# Kruskal's Principle and Collision Time for Monotone Transitive Walks on the Integers

Ravi Montenegro [*] Prasad Tetali [†]

September 7, 2014

**Abstract**

A set of positive generators $S \subset \mathbb{N}$ with a probability distribution $p$ on $S$ induces a monotone transitive walk on the integers $\mathbb{Z}$, with $\mathsf{P}(x, x+s) = p(s)$. Kruskal's principle observes that when two independent copies of the walk are started from nearby states then, with high probability, they do not have to travel far before visiting a common state ("collision distance"). We develop tools for determining the expected collision distance and the probability of collision within a certain distance. We then derive bounds in terms of "collision time"; These are used to prove that Pollard's Kangaroo method solves the discrete logarithm problem $g^x = h$ on a cyclic group in expected time $(2 + o(1))\sqrt{N}$, when $x$ is in an interval $[a, b]$ of size $N = b - a + 1$. We also resolve a conjecture of Pollard's by showing that the same bound holds, when step sizes are generalized from powers of 2 to powers of any fixed $n$.

## 1 Introduction

Probabilistic "paradoxes" can have unexpected applications in computational problems, but mathematical tools often do not exist to prove the reliability of the resulting computations, so instead practitioners have to rely on heuristics, intuition and experience. A case in point is Kruskal's Principle, also known as the Kruskal Count, a probabilistic observation reported by Martin Kruskal and popularized in a card trick by Martin Gardner. In a 1978 paper John Pollard applied the same trick to a mathematical problem related to code breaking, the Discrete Logarithm Problem: solve for the exponent $x$, given the generator $g$ of a cyclic group $G$ and an element $h \in G$ such that $g^x = h$. Variations on this important algorithm are still the fastest methods for breaking many codes, including instances of the Diffie-Hellman key exchange, ElGamal cryptosystem, the US government's DSA (Digital Signature Algorithm), and the Bitcoin protocol.

Pollard's Kangaroo method involves running two independent instances of the same random walks on a cyclic group $G$, one starting at a state $g^y$ with known exponent $Y_0 = y$ (the "tame kangaroo") and the other starting at state $h = g^x$ with unknown exponent $X_0 = x$ (the "wild kangaroo"), with the walks terminating after their first collision $g^{X_i} = g^{Y_j}$. If it is known that $|X_0 - Y_0| \leq N$ for some $N$ then this can be designed to have runtime order $\sqrt{N}$. This exploits the property that two independent instances of an increasing additive walk, $Z \to Z + s$ where $s \in S \subset \mathbb{N}$, will visit some common state $X_i = Y_j$ ("collide") fairly quickly when started at nearby

[*] Department of Mathematical Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA. Email: ravi_montenegro@uml.edu; research supported in part by JSPS fellowship L-13516 while in residence at Kyushu University.

[†] School of Mathematics and College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: tetali@math.gatech.edu; research supported in part by NSF grants DMS 0401239, 0701043.

states. Kruskal's Principle exploits the related property that with high probability the walks will collide without having traveled very far. (See Section 5.1 for a precise description of the Kangaroo method.)

Past work examining collision of walks seems to be of little help in understanding Pollard's Kangaroo method. Lagarias et.al. [3] study Kruskal's Principle, but only for uniform and geometric walks for which the expected number of steps until an intersection ("collision time") is easy to determine. Pollard used similar methods to study the collision time of the kangaroo method, but his results involved recurrence relations which can only be solved numerically on a case-by-case basis [6]. In contrast, we develop a method for showing explicit bounds on *expected collision time, expected collision distance*, and Kruskal's Principle which are asymptotically quite precise.

Our approach to these questions is motivated by our past work on Pollard's Rho algorithm for Discrete Logarithm, a problem which concerns self-intersection of a single walk [2]. However, a key part of that analysis involved examining transitions after the *mixing time* until a near stationary distribution has been reached, whereas in the kangaroo walk when $N \ll |G|$, the algorithm may terminate long before the walk has mixed. Instead we develop a notion of *mixing distance* $\mathcal{D}(\epsilon)$, the distance a walk on $\mathbb{Z}$ must travel until it has near uniform probability of hitting each subsequent state. A first moment argument leads to results on Kruskal's Principle and expected collision distance, with collision time following as a corollary.

An exact statement of our results involves a fair amount of notation, and so we leave it until later. However, a heuristic is instructive. Suppose the two walks have transitions $Z \to Z + s$, with $z \in S$ drawn from a distribution $p : S \to [0, 1]$ and average step size is $\bar{S} = \sum_{s \in S} p(s) \cdot s$. Each process visits a $1/\bar{S}$ fraction of states. So each state has probability $(1/\bar{S})^2$ of being visited by both walks. If we ignore dependencies between nearby states then we expect the walks to travel distance $\bar{S}^2$ before colliding ("collision distance"). Likewise, each step of the $X_i$ process has probability $1/\bar{S}$ of being visited by the $Y_j$ process, so ignoring dependencies then the two will collide after an average of $\bar{S}$ steps of the $X_i$ process, and by symmetry the same holds for the $Y_j$ process. We show that accounting for dependencies introduces a correction factor of $(1 + B_\epsilon)$, where $B_\epsilon$ is the maximum expected number of collisions between two *independent* processes starting at a nearby state and proceeding for the mixing distance $\mathcal{D}(\epsilon)$.

In the specific case of the walk involved in Pollard's Kangaroo method, our upper and lower bounds even agree on their lead constants, which is quite rare among the analysis of algorithms motivated by Markov chains. More specifically we have:

**Theorem 1.1.** *Consider the interval discrete logarithm problem on cyclic group $G$: Solve for unknown $x$ when $h = g^x$ for $g, h \in G$ and $x$ uniform in interval $[a, b]$. The Distinguished Points implementation of Pollard's Kangaroo method with powers of 2 step sizes has expected run time*

$$(2 + o(1))\sqrt{b - a}\,.$$

*If $x$ is not known to be uniform then the expected number of group operations is upper bounded by*

$$(3 + o(1))\sqrt{b - a}\,,$$

*with equality when $x = a$ or $x = b$.*

The closest previous result is due to Pollard [6] who gave a convincing but not completely rigorous argument for the first bound. Given the practical significance of Pollard's Kangaroo method for solving the discrete logarithm problem, we find it surprising that there has been no fully rigorous analysis of this algorithm, particularly since it has been over 30 years since it was first proposed in [5].

The paper proceeds as follows. In Section 2 we study the expected collision distance and the Kruskal count probability of collision within some distance $d$. In Section 3 we show how to bound quantities appearing in Section 2, and describe a process for calculating them in Section 3.3. The results are extended to collision time in Section 4. In Section 5 we apply the results of Section 3.3 and 4 to show sharp results for Pollard's Kangaroo method with step sizes of powers of two. The paper finishes in Section 6 with an extension of the results to powers of an arbitrary $n$, resolving a conjecture of Pollard's.

## 2 Collision Distance

Consider a Markov Chain on $\mathbb{Z}$ which is increasing and transitive:

$$\forall x, s \in \mathbb{Z} : \mathsf{P}(x, x+s) = \mathsf{P}(0, s) > 0\,, \quad \text{only if} \quad s > 0\,.$$

If the walk is aperiodic, i.e. $gcd(\{s : \mathsf{P}(0,s) > 0\}) = 1$, then two independent instances $\{X_i\}$ and $\{Y_j\}$ of the walk will collide eventually, i.e. $\exists i, j : X_i = Y_j$. It is natural to ask how far the walks will travel before colliding, and how many steps this will take, known as *collision distance* and *collision time* respectively. The probability that a collision occurs within some distance $d$ is sometimes known as the Kruskal Count, as it relates to a card trick popularized by Martin Gardner and credited to Martin Kruskal.

In this section we consider two forms of this question: the probability that a collision will occur within distance $d$, and the expected distance until the collision.

We first require some notation. Let $S = \{s > 0 : \mathsf{P}(0,s) > 0\} \subseteq \mathbb{N} \setminus \{0\}$ and $p(s) = \mathsf{P}(0,s)$ be a probability distribution on $S$. Then $(p, S)$ generates the original walk on $\mathbb{Z}$, with $\mathsf{P}(x,y) = \mathsf{P}(0, y-x) = p(y-x)$. Let $\bar{S} = \sum_{s\in S} s\, p(s)$ denote the average step size, and $S_{max} = \max_{s\in S} s$ be the largest step size. Asymptotically, each state has probability $1/\bar{S}$ of being hit by the $X_i$ walk:

$$\lim_{d\to\infty} \mathsf{Pr}\left(\exists i : X_i = X_0 + d\right) = 1/\bar{S}\,. \tag{1}$$

It is natural to ask the distance required to closely approximate this limiting behavior.

