$

Proof of Lemma: The Last Surviving Particle

Proof of Lemma

Just as before, think of the two-particle random walk with a and b as a single particle walk on [0, N], with absorbing barriers at both endpoints. However, this time we view the walk from the perspective of b, starting at 1 and conditioning on absorption at N, since we only care about the case where b overtakes a.



Proof of Lemma: The Last Surviving Particle

Suppose the particle is at position i at time t. We want to know its transition probabilities conditioned on absorbing at N. We start with the transition function p:

$$p(i,j) = \begin{cases} 1/2 & j = i+1,\ j = i-1 \\ 0 & otherwise \end{cases} \tag{11}$$

We can note that the probability of the particle reaching N before 0 starting at i is $\phi(i) = i/N$, and then define

$$q(i,j) = \mathbb{P}_i(X_1 = j | T_N < T_0) = \frac{\phi(j)}{\phi(i)} p(i,j) = \frac{j}{i} p(i,j)$$

where X_1 is the state of the particle after one step, and T_N and T_0 are the times until the particle reaches N and 0, respectively. So, $Q = \{q(i,j)\}_{i,j}$ is the transition matrix for our random walk conditioned on $T_N < T_0$.

Proof of Lemma: The Last Surviving Particle

We want to find $\mathbb{E}[e^{-\lambda T_N}]$ conditioned on $T_N < T_0$. To do this, we can consider the expectation starting at the point i and express it as a sum of the possible expectations after one step:

$$\mathbb{E}_i[e^{-\lambda T_N}] = q(i, i+1)\mathbb{E}_{i+1}[e^{-\lambda(1+T_N)}] + q(i, i-1)\mathbb{E}_{i-1}[e^{-\lambda(1+T_N)}] \tag{12}$$

Note that $\mathbb{E}[e^{-\lambda T_N}]$ is just $\mathbb{E}_i[e^{-\lambda T_N}]$ where $i = 1$.
After extensive manipulations of the terms in (12), we can solve for our desired moment generating function:

$$\mathbb{E}[e^{-\lambda T_N}] = N\frac{\sinh\left(\frac{1}{2}\log(x)\right)}{\sinh\left(\frac{N}{2}\log(x)\right)} \tag{13}$$

where

$$x = \frac{1-\sqrt{1-e^{-2\lambda}}}{1+\sqrt{1-e^{-2\lambda}}}$$

Theorem: The Last Surviving Particle

Theorem

Let $\mathcal{L}$ denote the number of steps taken by the last surviving particle. Then,

$$\mathbb{E}[\mathcal{L}] = \frac{N^2}{6} + \frac{N}{2} - \frac{2}{3}$$

$$\mathbb{E}[\mathcal{L}^2] = \frac{7N^4}{180} + \frac{N^3}{6} - \frac{N^2}{36} - \frac{2N}{3} + \frac{22}{45}$$

$$Var[\mathcal{L}] = \frac{(N^2-1)(N^2-4)}{90}$$

Proof of Theorem: The Last Surviving Particle

Proof of Theorem

Consider again the random walk on [0, N] described in the proof of the preceding lemma. Note that

$$T_N = R + L$$

where R and L are the number of right and left steps taken by the particle, respectively.

Further, since we've conditioned on the particle reaching N, we must have

$$R = L + N - 1$$

Combining these expressions, we arrive at the equality

$$R = \frac{T_N + N - 1}{2} \tag{14}$$

Proof of Theorem: The Last Surviving Particle

We can note that $\mathcal{L}$, the number of steps taken by the last surviving particle, is exactly R, and hence

$$
\begin{aligned}
\mathbb{E}[e^{-\lambda\mathcal{L}}] &= \mathbb{E}\left[e^{-\lambda\left(\frac{T_N+N-1}{2}\right)}\right] \\
&= \left(e^{\frac{-\lambda(N-1)}{2}}\right)\mathbb{E}\left[e^{\left(\frac{-\lambda}{2}\right)T_N}\right] \\
&= N\left(e^{\frac{-\lambda(N-1)}{2}}\right)\left(\frac{\sinh\left(\frac{1}{2}\log(x)\right)}{\sinh\left(\frac{N}{2}\log(x)\right)}\right) \qquad (15)
\end{aligned}
$$

where

$$
x = \frac{1-\sqrt{1-e^{-\lambda}}}{1+\sqrt{1-e^{-\lambda}}}
$$

The desired formulas (8), (9), and (10) follow immediately when we take the limits of the first and second derivatives of (15) as $\lambda \to 0$. □

Thank You!