How Long Does it Take to Catch a Wild Kangaroo?

Ravi Montenegro * Prasad Tetali †

ABSTRACT

The discrete logarithm problem asks to solve for the exponent x, given the generator g of a cyclic group G and an element $h \in G$ such that $g^x = h$. We give the first rigorous proof that Pollard's Kangaroo method finds the discrete logarithm in expected time $(3+o(1))\sqrt{b-a}$ for the worst value of $x \in [a, b]$, and $(2 + o(1))\sqrt{b-a}$ when $x \in_{uar} [a, b]$. This matches the conjectured time complexity and, rare among the analysis of algorithms based on Markov chains, even the lead constants 2 and 3 are correct.

ACM Classifiers: F.2.1, G.3
General Terms: Algorithms, Theory.
Keywords: Pollard's Kangaroo method, digital signature, discrete logarithm, Markov chain, mixing time.

1. INTRODUCTION

Cryptographic schemes are generally constructed in such a way that breaking them will likely require solving some presumably difficult computational problem, such as finding prime factors of a large integer or solving a discrete logarithm problem. Recall that the discrete logarithm problem asks to solve for the exponent x, given the generator g of a cyclic group G and an element $h \in G$ such that $g^x = h$. The Diffie-Hellman key exchange, ElGamal cryptosystem, and the US government's DSA (Digital Signature Algorithm) are all based on an assumption that discrete logarithm is difficult to find. Algorithms motivated by probabilistic intuition are often used to solve these problems and yet, although heuristics can be given for the time complexity of these methods, rigorous results are rare.

*Department of Mathematical Sciences, University of Massachusetts at Lowell, Lowell, MA 01854, USA. Email: ravi_montenegro@uml.edu
†School of Mathematics and School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: tetali@math.gatech.edu; research supported in part by NSF grants DMS 0401239, 0701043.

STOC'09, May 31–June 2, 2009, Bethesda, Maryland, USA.

In [2, 4] one such method is considered, namely Pollard's Rho Algorithm to find the discrete logarithm on a cyclic group G, verifying the correctness of commonly held intuition. This work generated further interest in the cryptography community, and Dan Boneh in particular encouraged us to analyze Pollard's Kangaroo method [6], due to its very many applications. While the Rho Algorithm is motivated by the Birthday Problem, the Kangaroo method works on the same principle as the Kruskal Count [7]. When the discrete logarithm x is known to lie in a small interval $[a, b]$ with $b - a \ll |G|$, this algorithm is expected to improve on the Rho algorithm, with a run time averaging $2\sqrt{b-a}$ steps, versus $\sqrt{(\pi/2)|G|}$ for the Rho algorithm. In fact, the Kangaroo method is often the most efficient means for finding discrete logarithm on an arbitrary cyclic group, as the Rho algorithm requires $|G|$ be known exactly and Shanks baby-step giant-step method requires too much memory.

Among the cases in which this would be useful, Boneh and Boyen [1] give a signature scheme in which a shorter signature can be transmitted if the recipient uses the Kangaroo method to determine certain omitted information. Verification of the time complexity of the Kangaroo method (as we do here) would then make rigorous their claim that the missing bits can be efficiently constructed. While the above is an application for signature communication, another natural application is in forging a signature. For instance, in order to speed up computation of a signature the secret key x may be chosen from an interval $[a, b]$ with $b - a \ll |G|$, or an attack might reveal a sequence of consecutive bits at the beginning or end of the key, in which cases the Kangaroo method can be used to find the key and forge a signature.

The Kangaroo method is based on running two independent sequences of hops (random walks), one starting at a known state (the "tame kangaroo") and the other starting at the unknown value of the discrete logarithm x (the "wild kangaroo"). The main result of this paper will be a bound on the expected number of steps required until these random walks intersect and the logarithm is determined.

THEOREM 1.1. *Suppose $g, h \in G$ are such that $h = g^x$ for some $x \in [a, b]$. The expected number of group operations required by the Distinguished Points implementation of the Kangaroo method is*

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

when $x = a$ or $x = b$, and is otherwise upper bounded by this. If x is a uniform random value in $[a, b]$, i.e. $x \in_{uar} [a, b]$, then the expected number of group operations is

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

We show matching upper and lower bounds, so the lead constants are sharp, which is quite rare among the analyses of algorithms based on Markov chains. Previously the first bound was known only by a rough heuristic, while Pollard [7] gives a convincing but not completely rigorous argument for the second. 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 30 years since it was first proposed in [6].

Past work on problems related to the Kruskal Count seem to be of little help here. Pollard's argument of [7] gives rigorous results for specific values of $\sqrt{b-a}$, but the recurrence relations he uses can only be solved on a case-by-case basis by numerical computation. Lagarias et.al. [3] used probabilistic methods to study the *distance traveled* before two walks intersect, but only for walks in which the number of steps until an intersection was simple to bound. Although our approach here borrows a few concepts from the study of the Rho algorithm in [2], such as the use of a second moment method to study the number of intersections, a significant complication in studying this algorithm is that, when $b-a \ll |G|$, the kangaroos will have proceeded only a small way around the cyclic group before the algorithm terminates. As such, mixing time is no longer a useful notion, and instead a notion of convergence is required which occurs long before the mixing time. We expect that the tools developed in this paper to avoid this problem will prove useful in examining other such randomized algorithms.

Our analysis assumes that the Kangaroo method involves a truly random hash function: if $\mathfrak{g} \in G$ then $F(\mathfrak{g})$ is equally likely to be any of the jump sizes, independent of all other $F(\mathfrak{g}')$. In practice different hash functions will be used on different groups – whether over a subgroup of integers mod p, elliptic curve groups, etc – but in general the hash is chosen to "look random." Since the Kangaroo method applies on all cyclic groups then a constructive proof would involve the impossible task of explicitly constructing a hash on every cyclic group, and so the assumption of a truly random hash is made in all attempts at analyzing it of which we are aware [8, 5, 7].