**Definition 2.1.** The *mixing distance* $\mathcal{D}(\epsilon)$, for $\epsilon \geq 0$, is the smallest integer such that

$$\forall d \geq \mathcal{D}(\epsilon) : \frac{1-\epsilon}{\bar{S}} \leq \mathsf{Pr}\left(\exists i,\, X_i = d \mid X_0 = 0\right) \leq \frac{1+\epsilon}{\bar{S}}\,.$$

If $d \geq \mathcal{D}(\epsilon)$ and $Y_0 < X_0$ then each state $X_0 + d$ has probability $\approx 1/\bar{S}$ of being hit by $X$ and an independent $\approx 1/\bar{S}$ chance of being hit by $Y$, so that the probability that both walks hit the state is $\approx 1/\bar{S}^2$. If we ignore correlations then a collision is expected within distance $\mathcal{D}(\epsilon) + \bar{S}^2$.

The frequency of collisions before distance $\mathcal{D}(\epsilon)$ will vary depending on the Markov chain, and so we quantify it separately. Let $B_\epsilon$ be the worst-case expected number of collisions between two independent walks before they travel distance $\mathcal{D}(\epsilon)$.

$$B_\epsilon = \max_{Y_0 \leq X_0 = 0} \mathbb{E} \sum_{d=0}^{\mathcal{D}(\epsilon)-1} \mathbf{1}_{\{\exists i,j\,:\,(X_i = Y_j = d)\wedge((i,j)\neq(0,0))\}}\,.$$

The main result of this section is to prove that when $\epsilon$ and $B_\epsilon$ are close to zero and $\mathcal{D}(\epsilon), S_{max} \ll \bar{S}^2$ then the walks travel an expected distance of $(1 + o(1))\bar{S}^2$ until a collision. More precisely,

**Theorem 2.2.** *Consider an increasing additive Markov chain on $\mathbb{Z}$ with generating set $S$, probability distribution $p : S \to [0,1]$, and transitions $\mathsf{P}(x, x+s) = p(s)$. Run two independent instances of the walk with starting states $Y_0 < X_0 = 0$. The expected distance the walks travel until a collision is:*

$$\begin{aligned}
\mathbb{E}\min\{d : \exists i,j,\, X_i = Y_j = d\} &\leq \bar{S}^2 \left( \frac{\sqrt{1+B_\epsilon} + \frac{1}{\bar{S}}\sqrt{\mathcal{D}(\epsilon) + S_{max}}}{1-2\epsilon} \right)^2 \\
\mathbb{E}\min\{d : \exists i,j,\, X_i = Y_j = d\} &\geq \bar{S}^2 \left( \frac{\max\{0,\, 1-\sqrt{B_\epsilon}\}}{1+\epsilon} \right)^2 .
\end{aligned}$$

*The probability there is no collision within distance $d$ is:*

$$\begin{aligned}
\mathsf{Pr}\left(\nexists i,j,\, X_i = Y_j \leq d\right) &\leq \exp\left( -d \Big/ \bar{S}^2 \left( \frac{\sqrt{1+B_\epsilon} + \frac{1}{\bar{S}}\sqrt{\mathcal{D}(\epsilon) + S_{max}}}{1-2\epsilon} \right)^2 \right) \\
\mathsf{Pr}\left(\nexists i,j,\, X_i = Y_j \leq d\right) &\geq \exp\left( -d \Big/ \bar{S}^2 \left( \frac{\max\{0,\, 1-\sqrt{B_\epsilon}\}}{1+\epsilon} \right)^2 \right) .
\end{aligned}$$

## 2.1 The proof

Roughly speaking our argument is as follows: break the walk into blocks covering some distance $\mathcal{D}(\epsilon) + \Delta$ each, determine the probability of no collision on a block, and then raise this to the $\ell^{th}$ power to find the probability of no collision within $\ell$ blocks. Boundary effects complicate this a little.

Let $I_D^\Delta$ count collisions with $X_i = Y_j \in [X_0 + D, X_0 + D + \Delta)$, so that

$$I_D^\Delta = \sum_{d=D}^{D+\Delta-1} \mathbf{1}_{\{\exists i,j,\, X_i = Y_j = X_0 + d\}} .$$

Our goal is to study $\mathbb{E}\min\{d : I_0^d > 0\}$. It is natural to expect that if $\mathbb{E}I_0^d$ is large then $\mathsf{Pr}\left(I_0^d > 0\right)$ is going to be large as well. We use a first moment argument to show this. In particular, a non-negative random variable $Z \geq 0$ satisfies

$$\mathbb{E}[Z] = \mathsf{Pr}\left(Z = 0\right)\, \mathbb{E}[Z \mid Z = 0] + \mathsf{Pr}\left(Z > 0\right)\, \mathbb{E}[Z \mid Z > 0] ,$$

and so

$$\mathsf{Pr}\left(Z > 0\right) = \frac{\mathbb{E}[Z]}{\mathbb{E}[Z \mid Z > 0]} .$$

We apply this with $Z = I_0^{\mathcal{D}(\epsilon)+\Delta}$ to obtain the following:

**Lemma 2.3.** *Under the conditions of Theorem 2.2, if $\epsilon \leq 1/2$ and $\Delta \geq 0$ then*

$$\begin{aligned}
\mathsf{Pr}\left(I_0^{\mathcal{D}(\epsilon)+\Delta} > 0 \mid Y_0 < X_0 = 0\right) &\leq B_\epsilon + \frac{\Delta}{\bar{S}^2}(1+\epsilon)^2 \\
\mathsf{Pr}\left(I_0^{\mathcal{D}(\epsilon)+\Delta} > 0 \mid Y_0 < X_0 = 0\right) &\geq \frac{\Delta}{\bar{S}^2} \frac{(1-2\epsilon)^2}{1 + B_\epsilon + \frac{\Delta}{\bar{S}^2}} .
\end{aligned}$$

To show this we first need bounds on $\mathbb{E}[Z]$ and $\mathbb{E}[Z \mid Z > 0]$. It is easier to consider $I_{\mathcal{D}(\epsilon)}^\Delta$ and fill in the omitted intersections $I_0^{\mathcal{D}(\epsilon)}$ later.

**Lemma 2.4.** *Under the conditions of Theorem 2.2, if* $\Delta \geq 0$ *then*

$$\begin{aligned} (1-\epsilon)^2\,\frac{\Delta}{\bar{S}^2} \leq \quad & \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid Y_0 < X_0 = 0\right] && \leq (1+\epsilon)^2\,\frac{\Delta}{\bar{S}^2} \\ & \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid I^{\Delta}_{\mathcal{D}(\epsilon)} > 0,\, Y_0 < X_0 = 0\right] && \leq 1 + B_\epsilon + \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid X_0 = Y_0 = 0\right] . \end{aligned}$$

*Proof of Lemma 2.4.* The walks $X_i$ and $Y_j$ are independent and so if $d \geq \mathcal{D}(\epsilon)$ then

$$\begin{aligned} & \mathsf{Pr}\left(\exists i,j,\, X_i = Y_j = d \mid Y_0 < X_0 = 0\right) \\ = \quad & \mathsf{Pr}\left(\exists i,\, X_i = d \mid X_0 = 0\right) \mathsf{Pr}\left(\exists j,\, Y_j = d \mid Y_0 < 0\right) \\ \geq \quad & \left(\frac{1-\epsilon}{\bar{S}}\right)^2 . \end{aligned}$$

The expectation $\mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid Y_0 < X_0 = 0\right]$ then satisfies

$$\begin{aligned} \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid Y_0 < X_0 = 0\right] &= \mathbb{E}\left[\sum_{d=\mathcal{D}(\epsilon)}^{\mathcal{D}(\epsilon)+\Delta-1} \mathbf{1}_{\{\exists i,j,\, X_i = Y_j = d\}} \mid Y_0 < X_0 = 0\right] \\ &= \sum_{d=\mathcal{D}(\epsilon)}^{\mathcal{D}(\epsilon)+\Delta-1} \mathsf{Pr}\left(\exists i,j,\, X_i = Y_j = d \mid Y_0 < X_0 = 0\right) \\ &\geq \Delta\left(\frac{1-\epsilon}{\bar{S}}\right)^2 . \end{aligned}$$

The upper bound on $\mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid Y_0 < X_0 = 0\right]$ is similar.

