Def: Given \((X_t), \mathcal{F}_t\), the cover time \(T_{cov}\) is the first time that all states have been visited.

We are interested in the expected value of \(T_{cov}\) from the “worst case” starting position. Formally, this is:

\[
t_{cov} = \max_{x \in \mathcal{X}} E_x T_{cov}.
\]

**Example:** Random walk on \(n\)-cycle

This can be thought of as a random walk on \(\mathbb{Z}\) mod \(n\).

\(T_{cov}\) is the same as the expected time it takes to visit \(n\) states in a walk on \(\mathbb{Z}\), since they will all be contiguous and therefore different mod \(n\).

How long does it take to visit a new state?

If we have just reached the \((k-1)\)th state for the first time, and are looking to get to the \(k\)th state, we have this picture:

![Diagram showing a random walk on a cycle with \(k-1\) visited states and \(k\)th state not yet visited.]

We are at the end of the contiguous region and want to get to one of the white states. This is the gambler's ruin problem where we start with \(k+1\) dollars and stop at \(k\) or when we are broke. The expected time is \((1)(k-1) = k-1\).

So the expected time to visit \(n\) new states is

\[
\sum_{k=1}^{n} \frac{n(n-1)}{2}.
\]
Recall: If we look at any two points $a, b \in S$ we can ask the expected time to go from $a$ to $b$. The max over all pairs is $t_{hit}$. Formally,

$$t_{hit} := \max_{a,b \in S} E_a \tau_b$$

The bound for the cover time is similar to the coupon collector problem, except instead of taking expected time $n$ to collect a particular coupon it takes $\leq t_{hit}$, and instead of taking $\log n$ to collect all of them, it takes $t_{hit} \log n$.

**PF:** Assume $s = E^n I$ and let $x \in S$ be an arbitrary starting state, and choose $s \in S_n$ uniformly at random. (Think of $s$ as an ordering of the states, we will look to collect these states in $s$ order).

Let $T_k$ be the first time that we have collected all of $s(1), s(2), \ldots, s(k)$. (so the quantity we are looking for is $t_{cov} = E_x [T_k]$.)

What is $E[T_1]$?

$E_x [T_1] = E_x \tau_{s(1)} \leq t_{hit}$
Do this more formally, we need exercise 11.1a

Claim: Let $Y$ be a random variable on a probability space and let $B = \bigcup_j B_j$ be a partition of $B$.

Then if $E(Y | B_j) \leq M$ for all $j$, then $E(Y | B) \leq M$.

Proof:

$$E(Y | B) = \frac{1}{P(B)} \sum_j P(B_j) E(Y | B_j) \leq \frac{1}{P(B)} \sum_j P(B_j) M = M$$

So if $s \in \Omega$,

$$E_x(T_1 | \sigma(1) = s) = E_x \gamma_s \leq t_{hit}.$$ 

And since the events $\sigma(1) = s$ partition the space $\Omega$,

$$E_x(T_1) = E_x(T_1 | \sigma \in \Omega) \leq t_{hit}.$$ 

Now we ask what is $E_x[T_2 - T_1]$?

If we hit $\sigma(2)$ before $\sigma(1)$, then $T_2 - T_1 = 0$.

Else, by similar reasoning as before, $E[T_2 - T_1] = E[\gamma_{\sigma(1)}] \leq t_{hit}$.

Let $r, s \in \Omega$, and WLOG, assume our chain hits $r$ before $s$.

Then

$$E[T_2 - T_1 | \{\sigma(1), \sigma(2)\} = \{r, s\}] = \frac{1}{2} E[T_2 - T_1 | \sigma(1) = r] + \frac{1}{2} E[T_2 - T_1 | \sigma(1) = s]$$

$$\leq \frac{1}{2} t_{hit} + \frac{1}{2} 0$$

Since we chose $r$ uniformly at random.

Again, by exercise 11.1a, this shows that $E[T_2 - T_1] \leq \frac{1}{2} t_{hit}$.

So $E[T_2] = E[T_1] + E[T_2 - T_1] \leq t_{hit}(1 + \frac{1}{2})$. 

What is $E[T_3 - T_2]$?

If we hit $\sigma(3)$ before $\sigma(1)$ or $\sigma(2)$ then $T_3 - T_2 = 0$.
Else, $E[T_3 - T_2] = E_{\sigma(2), \sigma(3)} \mathbb{1}_{\sigma(3) \neq \sigma(1)} \leq t_{hit}$

Let $g, r, s \in \mathbb{Z}$ and WLOG assume we hit $g$ then $r$ then $s$.

$E(T_3 - T_2 | \sigma(1), \sigma(2), \sigma(3) = \{g,r,s\})$

$= \frac{1}{3} E(T_3 - T_2 | \sigma(3) = s, \sigma(1), \sigma(2) = \{g,r\}) + \frac{2}{3} E(T_3 - T_2 | \sigma(3) = s, \sigma(1), \sigma(2) = \{g,r,s\})$

$= \frac{1}{3} \cdot t_{hit} + \frac{2}{3} \cdot 0$

So $E[T_3] = E[T_1] + E[T_2 - T_1] + E[T_3 - T_2]$

$= t_{hit} \left( 1 + \frac{1}{2} + \frac{1}{3} \right)$

by induction, we can show that $E[T_n] = t_{hit} \left( 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \right)$

Example 2: Random walk on a complete graph with self loops

For all states $x, y \in \mathbb{Z}$, $E_X \mathbb{1}_Y = t_{hit}$

so in the previous proof we can replace all $\mathbb{1}_z \leq \mathbb{1}_w =$, so our bound is tight.
Lower Bounds

Let $A \subseteq \mathcal{L}$ and $t_{\min}^A = \min_{a \in \mathcal{A}} E_a X_b$.