Now consider the conditional case of $\mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid I^{\Delta}_{\mathcal{D}(\epsilon)} > 0,\, Y_0 < X_0 = 0\right]$. By transitivity, the number of intersections after the first collision is at most the number if $X_0 = Y_0 = 0$.

$$\begin{aligned} & \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid I^{\Delta}_{\mathcal{D}(\epsilon)} > 0,\, Y_0 < X_0 = 0\right] \\ \leq \quad & 1 + \mathbb{E}\left[\sum_{d=1}^{\Delta-1} \mathbf{1}_{\{\exists i,j,\, X_i = Y_j = d\}} \mid X_0 = Y_0 = 0\right] \\ \leq \quad & 1 + B_\epsilon + \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid X_0 = Y_0 = 0\right] . \end{aligned}$$

□

*Proof of Lemma 2.3.* As before, let $Z = I_0^{\mathcal{D}(\epsilon)+\Delta}$, so that

$$\begin{aligned} \mathsf{Pr}\left(I_0^{\mathcal{D}(\epsilon)+\Delta} > 0 \mid Y_0 < X_0 = 0\right) &\geq \mathsf{Pr}\left(I^{\Delta}_{\mathcal{D}(\epsilon)} > 0 \mid Y_0 < X_0 = 0\right) \\ &\geq \frac{\mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid Y_0 < X_0 = 0\right]}{1 + B_\epsilon + \mathbb{E}\left[I^{\Delta}_{\mathcal{D}(\epsilon)} \mid X_0 = Y_0 = 0\right]} \\ &\geq \frac{(1-\epsilon)^2\Delta/\bar{S}^2}{1 + B_\epsilon + (1+\epsilon)^2\Delta/\bar{S}^2} \geq \left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \frac{\Delta/\bar{S}^2}{\frac{1+B_\epsilon}{(1+\epsilon)^2} + \Delta/\bar{S}^2} . \end{aligned}$$

For the upper bound

$$\begin{aligned}
\mathsf{Pr}\left(I_0^{\mathcal{D}(\epsilon)+\Delta} > 0 \mid Y_0 < X_0 = 0\right) &\leq \mathbb{E}\left[I_0^{\mathcal{D}(\epsilon)+\Delta} \mid Y_0 < X_0 = 0\right] \\
&= \mathbb{E}\left[I_0^{\mathcal{D}(\epsilon)} + I_{\mathcal{D}(\epsilon)}^{\Delta} \mid Y_0 < X_0 = 0\right] \\
&\leq B_\epsilon + \Delta\,\frac{(1+\epsilon)^2}{\bar{S}^2}\,.
\end{aligned}$$

□

Lemmas 2.4 and 2.3 will now be combined to prove the theorem.

*Proof of Theorem 2.2.* Recall from our sketch at the beginning of the section that we first break the walk into segments covering some distance $\mathcal{D}(\epsilon)+\Delta$ each. Let $D_0 = 0$, let $\mathcal{D}(\epsilon)$ be the mixing distance, and set $D_k = D_{k-1} + \mathcal{D}(\epsilon) + \Delta$. Also, let $X_0^{(k)}$ denote the first $X_i \geq D_k$, and likewise $Y_0^{(k)}$ is the first $Y_j \geq D_k$, with $X_0^{(0)} = X_0$ and $Y_0^{(0)} = Y_0$. Let $F_k$ denote the event that these are distinct, i.e. $X_0^{(k)} \neq Y_0^{(k)}$. Then

$$\begin{aligned}
\mathsf{Pr}\left(I_0^{D_\ell} = 0 \mid F_0\right) &= \prod_{k=0}^{\ell-2} \mathsf{Pr}\left(I_{D_k}^{\mathcal{D}(\epsilon)+\Delta} = 0 \wedge F_{k+1} \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad \times \mathsf{Pr}\left(I_{D_{\ell-1}}^{\mathcal{D}(\epsilon)+\Delta} = 0 \mid I_0^{D_{\ell-1}} = 0 \wedge F_{\ell-1}\right)\,.
\end{aligned}$$

When $F_k = 0$ then assume that $X_0^{(k)} > Y_0^{(k)}$, and if not then exchange the labels of $X$ and $Y$. By Lemma 2.3, with $\Delta$ replaced by $\Delta + S_{max}$,

$$\begin{aligned}
&\mathsf{Pr}\left(I_{D_k}^{\mathcal{D}(\epsilon)+\Delta} = 0 \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\geq \mathsf{Pr}\left(I_{D_k}^{\mathcal{D}(\epsilon)+\Delta} = 0 \wedge F_{k+1} \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\geq \mathsf{Pr}\left(I_{X_0^{(k)}}^{\mathcal{D}(\epsilon)+\Delta+S_{max}} = 0 \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\geq 1 - B_\epsilon - \frac{\Delta + S_{max}}{\bar{S}^2}\,(1+\epsilon)^2\,.
\end{aligned}$$

In the other direction,

$$\begin{aligned}
&\mathsf{Pr}\left(I_{D_k}^{\mathcal{D}(\epsilon)+\Delta} = 0 \wedge F_{k+1} \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\leq \mathsf{Pr}\left(I_{D_k}^{\mathcal{D}(\epsilon)+\Delta} = 0 \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\leq \mathsf{Pr}\left(I_{X_0^{(k)}}^{\mathcal{D}(\epsilon)+\Delta} = 0 \mid I_0^{D_k} = 0 \wedge F_k\right) \\
&\quad\leq 1 - \frac{\Delta - S_{max}}{\bar{S}^2}\,\frac{(1-2\epsilon)^2}{1 + B_\epsilon + \frac{\Delta - S_{max}}{\bar{S}^2}}\,.
\end{aligned}$$

From this we obtain the probability of no collision in $\ell$ segments:

$$\left(1 - B_\epsilon - \frac{\Delta + S_{max}}{\bar{S}^2}\,(1+\epsilon)^2\right)^\ell \leq \mathsf{Pr}\left(I_0^{D_\ell} = 0\right) \leq \left(1 - \frac{\Delta - S_{max}}{\bar{S}^2}\,\frac{(1-2\epsilon)^2}{1 + B_\epsilon + \frac{\Delta - S_{max}}{\bar{S}^2}}\right)^\ell\,. \tag{2}$$

To bound Kruskal count observe that, if $E_0 \subseteq E_1 \subseteq E_2 \subseteq \cdots$ is an increasing family of events, then $\mathbb{E} \min\{d : E_d\} = \sum_{d=0}^{\infty} \mathsf{Pr}\left(\neg E_d\right)$, where we let $E_d = \{I_0^d > 0\}$ be the event that a collision occurs within distance $d$.

$$\begin{aligned}
\mathbb{E} \min\{d : I_0^d > 0\} &= \sum_{d=0}^{\infty} \mathsf{Pr}\left(I_0^d = 0\right) \leq \sum_{\ell=0}^{\infty} (\mathcal{D}(\epsilon) + \Delta)\, \mathsf{Pr}\left(I_0^{D_\ell} = 0\right) \\
&\leq (\mathcal{D}(\epsilon) + \Delta) \sum_{\ell=0}^{\infty} \left(1 - \frac{\Delta - S_{max}}{\bar{S}^2} \frac{(1-2\epsilon)^2}{1 + B_\epsilon + \frac{\Delta - S_{max}}{\bar{S}^2}}\right)^{\ell} \\
&= (\mathcal{D}(\epsilon) + \Delta) \frac{\bar{S}^2}{\Delta - S_{max}} \frac{1 + B_\epsilon + \frac{\Delta - S_{max}}{\bar{S}^2}}{(1-\epsilon)^4} \,. \qquad (3)
\end{aligned}$$

This is minimized when $\Delta - S_{max} = \bar{S} \sqrt{(\mathcal{D}(\epsilon) + S_{max})(1 + B_\epsilon)}$. Substituting this into (3) gives the upper bound on expected run time. Substituting this into (2) and using the relation $1 - x \leq e^{-x}$ gives the upper bound on probability of no collision.

The lower bound is similar.

$$\begin{aligned}
\mathbb{E} \min\{d : I_0^d > 0\} &= \sum_{d=0}^{\infty} \mathsf{Pr}\left(I_0^d = 0\right) \geq \sum_{\ell=1}^{\infty} (\mathcal{D}(\epsilon) + \Delta)\, \mathsf{Pr}\left(I_0^{D_\ell} = 0\right) \\
&\geq (\mathcal{D}(\epsilon) + \Delta) \sum_{\ell=1}^{\infty} \left(1 - B_\epsilon - \frac{\Delta + S_{max}}{\bar{S}^2} (1+\epsilon)^2\right)^{\ell} \\
&= (\mathcal{D}(\epsilon) + \Delta) \left(\frac{1}{B_\epsilon + \frac{\Delta + S_{max}}{\bar{S}^2} (1+\epsilon)^2} - 1\right) . \qquad (4)
\end{aligned}$$

This is maximized when $\Delta + S_{max} = \frac{\bar{S}^2}{(1+\epsilon)^2} \left(\sqrt{B_\epsilon - \frac{\mathcal{D}(\epsilon) - S_{max}}{\bar{S}^2}(1+\epsilon)^2} - B_\epsilon\right)$. Substituting this into (4) gives a lower bound, after some straightforward simplification. Substituting this into (2) and using the relation: $1 - x \geq e^{-x/(1-x)}$ (valid for all $x \in [0, 1)$)), gives the lower bound on probability of no collision. □

## 3 Calculating Collision Distance

The bounds of Theorem 2.2 require $\mathcal{D}(\epsilon)$ and $B_\epsilon$. We develop tools for bounding these in Sections 3.1 and 3.2 respectively. This culminates in Section 3.3, where we develop a step-by-step process for applying Theorem 2.2. We illustrate its application in Section 3.4 by examining the uniform walk.

### 3.1 Mixing Distance

In order to understand our approach it is helpful to begin with some asymptotic properties of the walk.