Then $t_{\text{cov}} \leq t_{\text{cov}}^A$ (it takes at least as long to cover $\mathcal{L}$ as it does to cover $A$).

and by a similar coupon collecting argument, we can show

$t_{\text{cov}} \geq t_{\text{cov}}^A \geq t_{\min}^A \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{|A|-1}\right)$

which gives us a lower bound for each set $A \subseteq \mathcal{L}$. There is some cleverness required for choosing $A$ because if you make it too small then $(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{|A|-1})$ will be small, but if you make $A$ big, then you will be able to minimize $t_{\min}^A$ more.
Random walk on binary trees.

Which nodes maximize \( t_{\text{hit}} \)?

2 nodes who's only common ancestor is the root.

The expected time to get between these two nodes is the same as the time to get from a leaf, to the root, and back to that same leaf. This is the commute time from the root to the leaf.

Recall: The commute time, the time it takes to get from state \( a \) to state \( b \) and back to state \( a \) is

\[
E_a(\tau_{ab}) = C_a \cdot R(a \leftrightarrow b) \quad (\text{Prop 10.6})
\]

So

\[
E_a(\tau_{\text{root, root}}) = 2(\# \text{edges}) \cdot (\text{depth}) = 2(n-1)k
\]

so \( t_{\text{cov}} \leq 2(n-1)k \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n} \right) = (2+o(1)) (n-1) k \cdot \log 2^{k-1}

= (2+o(1)) (n-1) k \cdot \log 2

= (2+o(1)) \cdot \log 2 \cdot nk^2.

To get a lower bound, we need to choose a good \( A \). we don't want \( A = \{ \text{leaves} \} \) because then we will get a small \( t_{\text{min}} \) by choosing two leaves with the same parent.

Instead, let \( h \) be some height, and let \( A \) be a set of \( 2^h \) leaves, each with a unique ancestor at height \( h \). We will decide how to choose \( h \) later.

The closest two vertices have a common ancestor at height \( h-1 \), so the commute time identity gives

\[
t_{\text{min}} = 2(n-1) - (k - (h-1))
\]
again, \( t_{min}^A = 2(n-1)(k-h+1) \), and \( |A| = 2^n \)

so \( t_{cov} \geq 2(n-1)(k-h+1) \left( 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^{h-1}} \right) \)

\[ = (2 + o(1)) \cdot n \cdot (k-h) \cdot h \cdot \log_2 \]

which is maximized when \( h = \frac{k}{2} \). Then

\[ t_{cov} \geq \frac{1}{4} (2 + o(1)) \left( \log_2 \right) n k^2 \]

(our upper bound was the same, without the \( \frac{1}{4} \))
eq 4: Random walk on d dimensional torus

\textbf{Case 1: } \text{d} \geq 3

\textbf{Recall:} (prop 10.13) There exist constants \(c_d, C_d\) such that

\[ c_d n^d \leq \mathbb{E}_x (\zeta_y) \leq C_d n^d \]

These bounds do not depend on \(x\) and \(y\).

Thus

\[ t_{\text{cov}} \leq C_d n^d \cdot (1 + \frac{1}{2} + \ldots + \frac{1}{n^d}) = C_d d (\log n) n^d (1 + o(1)) \]

Taking

\[ A = \mathbb{Z}_n^d \] (the entire torus)

Then

\[ t_{\text{cov}} \geq C_d n^d \cdot (1 + \frac{1}{2} + \ldots + \frac{1}{n^d}) = C_d d (\log n) n^d (1 + o(1)) \]

\textbf{Case 2: } \text{d} = 2

\textbf{Recall:} (prop 10.13) If \(x, y \in \mathbb{Z}_n^d\) are distance \(k\) apart,

\[ c_2 n^2 \cdot \log k \leq \mathbb{E}_x \zeta_y \leq C_2 n^2 \cdot \log k \]

so since the max \(k\) can be is \(n\)

\[ t_{\text{cov}} \leq C_2 n^2 \cdot \log n \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n^2} \right) = (1 + o(1)) C_2 n^2 \cdot 2 \cdot (\log n)^2 \]

for lower bound, Assume \(n\) is a perfect square, and let \(A = \{ \text{coordinates that are multiples of } \mathbb{Z}_n^2 \} \)

So \(|A| = n^2\), and

\[ t_{\text{min}} \geq C_2 n^2 \cdot \log n \]

So

\[ t_{\text{cov}} \geq C_2 n^2 \cdot \log n \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n^2} \right) = (1 + o(1)) C_2 n^2 \cdot \frac{1}{2} \cdot (\log n)^2 \]

= \( (1 + o(1)) C_2 \cdot \frac{1}{2} \cdot n^2 (\log n)^2 \)