Recall that asymptotically each state has probability $1/\bar{S}$ of being hit by the $X_i$ walk, so that $\lim_{D \to \infty} \mathsf{Pr}\left(\exists j : X_j = X_0 + D\right) = 1/\bar{S}$. The probability that the first $X_j \geq X_0 + \mathcal{D}$ is given by $X_j = X_0 + \mathcal{D} + \delta$ will be

$$F(\mathcal{D}, \delta) = \sum_{s \in S,\, s > \delta} \mathsf{Pr}\left(\exists j : X_j = X_0 + \mathcal{D} + \delta - s\right) p(s) \,.$$

In the limit this approaches

$$F(\delta) = \lim_{\mathcal{D}\to\infty} F(\mathcal{D},\delta) = \frac{1}{\bar{S}} \sum_{s\in S,\, s>\delta} p(s)\,. \tag{5}$$

It follows that if $\mathcal{T}$ satisfies $\mathsf{Pr}\left(X_{\mathcal{T}} = X_0 + \mathcal{D} + \delta\right) = F(\delta)$ for some $\mathcal{D}$ and all $\delta \in [0, S_{max})$ then

$$\forall d \geq \mathcal{D} :\ \mathsf{Pr}\left(\exists j :\ X_j = X_0 + d\right) = 1/\bar{S}\,. \tag{6}$$

It will almost never be the case that $\mathsf{Pr}\left(X_{\mathcal{T}} = X_0 + \mathcal{D} + \delta\right) = F(\delta)$ holds for fixed values of $\mathcal{D}$ and $\mathcal{T}$. However, the relation still holds if $\mathcal{D}$ and $\mathcal{T}$ are random variables, as long as $\mathcal{D}$ is independent of $\delta$. We show that (6) can be made rigorous when $\mathcal{T}$ is a stopping time:

**Definition 3.1.** A *stopping time* for a random walk $\{X_i\}_{i=0}^{\infty}$ is a random variable $\mathcal{T} \in \mathbb{N}$ such that the event $\{\mathcal{T} = t\}$ depends only on $X_0,\, X_1,\, \ldots,\, X_t$.

**Lemma 3.2.** *Consider a stopping time $\mathcal{T}$ and associated random variable $\mathcal{D}$ such that*

$$\forall \delta \geq 0 :\ \mathsf{Pr}\left(X_{\mathcal{T}} = X_0 + \mathcal{D} + \delta\right) = F(\delta) = \frac{1}{\bar{S}} \sum_{s\in S,\, s>\delta} p(s)\,.$$

*Then*

$$\forall d \geq \mathcal{D} :\ \mathsf{Pr}\left(\exists j \geq \mathcal{T} :\ X_j = X_0 + d\right) = 1/\bar{S}\,.$$

*Proof.* The proof will be inductive. The base case is when $d = \mathcal{D}$. By assumption

$$\mathsf{Pr}\left(\exists j \geq \mathcal{T},\, X_j = X_0 + d\right) = F(0) = \bar{S}^{-1}\,.$$

When $d > \mathcal{D}$ then by induction assume $\forall c \in [\mathcal{D}, d) :\ \mathsf{Pr}\left(\exists j \geq \mathcal{T},\, X_j = X_0 + c\right) = \bar{S}^{-1}$. Then

$$\begin{aligned}
&\mathsf{Pr}\left(\exists j \geq \mathcal{T},\, X_j = X_0 + d\right) \\
&\quad= \ F(d-\mathcal{D}) + \sum_{c\in[\mathcal{D},d)} \mathsf{Pr}\left(\exists t \geq \mathcal{T},\, X_t = X_0 + c\right)\, p(d-c) \\
&\quad= \ \frac{\sum_{s>d-\mathcal{D}} p(s)}{\bar{S}} + \sum_{s\leq d-\mathcal{D}} \frac{1}{\bar{S}}\, p(s) = \frac{1}{\bar{S}}\,.
\end{aligned}$$

The lemma follows. □

In particular, if $\mathcal{D}(\epsilon)$ is such that $\mathsf{Pr}\left(\mathcal{D} > \mathcal{D}(\epsilon)\right) \leq \epsilon$ then

$$\forall d \geq \mathcal{D} :\ \frac{1-\epsilon}{\bar{S}} \leq \mathsf{Pr}\left(\exists j \geq \mathcal{T} :\ X_j = X_0 + d\right) \leq \frac{1+\epsilon}{\bar{S}}\,.$$

## 3.2 Bounding $B_\epsilon$

Once we have a bound on some $\mathcal{D}(\epsilon)$ then $B_\epsilon$ needs to be determined. The following lemma reduces this to a problem of showing that the $t$-step transition probabilities decrease quickly in $t$.

**Lemma 3.3.** *If $\mathcal{D}(\epsilon)$ is the mixing distance, $S$ the set of generators for the additive walk, and $N$ is any positive integer then*

$$\frac{1}{|S|} \leq B_\epsilon \leq B_T + \frac{2}{N} \frac{\mathcal{D}(\epsilon)}{\mathcal{D}(\epsilon) - S_{max}}\,,$$

*where*

$$T(\epsilon, N) = \frac{2\mathcal{D}(\epsilon)}{\bar{S}} + \frac{1}{2}\left(\frac{S_{max}}{\bar{S}}\right)^2 \log(\mathcal{D}(\epsilon)N)\,,$$

*and $B_T$, the expected number of collisions before time $T(\epsilon, N)$, satisfies*

$$B_T \le 2 \sum_{i=1}^{T(\epsilon,N)} i \max_v \mathsf{P}^i(0, v)\,.$$

*Proof.* We will use Hoeffding's Inequality, that if $Z$ is the sum of $n$ independent random variables with values in $[a, b]$ then for any $t \ge 0$

$$\mathsf{Pr}\left(|Z - \mathbb{E}Z| \ge t\right) \le \exp\left(\frac{-2t^2}{n(b-a)^2}\right)\,.$$

For simplicity, let $T = T(\epsilon, N)$. Since $X_T - X_0 = \sum_{j=1}^{T} X_j - X_{j-1}$ is the sum of $T$ independent random variables from $S$, with $\mathbb{E}X_T - X_0 = T\bar{S}$, then

$$\begin{aligned}
\mathsf{Pr}\left(X_T < X_0 + \mathcal{D}(\epsilon)\right) &= \mathsf{Pr}\left((X_T - X_0) - T\bar{S} < -(T - \mathcal{D}(\epsilon)/\bar{S})\,\bar{S}\right) \\
&\le \exp\left(\frac{-2(T - \mathcal{D}(\epsilon)/\bar{S})^2(\bar{S}/S_{max})^2}{T}\right) \le \frac{1}{\mathcal{D}(\epsilon)N}\,.
\end{aligned}$$

Likewise, $\mathsf{Pr}\left(Y_T < X_0 + \mathcal{D}(\epsilon)\right) \le \frac{1}{(\mathcal{D}(\epsilon) - S_{max})N}$.

Then

$$\begin{aligned}
B_\epsilon &\le B_T + \mathsf{Pr}\left(X_T < X_0 + \mathcal{D}(\epsilon)\right) \left|\{Y_j : Y_j < X_0 + \mathcal{D}(\epsilon)\}\right| \\
&\quad + \mathsf{Pr}\left(Y_T < X_0 + \mathcal{D}(\epsilon)\right) \left|\{X_i : X_i < X_0 + \mathcal{D}(\epsilon)\}\right| \\
&\le B_T + \frac{1}{\mathcal{D}(\epsilon)N}\left(\mathcal{D}(\epsilon) + S_{max}\right) + \frac{1}{(\mathcal{D}(\epsilon) - S_{max})N}\,\mathcal{D}(\epsilon) \\
&\le B_T + \frac{2\mathcal{D}(\epsilon)}{(\mathcal{D}(\epsilon) - S_{max})N}\,,
\end{aligned}$$

and

$$\begin{aligned}
B_T &= \max_{-S_{max} < Y_0 \le 0 = X_0} \sum_{i=0}^{T-1}\sum_{j=1}^{T-1} \mathsf{Pr}\left(X_i = Y_j\right) \\
&= \max_{-S_{max} < Y_0 \le 0 = X_0} \sum_{i=0}^{T-1}\sum_{j=1}^{T-1}\sum_{w} \mathsf{P}^i(X_0, w)\mathsf{P}^j(Y_0, w) \\
&\le \sum_{i=1}^{T} \max_v \mathsf{P}^i(0, v) \sum_{j=0}^{i}(1 + \mathbf{1}_{\{0<j<i\}}) \sum_w \mathsf{P}^j(0, w) \\
&= \sum_{i=1}^{T} 2i \max_v \mathsf{P}^i(0, v)\,.
\end{aligned}$$

The inequality follows by letting $i$ denote the larger of the two indices and $j$ the smaller, while the final equality is because $\sum_w \mathsf{P}^j(0, w) = 1$.

For the lower bound let $X_0 = Y_0$ so that $B_\epsilon \ge \mathsf{Pr}\left(X_1 = Y_1\right) = \sum_{s\in S} p(s)^2$. By Cauchy-Schwarz

$$1 = \sum_{s\in S} p(s) \times 1 \le \sqrt{\sum_{s\in S} p(s)^2}\sqrt{\sum_{s\in S} 1^2}\,,$$

and so $\sum_{s\in S} p(s)^2 \ge 1/|S|$. Then $B_\epsilon \ge 1/|S|$. □

### 3.3 One approach to bounding collision distance

We combine the results of the previous two subsections to give an approach to applying Theorem 2.2 that seems to work for many problems. The first step is to construct a stopping time and use it to upper bound mixing distance.

(Ia). Construct a "tentative" stopping time $\mathcal{T}_1$ which is uniform on some $[d, d+S_{max})$. The value of $d$ may be a random variable which depends on much of the walk until time $\mathcal{T}_1$.

(Ib). If $X_{\mathcal{T}_1} = X_0 + d + \delta$ then "accept" this with probability

$$\frac{F(\delta)}{F(0)} = \sum_{s>\delta} p(s)\,,$$

where $F(\delta)$ is as in (5). If accepted then set $\mathcal{T} = \mathcal{T}_1$ and $\mathcal{D} = d$. When it is not accepted then start over from $X_{\mathcal{T}_1}$, use the same procedure as in (Ia) to construct a new tentative stopping time which is uniform, again decide whether to accept/reject, and repeat until a stopping time is accepted.

(Ic). How many steps will it take to find the stopping time?

Since $\delta$ is uniform in $[0, S_{max})$, the probability a tentative stopping time is accepted is

$$\begin{aligned}\mathbb{E}\frac{F(\delta)}{F(0)} &= \sum_{c=0}^{S_{max}-1} \mathsf{Pr}\left(\delta = c\right)\frac{F(c)}{F(0)} \\ &= \sum_{\delta=0}^{S_{max}-1} \frac{1}{S_{max}} \sum_{s>c} p(s) \\ &= \frac{\sum_s p(s)s}{S_{max}} = \frac{\bar{S}}{S_{max}}\,.\end{aligned}$$

It follows that

$$\mathsf{Pr}\left(\mathcal{T} = \mathcal{T}_k \mid \mathcal{T} > \mathcal{T}_{k-1}\right) = \mathbb{E}\frac{F(\delta)}{F(0)} = \frac{\bar{S}}{S_{max}}\,.$$

If $M = \frac{S_{max}}{\bar{S}}\log(\bar{S}/\epsilon)$ then

$$\mathsf{Pr}\left(\mathcal{T} > \mathcal{T}_M\right) \le \prod_{k=1}^{M} \mathsf{Pr}\left(\mathcal{T} > \mathcal{T}_k \mid \mathcal{T} > \mathcal{T}_{k-1}\right) \le \frac{\epsilon}{\bar{S}}\,.$$

If $\mathcal{T}_k - \mathcal{T}_{k-1} \le \Delta$ for all $k$ then $\mathcal{T}_M \le \Delta M$, and so $\mathsf{Pr}\left(\mathcal{T} > \Delta M\right) \le \epsilon/\bar{S}$ is bounded.

(Ic'). When $\mathcal{T}_k - \mathcal{T}_{k-1}$ is unbounded then we modify this slightly. Let

$$\Delta(\xi) = \max_k \min\{t :\ \mathsf{Pr}\left(\mathcal{T}_k - \mathcal{T}_{k-1} > t\right) \le \xi\}\,.$$

If $\mathcal{T}_k$ is not determined by time $\mathcal{T}_{k-1} + \Delta(\xi)$ then set $\mathcal{T}_k = \mathcal{T}_{k-1} + \Delta(\xi)$, automatically reject this and construct $\mathcal{T}_{k+1}$ starting from this time. Then $\mathsf{Pr}\left(\mathcal{T} = \mathcal{T}_k \mid \mathcal{T} > \mathcal{T}_{k-1}\right) \ge \bar{S}/S_{max} - \xi$.

In this case, if $M = 2\frac{S_{max}}{\bar{S}}\log(\bar{S}/\epsilon)$ and rounds have length $\mathcal{T}_k - \mathcal{T}_{k-1} \leq \Delta(\bar{S}/2S_{max})$ then

$$\begin{aligned}\mathsf{Pr}\left(\mathcal{T} > \mathcal{T}_M\right) &\leq \prod_{k=1}^{M} \mathsf{Pr}\left(\mathcal{T} > \mathcal{T}_k \mid \mathcal{T} > \mathcal{T}_{k-1}\right) \\ &\leq \left(1 - \frac{\bar{S}}{S_{max}} + \frac{\bar{S}/S_{max}}{2}\right)^{2(S_{max}/\bar{S})\log(\bar{S}/\epsilon)} \leq \frac{\epsilon}{\bar{S}}\,.\end{aligned}$$

(Id). Mixing distance satisfies one of the following:

$$\begin{aligned}(Ic)\ \mathcal{D}(\epsilon) &\leq M\,\Delta\,S_{max} = \Delta\left(\frac{S_{max}}{\bar{S}}\right)^2 \log(\bar{S}/\epsilon)\,\bar{S} \\ (Ic')\ \mathcal{D}(\epsilon) &\leq M\,\Delta(\bar{S}/2S_{max})\,S_{max} = 2\Delta(\bar{S}/2S_{max})\left(\frac{S_{max}}{\bar{S}}\right)^2 \log(\bar{S}/\epsilon)\,\bar{S}\,.\end{aligned}$$

Next, we upper bound collision number $B_\epsilon$:

(IIa). The appropriate value of $T$ is given in Lemma 3.3. It may be simplified using the relations $S_{max}\,p_{min} \leq \bar{S} \leq S_{max}$, where $p_{min} = \min_{s\in S} p(s)$.

(IIb). It is helpful to note that $\max_{u,v} \mathsf{P}^k(u,v)$ is non-increasing in $k$. To see this, suppose that $\max_v \mathsf{P}^k(0,v) \leq C$. By induction, if $i > k$ then

$$\max_v \mathsf{P}^i(0,v) = \max_v \sum_w \mathsf{P}^{i-k}(0,w)\mathsf{P}^k(w,v) \leq C\,\max_v \sum_w \mathsf{P}^{i-k}(0,w) = C\,. \tag{7}$$

(IIc). In particular, if $\max_v \mathsf{P}(0,v) \leq \frac{\alpha}{|S|}$ while $\max_v \mathsf{P}^2(0,v) \leq \beta$ and $\max_v \mathsf{P}^k(0,v) \leq \xi$ for some $k \geq 2$ then

$$\begin{aligned}\frac{1}{|S|} \leq B_\epsilon &\leq \frac{2\alpha}{|S|} + 2\sum_{i=2}^{k-1} i\,\beta + 2\sum_{i=k}^{T} i\,\xi + \frac{2}{N}\frac{\mathcal{D}(\epsilon)}{\mathcal{D}(\epsilon) - S_{max}} \\ &\leq \frac{2\alpha}{|S|} + k^2\,\beta + T^2\,\xi + \frac{2}{N}\frac{\mathcal{D}(\epsilon)}{\mathcal{D}(\epsilon) - S_{max}}\,.\end{aligned} \tag{8}$$

This reduces the problem to one of studying single-step transitions, which are stated in each problem, two-step transitions which are easy enough to compute, and then showing that some moderately sized exponent has $\mathsf{P}^k(0,v) \ll 1/T^2|S|$.

### 3.4 Uniform Distribution

We now examine collision of a walk with uniform transitions $X \to X + i$ with $i \in [1, U]$ for some integer $U \in \mathbb{N}$, so that $\bar{S} = \frac{1+U}{2}$ and $S_{max} = U$. Previously Lagarius et.al. [3] used a coupling method to show that

$$\mathsf{Pr}\left(\nexists i, j,\, X_i = Y_j \leq d\right) \leq \exp\left(-\frac{4}{U^2}\left(1 + o(1)\right)d\right)\,.$$

Haga and Robbins used more direct methods to show this to be an equality [1]. We show this same equality by use of our Theorem 2.2, and also show an equality for expected distance until collision.

Our goal here is not to show a new result but to demonstrate methods we will later apply to the harder Pollard's Kangaroo method.

As discussed earlier, a simple heuristic suggests a collision will occur after an average distance of $\bar{S}^2$ steps. We show this is true by using Theorem 2.2 and the methods of Section 3.3. We use notation (Ia), (Ib), etc to denote step (Ia), (Ib) etc described in Section 3.3.

(Ia) Let $X_0 = 0$. Then $X_1 = X_0 + 1 + \delta$ is uniformly random in $[1, 1+U)$, or equivalently $\delta$ is uniform in $[0, U)$. It follows that $\mathcal{T}_1 = 1$ is a tentative stopping time.

(Ib) The probability we accept this stopping time and set $\mathcal{T} = 1$ and $\mathcal{D} = 1$ is

$$\mathsf{Pr}\left(\mathcal{T} = \mathcal{T}_1 = 1\right) = \frac{\bar{S}}{S_{max}} = \frac{1+U}{2U} > \frac{1}{2}$$

(Ic) More generally, if $M = \log_2(\bar{S}/\epsilon)$ then

$$\mathsf{Pr}\left(\mathcal{T} > \mathcal{T}_M\right) \leq \left(1 - \frac{1}{2}\right)^M \leq \frac{\epsilon}{\bar{S}}$$

(Id) It follows that

$$\mathcal{D}(\epsilon) \leq M\, S_{max} \leq 2\bar{S}\log_2(\bar{S}/\epsilon)$$

Next we study $B_\epsilon$.

(IIa) If $N = \bar{S}$ then in Lemma 3.3

$$\begin{aligned} T(\epsilon, \bar{S}) &= \frac{2\mathcal{D}(\epsilon)}{\bar{S}} + \frac{1}{2}\left(\frac{S_{max}}{\bar{S}}\right)^2 \log(\mathcal{D}(\epsilon)N) \\ &\leq 4\log_2(\bar{S}/\epsilon) + \frac{1}{2}\cdot 4 \cdot \log(2\bar{S}^2\log_2(\bar{S}/\epsilon)) \\ &< 2.2 + 10\log_2(\bar{S}/\epsilon) \end{aligned}$$

(IIb) Obviously $\max_v \mathsf{P}^1(0,v) = 1/U$, and so $\forall i:\ \mathsf{P}^i(0,v) \leq 1/U$ as well.

(IIc) Using just the crude bound $\forall i:\ \mathsf{P}^i(0,v) \leq 1/U$ then and

$$B_{(1/\bar{S}S_{max})} \leq 2\sum_{i=1}^{T} \frac{i}{U} + \frac{4}{U}\left(1 + 1/\log_2(\bar{S}/2\epsilon)\right) = O\left(\frac{\log^2 U}{U}\right)$$

Since $B_\epsilon \geq 1/|S|$ then $B_{(1/\bar{S}S_{max})} = \Theta^*(1/U)$.

We can conclude from Theorem 2.2 that

$$\begin{aligned} \mathbb{E}\min\{d:\ \exists i,j,\ X_i = Y_j = d\} &= (1 + O^*(1/\sqrt{U}))\,\frac{U^2}{4} \\ \mathsf{Pr}\left(\not\exists i,j,\ X_i = Y_j \leq d\right) &= \exp\left(-4d/U^2(1 + O^*(1/\sqrt{U}))\right) \end{aligned}$$

# 4 Collision Times

The Kangaroo walk travels an average distance $\bar{S}$ in each step, and so naturally the expected distance traveled until collision will be roughly $\bar{S}$ times larger than the expected run time. In fact, it is exactly that much larger.

**Corollary 4.1.** *Given an increasing transitive Markov chain on $\mathbb{Z}$ with mixing distance $\mathcal{D}(\epsilon)$, run two independent instances of the walk with starting states $Y_0 < X_0 = 0$. The expected time until a collision is:*

$$\begin{aligned}\mathbb{E}\min\{i : \exists j,\, X_i = Y_j\} &\leq 1 + \bar{S}\left(\frac{\sqrt{1+B_\epsilon} + \frac{1}{\bar{S}}\sqrt{\mathcal{D}(\epsilon) + S_{max}}}{1-2\epsilon}\right)^2 . \\ \mathbb{E}\min\{i : \exists j,\, X_i = Y_j\} &\geq \bar{S}\left(\frac{\max\{0,\, 1-\sqrt{B_\epsilon}\}}{1+\epsilon}\right)^2 .\end{aligned}$$

The corollary can be restated entirely in terms of time. Define the *intersection mixing time* $T(\epsilon)$ to be the smallest integer with

$$\forall i \geq T(\epsilon),\, \forall Y_0 \leq X_0 : \frac{1-\epsilon}{m} \leq \mathsf{Pr}\left(\exists j : X_i = Y_j\right) \leq \frac{1+\epsilon}{m} .$$

**Theorem 4.2.** *Given an increasing transitive Markov chain on $\mathbb{Z}$ with intersection mixing time $T(\epsilon)$, run two independent instances of the walk with starting states $Y_0 < X_0 = 0$. The expected time until a collision is:*

$$\mathbb{E}\min\{i > 0 : \exists j,\, X_i = Y_j\} \leq 1 + \bar{S}\left(\frac{\sqrt{1+B_T} + \sqrt{T(\epsilon)/\bar{S}}}{1-\epsilon}\right)^2 ,$$

*where $B_T$ is the expected number of collisions before time $T(\epsilon)$.*

The proof requires only fairly straightforward changes from that of Theorem 2.2. We do not use this form and so we omit the proof.

*Proof of Corollary 4.1.* Let $\mathcal{T} = \min\{i : \exists j,\, X_i = Y_j\}$ be the time of the first collision, and $\mathcal{D} = X_{\mathcal{T}} - X_0$ the distance traveled by the $X$ walk until collision. Theorem 2.2 bounds $\mathbb{E}\mathcal{D}$. It is natural to expect that $\mathbb{E}\mathcal{T} = (\mathbb{E}\mathcal{D})/\bar{S}$. We show that this is true.

A random variable $Z_t$ is a Martingale if $\mathbb{E}[Z_{t+1} \mid Z_0, Z_1, \ldots, Z_t] = Z_t$. The optional stopping theorem says that if $\mathcal{T}$ is a bounded stopping time and $Z_t$ a Martingale then $\mathbb{E}Z_{\mathcal{T}} = \mathbb{E}Z_0$. This says that no strategy for stopping the process in finite time can do better than stopping at time 0.

The process $Z_t = (X_t - X_0) - t\bar{S}$ is a Martingale. By the optional stopping theorem, $\forall n \geq 0 : \mathbb{E}Z_{\min\{\mathcal{T},n\}} = Z_0 = 0$, so that

$$\mathbb{E}X_{\min\{\mathcal{T},n\}} = X_0 + (\mathbb{E}\min\{\mathcal{T},n\})\bar{S} .$$

Since $X_i$ is increasing then $\mathbb{E}X_{\min\{\mathcal{T},n\}} \leq \mathbb{E}X_{\mathcal{T}} = \mathbb{E}\mathcal{D}$, and so by Dominated Convergence this increasing bounded sequence has limit $\lim_{n\to\infty}\mathbb{E}X_{\min\{\mathcal{T},n\}} = \mathbb{E}X_{\mathcal{T}}$. Likewise $\lim_{n\to\infty}\mathbb{E}\min\{\mathcal{T},n\} = \mathbb{E}\mathcal{T}$. It follows that

$$\begin{aligned}\mathbb{E}X_{\mathcal{T}} &= X_0 + (\mathbb{E}\mathcal{T})\bar{S} \\ \Rightarrow\ \mathbb{E}\mathcal{T} &= \frac{\mathbb{E}X_{\mathcal{T}} - X_0}{\bar{S}} = \frac{\mathbb{E}\mathcal{D}}{\bar{S}} .\end{aligned}$$

□

# 5 Catching Kangaroos

Our primary reason for studying collision time and distance was to give a rigorous proof of the time-complexity of Pollard's Kangaroo Method for discrete logarithm. We describe the algorithm in more detail here and then give our proof.

### 5.1 Pollard's Kangaroo Method

We describe here the Distinguished Points implementation of van Oorschot and Wiener [4] because it is more efficient than Pollard's original implementation in [5].

**Problem:** *Given $g, h \in G$, solve for $x \in [a, b]$ with $h = g^x$.*

**Method:** Pollard's Kangaroo method (distinguished points version).

**Preliminary Steps:**

- Define a set $\mathsf{P} \subset G$ of "distinguished points," with $\omega(1)$ of every $\sqrt{b-a}$ distinguished. Our analysis requires only that $\log^3(b-a)$ of every $\sqrt{b-a}$ be distinguished.
- A set of jump sizes $S = \{s_0, s_1, \ldots, s_m\}$ with probability distribution $p : S \to [0,1]$. We analyze Pollard's suggestion to use powers of two: $S = \{2^k\}_{k=0}^m$ with $m \approx \log_2 \sqrt{b-a} + \log_2 \log_2 \sqrt{b-a} - 2$ chosen so that average $\bar{S} = \sum_{s \in S} s\, p(s) \approx \frac{\sqrt{b-a}}{2}$.
- A hash $F : G \to S$ giving pseudo-random jumps such that $\forall \mathfrak{g} \in G : \mathsf{Pr}\left(F(\mathfrak{g}) = s\right) \approx p(s)$.

**The Algorithm:**

- Let $Y_0 = \frac{a+b}{2}$, $X_0 = x$, and $d_0 = 0$. Observe that $g^{X_0} = hg^{d_0}$.
- Transitions $Y_{j+1} = Y_j + F(g^{Y_j})$ and $X_{i+1} = X_i + F(g^{X_i})$. Observe that if $d_{i+1} = d_i + F(hg^{d_i})$ then $X_{i+1} = x + d_{i+1}$.
- If $g^{Y_j} \in \mathsf{P}$ then store the pair $(g^{Y_j}, Y_j - Y_0)$ with an identifier $T$ (for tame). Likewise if $g^{X_i} = hg^{d_i} \in \mathsf{P}$ then store $(g^{X_i}, d_i)$ with an identifier $W$ (for wild).
- Once some distinguished point has been stored with both identifiers $T$ and $W$, say $g^{X_i} = g^{Y_j}$ where $(g^{X_i}, d_j)$ and $(g^{Y_j}, Y_j - Y_0)$ were stored, then

$$\begin{aligned} Y_j \equiv X_i \equiv x + d_i \quad &\bmod |G| \\ \implies x \equiv Y_j - d_i \quad &\bmod |G|. \end{aligned}$$

The $X_i$ and $Y_j$ walks are often called the "wild" and "tame" kangaroos, respectively.

We will make rigorous the following commonly used heuristic: If $X_0 \in [a, b]$ is a uniform random value then $\mathbb{E}\,|Y_0 - X_0|/\bar{S} = (b-a)/(4\bar{S})$ steps are required for the smaller of $X_0$ and $Y_0$ to reach the larger value. Subsequently each kangaroo visits a $1/\bar{S}$ fraction of states, so $\forall j : \mathsf{Pr}\left(\exists i : X_i = Y_j\right) \approx 1/\bar{S}$, and an average of $\bar{S}$ states are visited by the tame kangaroo until collision with the wild one. By symmetry the wild kangaroo also visits $\bar{S}$ states. So if walks are incremented simultaneously, $X_i \to X_{i+1}$ and $Y_i \to Y_{i+1}$, then in total each walk visits $(b-a)/(4\bar{S}) + \bar{S}$ states. This is minimized when $\bar{S} = \sqrt{b-a}/2$, for a total of $\sqrt{b-a}$ steps per kangaroo.

### 5.2 Analysis of the Kangaroo Method

We now turn our attention to the analysis of the Kangaroo Method. It is typically assumed that all transitions $X \to X + s$ with $s \in \{2^k\}_{k=0}^d$ are equally likely. However, some non-uniformity is needed to have exactly $\bar{S} = \sqrt{b-a}/2$ and so we state a result which allows non-uniformity.

In order to apply Corollary 4.1 we require an upper bound on collision distance $\mathcal{D}(\epsilon)$ and collision number $B_\epsilon$.

**Lemma 5.1.** *Consider a Kangaroo walk with step sizes $S = \{2^k\}_{k=0}^m$ and transition probabilities*

$$\frac{\gamma}{m+1} \geq p(s) \geq \frac{\gamma^{-1}}{m+1}\,,$$

*for some constant $\gamma \geq 1$. Then*

$$\mathcal{D}((m+1)^{-1}) \leq 14\gamma^{3.5}\,(m+1)^5\,\bar{S}\,.$$

*Proof.* The methods developed in Section 3.3 will be used, so when we write (Ib) then this is step (Ib) in Section 3.3. We encourage the reader to review our application of this method to the uniform distribution in Section 3.4.

It is easier to construct a stopping time on the slower "lazy" walk $\tilde{X}$ with $\tilde{X}_0 = X_0$, $\tilde{p}(s) = p(s)/2$ and $\tilde{p}(0) = 1/2$. This is just the kangaroo process slowed in time by a factor of two. Although slowing the walk doubles mixing time, it does not effect the sequence of distinct states visited by the walk, and so the mixing distance will be unchanged.

(Ia) Run walk $\tilde{X}$ as follows: given $\tilde{X}_i$ choose $s \in S$ with probability $p(s)$, half the time set $\tilde{X}_{i+1} = \tilde{X}_i + s$ and half the time set $\tilde{X}_{i+1} = \tilde{X}_i$. Let $\mathcal{T}_1$ be the first time every $s \in S \setminus \{2^m\}$ has been chosen at least once. Let $\delta_s \in \{0, s\}$ be the step size taken the first time $s \in S$ is chosen, so that $\mathsf{Pr}\,(\delta_s = 0) = \mathsf{Pr}\,(\delta_s = s) = 1/2$. Then $\delta = \sum_{s \in S\setminus\{2^m\}} \delta_s$ is uniform in $[0, 2^m)$, and if $\mathcal{D}_1 = \tilde{X}_{\mathcal{T}_1} - \tilde{X}_0 - \delta$ then $\tilde{X}_{\mathcal{T}} - \tilde{X}_0$ is a uniform sampler over interval $[\mathcal{D}_1, \mathcal{D}_1 + 2^m)$.

(Ib) Nothing to do. This step determines random $\mathcal{T}$ and $\mathcal{D}$ such that

$$\forall d \geq \mathcal{D} :\ \mathsf{Pr}\left(\exists j \geq \mathcal{T} :\ \tilde{X}_j - \tilde{X}_0 = d\right) = 1/\bar{S}\,.$$

(Ic') Let $p_{min} = \min_{s \in S} p(s)$ and $\Delta(\xi) = \frac{\log(m/\xi)}{p_{min}}$.

$$\begin{aligned}
\mathsf{Pr}\,(\mathcal{T}_1 > \Delta(\xi)) &= \mathsf{Pr}\left(\bigcup_{s \in S\setminus\{2^m\}} (s \text{ has not been chosen in } \Delta(\xi) \text{ steps})\right) \\
&\leq \sum_{s \in S\setminus\{2^m\}} \mathsf{Pr}\,(s \text{ has not been chosen in } \Delta(\xi) \text{ steps}) \\
&\leq m\,(1 - p_{min})^{\Delta(\xi)} \leq \xi\,. \qquad (9)
\end{aligned}$$

Then $\mathsf{Pr}\,(\mathcal{T} > \mathcal{T}_M) \leq \epsilon/\bar{S}$ when $M = 2\frac{S_{max}}{\bar{S}} \log(\bar{S}/\epsilon)$ and tentative stopping times are separated by at most $\Delta(\bar{S}/2S_{max})$ steps.

(Id) Finally, collision distance. Since $S_{max} \geq \bar{S} \geq p_{min}\,S_{max}$ then

$$\begin{aligned}
\mathcal{D}(\epsilon) &\leq 2\,\frac{\log(2mS_{max}/\bar{S})}{p_{min}}\left(\frac{S_{max}}{\bar{S}}\right)^2 \log(\bar{S}/\epsilon)\,\bar{S} \\
&\leq \frac{2}{p_{min}^3}\,\log\frac{2(m+1)}{p_{min}}\,\log\frac{\bar{S}}{\epsilon}\,\bar{S} \\
&\leq 2\gamma^3\,(m+1)^3\,\log(2\gamma(m+1)^2)\,(m\log 2 + \log(1/\epsilon))\,\bar{S}\,.
\end{aligned}$$

Finish the proof with the relation $\log(1 + x) \leq x$. □

To apply Theorem 2.2 it remains only to show that $B_{2/(m+1)}$ is small, i.e. one collision is unlikely to be quickly followed by another.

**Lemma 5.2.** *The kangaroo walk, as described in Lemma 5.1, has collision number*

$$B_{1/(m+1)} = \Theta\left(\frac{1}{m+1}\right) = o_m(1)\,.$$

*Proof.* This will be shown by applying Lemma 3.3 as sketched in Section 3.4.

(IIa) Determine $T$ using Lemma 3.3. Since $\mathcal{D}((m+1)^{-1}) \leq 14\gamma^{3.5}\,(m+1)^5\,\bar{S}$ and $\Delta(\xi) \leq \frac{\ln(m/\xi)}{p_{min}}$ then

$$\begin{aligned} T &\leq 2 \times 14\gamma^{3.5}\,(m+1)^5 + \frac{1}{2}\,(m+1)^2\,\log\left(14\gamma^{3.5}\,(m+1)^5\,2^m\,N\right) \\ &\leq 32\gamma^{3.5}\,(m+1)^5 + \frac{1}{2}\,(m+1)^2\,\log N\,. \end{aligned}$$

(IIc) Begin by upper bounding $\max_{u,v} \mathsf{P}^i(u,v)$.

Suppose $\mathsf{P}$ is uniform with $p(s) = 1/(m+1)$ for every $s \in S$. Then $\mathsf{P}^i(0,u) = \frac{c_i(u)}{(m+1)^i}$ where $c_i(u) \leq 2^{(m+1)}$ is the number of ways to write $u$ as the sum of $i$ non-distinct ordered elements of $S = \{2^k\}_{k=0}^m$. The binary expansion of $u$ has at most $i$ non-zero bits, each of is influenced only by powers of two drawn from $\{2^k\}_{k=\ell-i+1}^{\ell}$. So there are only $i^2$ bits which can be involved in an $i$-term summation to $u$. It follows that $c_i(u) \leq (i^2)^i$ and so $\max_{u,v} \mathsf{P}^i(v,u) = \max_u \mathsf{P}^i(0,u) = O_m((m+1)^{-i})$.

In the non-uniform case $\mathsf{P}^i(u,v) \leq \gamma^i\,\frac{c_i}{(m+1)^i} = O_m((m+1)^{-i})$.

It follows from (IIc) that if $k = 12$ and $N = (m+1)^2$ then

$$\begin{aligned} \frac{1}{m+1} \leq B_\epsilon &\leq \frac{2\gamma}{m+1} + 12^2\,\frac{c_2}{(m+1)^2} + T^2\,\frac{c_{12}}{(m+1)^{12}} + \frac{2}{N} \\ &= \frac{2\gamma}{m+1} + O((m+1)^{-2})\,. \end{aligned}$$

□

We now apply Corollary 4.1 to bound collision time for Pollard's Kangaroo method. As noted in our description of Pollard's algorithm, for our analysis we require only that roughly $\log^3(b-a)$ points are distinguished points in every interval of $\sqrt{b-a}$, so a little-o fraction with relatively weak uniformity requirement.

*Proof of Theorem 1.1.* The group elements $\left\{g^{(2^k)}\right\}_{k=0}^m$ can be pre-computed so that each step of a kangaroo requires only a single group multiplication. It then suffices to count the number of group multiplications as this is the number of steps of the walks in question.

As discussed in the heuristic argument of Section 5.1, an average of $\frac{|Y_0-X_0|}{\bar{S}}$ steps are needed to put the smaller of the starting states (e.g. $Y_0 < X_0$) within $S_{max} = 2^m$ of the one that started ahead. More precisely, average step size is $\bar{S} \approx \frac{\sqrt{b-a}}{2}$ with a standard deviation of $O^*(\sqrt{b-a})$, where $O^*$ indicates there may be extra logarithmic terms. In $(1+\epsilon)\frac{|Y_0-X_0|}{\bar{S}}$ steps the walk will have proceeded an expected distance of $(1+\epsilon)|Y_0 - X_0| = \omega((b-a)^{4/5})$ with probability 1, and have standard deviation of $O^*(\sqrt{|Y_0 - X_0|}\,\sqrt[4]{b-a}) = o((b-a)^{4/5})$, and so $(1+o(1))\frac{|Y_0-X_0|}{\bar{S}}$ steps suffice with probability 1.

Assume now that $|Y_0 - X_0| < S_{max}$, i.e. the rear walk has nearly caught up with the lead walk. Then Corollary 4.1, with $\epsilon = (m+1)^{-1}$, along with the bounds of Lemmas 5.1 and 5.2, shows an upper bound on expected collision time of $\bar{S}\,(1+\Theta(1/m))$ and a lower bound of $\bar{S}\,(1-\Theta(1/\sqrt{m}))$.

It remains to count the number of steps until a distinguished point has been reached. In step (Ia) of the proof of Lemma 5.1 we construct a stopping time $\mathcal{T}_1$ which samples uniformly on some interval $[d, d+S_{max})$. The expected run time $\mathbb{E}\mathcal{T}_1$ is a coupon collectors problem: each step of the $\tilde{X}$ walk has probability $\geq \frac{1}{2\gamma}\,\frac{m}{m+1}$ of sampling an element of $S\setminus\{2^m\}$, for an average of $2\gamma\,\frac{m+1}{m}$ steps to sample one element from $S\setminus\{2^m\}$, then $2\gamma\,\frac{m+1}{m-1}$ steps to sample another element of $S\setminus\{2^m\}$, etc., so that

$$\mathbb{E}\mathcal{T}_1 \leq 2\gamma\sum_{k=0}^{m-1}\frac{m+1}{m-k} = 2\gamma(m+1)\sum_{k=1}^{m}\frac{1}{k} \leq 2\gamma(m+1)(1+\ln m)\,.$$

Interval $[d, d+S_{max})$ is length $S_{max} = 2^m \sim \frac{1}{4}\sqrt{b-a}\log_2\sqrt{b-a}$ with $\Omega(\log^3(b-a))$ distinguished points, so with probability $\Omega(\log^2(b-a)/\sqrt{b-a})$ this is at a distinguished point. If not then repeat the process. This requires $O(\sqrt{b-a}/\log^2(b-a))$ rounds of $O(\gamma m\log m) = O(\log(b-a)\log\log(b-a))$ steps each, for a total of $o(\sqrt{b-a})$ additional steps.

The theorem follows by adding together the time in the catch-up phase, the expected collision time phase, and the distinguished points phase. □

## 6 Resolution of a Conjecture of Pollard

Pollard conjectured in [6] that Theorem 1.1 also holds for powers of any integer $n \geq 2$, as long as $\bar{S} \approx \frac{\sqrt{b-a}}{2}$. We now show his conjecture to be correct.

**Theorem 6.1.** *Consider a Kangaroo walk with step sizes $S = \{n^k\}_{k=0}^m$ and transition probabilities*

$$\frac{\gamma}{m+1} \geq p(s) \geq \frac{\gamma^{-1}}{m+1}$$

*such that $\bar{S} = \frac{\sqrt{b-a}}{2}$. Then Theorem 1.1 still holds.*

*Proof.* We detail only the differences from the case when $n = 2$, considered in Section 5.2.

Once again consider the lazy walk $\tilde{X}_t$ which half the time does nothing.

(Ia) Once a generator $s$ is chosen $(n-1)$ times then define $\Delta_s \in \{0, s, 2s, \ldots, (n-1)s\}$ to be the sum of the step sizes taken the first $(n-1)$ times that $s$ is chosen. Observe that $\mathsf{Pr}\,(\Delta_s = ks) = \binom{n-1}{k}/2^{n-1}$. "Accept" this with probability $\binom{n-1}{k}^{-1}$, in which case we define $\delta_s = \Delta_s = ks$. So the probability that $\delta_s$ is defined after $s$ is chosen $(n-1)$ times is $n/2^{n-1}$. Otherwise let $\Delta_s \in \{0, s, 2s, \ldots, (n-1)s\}$ be the sum of the step sizes taken the next $(n-1)$ times that $s$ is chosen, again check whether it's time to set $\delta_s = \Delta_s$, and if not repeat this procedure until $\delta_s$ has been defined. Once every $\delta_s$ has been defined then the sum $\delta = \sum_{s\in S\setminus\{n^m\}}\delta_s$ is a uniformly random $d$ digit number in base $n$. Let $\mathcal{T}_1$ denote the time when $\delta$ is finally defined.

(Ib) Nothing changes.

(Ic') By equation (9), in $j$ steps the probability that every generator $s \in S\setminus\{n^m\}$ has been chosen at least once is at least

$$1 - m(1 - 1/\gamma(m+1))^j \geq 1 - me^{-j/\gamma(m+1)}\,.$$

In $\alpha j$ steps the probability that each has been chosen at least $\alpha$ times is then at least

$$(1 - m\,e^{-j/\gamma(m+1)})^\alpha \geq 1 - \alpha\,m\,e^{-j/\gamma(m+1)}\,.$$

As discussed in (Ia), once generator $s$ is chosen $(n-1)$ times then we set $\delta_s = \Delta_s$ with probability $n/2^{n-1}$. So once $s$ is chosen $\alpha = 2^{n-1}\log(4\gamma(m+1)^2)$ times then $\delta_s$ is undefined with probability at most

$$\Big(1-\frac{n}{2^{n-1}}\Big)^{\alpha/(n-1)} \leq \frac{1}{4\gamma(m+1)^2}\,.$$

So once every $s \in S$ is chosen $\alpha$ times then some $\delta_s$ is undefined with probability at most $1/4\gamma(m+1)$.

It follows that if $j = \gamma(m+1)\log(\alpha m \cdot 4\gamma(m+1))$ then

$$\mathsf{Pr}\,(\mathcal{T}_1 > \alpha j) \leq \alpha\, m\, e^{-j/\gamma(m+1)} + \frac{1}{4\gamma(m+1)} \leq \frac{1}{2\gamma(m+1)}\,.$$

(Id), (II), Lemma 5.2: The remaining changes are straightforward consequences of the change in $\mathsf{Pr}\,(\mathcal{T}_1 > \alpha j)$. □

# References

[1] W. Haga and S. Robins, "On Kruskal's Principle," in *Organic Mathematics: Canadian Math. Society Conference Proceedings*, vol. 20, pp. 407–412 (1997).

[2] J-H. Kim, R. Montenegro, Y. Peres and P. Tetali, "A Birthday Paradox for Markov chains, with an optimal bound for collision in the Pollard Rho Algorithm for Discrete Logarithm," *The Annals of Applied Probability*, vol. 20(2), pp. 495–521 (2010).

[3] J. Lagarias, E. Rains and R.J. Vanderbei, "The Kruskal Count," in *The Mathematics of Preference, Choice and Order. Essays in Honor of Peter J. Fishburn*, (Stephen Brams, William V. Gehrlein and Fred S. Roberts, Eds.), Springer-Verlag: Berlin Heidelberg, pp. 371–391 (2009).

[4] P.C. van Oorschot and M.J. Wiener, "Parallel collision search with cryptanalytic applications," *Journal of Cryptology*, vol. 12(1), pp. 1–28 (1999).

[5] J. Pollard, "Monte Carlo methods for index computation mod p," *Mathematics of Computation*, vol. 32(143), pp. 918–924 (1978).

[6] J. Pollard, "Kangaroos, Monopoly and Discrete Logarithms," *Journal of Cryptology*, vol. 13(4), pp. 437–447 (2000).

[7] E. Teske, "Square-root Algorithms for the Discrete Logarithm Problem (A Survey)," in *Public-Key Cryptography and Computational Number Theory*, Walter de Gruyter, Berlin - New York, pp. 283–301 (2001).