The paper proceeds as follows. In Section 2 we introduce the Kangaroo method. A general framework for analyzing intersection of independent walks on the integers is constructed in Section 3. This is followed by a detailed analysis for the Kangaroo method in Section 4.

2. PRELIMINARIES

We describe here the Kangaroo method, originally known as the Lambda method for catching Kangaroos. The Distinguished Points implementation of [5] is given because it is more efficient than the original implementation of [6].

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 $D \subset G$ of "distinguished points", with $\frac{|D|}{|G|} = \frac{c}{\sqrt{b-a}}$ for some constant c.
- Define a set of jump sizes $S = \{s_0, s_1, \ldots, s_d\}$. We consider powers of two, $S = \{2^k\}_{k=0}^{d}$, with $d \approx \log_2 \sqrt{b-a} + \log_2 \log_2 \sqrt{b-a} - 2$, chosen so that elements of S average to a jump size of $\frac{\sqrt{b-a}}{2}$.
- Finally, a hash function $F : G \to S$.

The Algorithm:

- Let $Y_0 = \frac{a+b}{2}$, $X_0 = x$, and $d_0 = 0$. Observe that $g^{X_0} = hg^{d_0}$.
- Recursively define $Y_{j+1} = Y_j + F(g^{Y_j})$ and likewise $d_{i+1} = d_i + F(hg^{d_i})$. This implicitly defines $X_{i+1} = X_i + F(g^{X_i}) = x + d_{i+1}$.
- If $g^{Y_j} \in D$ then store the pair $(g^{Y_j}, Y_j - Y_0)$ with an identifier T (for tame). Likewise if $g^{X_i} = hg^{d_i} \in D$ 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

$$\begin{aligned} Y_j \equiv X_i \equiv x + d_i \mod |G| \\ \implies x \equiv Y_j - d_i \mod |G| \end{aligned}$$

The Y_j walk is called the "tame kangaroo" because its position is known, whereas the position X_i of the "wild kangaroo" is to be determined by the algorithm. This was originally known as the Lambda method because the two walks are initially different, but once $g^{Y_j} = g^{X_i}$ then they proceed along the same route, forming a λ shape.

Theorem 1.1 makes rigorous the following commonly used rough heuristic: Suppose $Y_0 \geq X_0$. Run the tame kangaroo infinitely far. Since the kangaroos have an average step size $\frac{\sqrt{b-a}}{2}$, one expects the wild kangaroo requires $\frac{Y_0 - X_0}{\sqrt{b-a}/2}$ steps to reach Y_0. Subsequently, at each step the probability that the wild kangaroo lands on a spot visited by the tame kangaroo is roughly $p = \frac{1}{\sqrt{b-a}/2}$, so the expected number of additional steps by the wild kangaroo until a collision is then around $p^{-1} = \frac{\sqrt{b-a}}{2}$. By symmetry the tame kangaroo also averaged p^{-1} steps. About $\frac{\sqrt{b-a}}{c}$ additional steps are required until a distinguished point is reached. Since X_i and Y_j are incremented simultaneously the total number of steps taken is then

$$2\left(\frac{|Y_0 - X_0|}{\sqrt{b-a}/2} + p^{-1} + \frac{\sqrt{b-a}}{c}\right) \leq (3 + 2c^{-1})\sqrt{b-a}$$

If $X_0 = x \in_{uar} [a,b]$ then $E\frac{|Y_0 - X_0|}{\sqrt{b-a}/2} = \frac{\sqrt{b-a}}{2}$ and the bound is $(2 + 2c^{-1})\sqrt{b-a}$.

Recall we assume a random hash function F: if $\mathfrak{g} \in G$ then $F(\mathfrak{g})$ is equally likely to be any value in S, independent of all other $F(\mathfrak{g}')$. A second assumption is that the distinguished points are well distributed with $c \xrightarrow{(b-a)\to\infty} \infty$; either they are chosen uniformly at random, or if $c = \Omega(d^2 \log d)$ then roughly constant spacing between points will suffice. The assumption on distinguished points can be dropped if one instead analyzes Pollard's (slower) original algorithm, to which our methods also apply.

3. UNIFORM INTERSECTION TIME AND A COLLISION BOUND

In order to understand our approach to bounding time until the kangaroos have visited a common location, which we call a *collision*, it will be helpful to consider a simplified version of the Kangaroo method. First, observe that because hash values $F(\mathfrak{g})$ are independent then X_i and Y_j are independent random walks at least until they collide, and so to bound time until this occurs it suffices to assume they are independent random walks even after they have collided. Second, these are random walks on $\mathbb{Z}/|G|\mathbb{Z}$, so if we drop the modular arithmetic and work on $\mathbb{Z}$ then the time until a collision can only be made worse. Third, since the walks proceed strictly in the positive direction on $\mathbb{Z}$ then in order to determine the number of hops the "wild kangaroo" (described by X_i) takes until it is caught by the "tame kangaroo" (i.e. $X_i = Y_j$ on $\mathbb{Z}$), it suffices to run the tame kangaroo infinitely long and only after this have the wild kangaroo start hopping.

With these simplifications the problem reduces to one about intersection of walks X_i and Y_j, both proceeding in the positive direction on the integers, in which Y_j proceeds an infinite number of steps and then X_i proceeds until some $X_i = Y_j$. Thus, rather than considering a specific probability $\mathsf{Pr}(X_i = Y_j)$ it is better to look at $\mathsf{Pr}(\exists j : X_i = Y_j)$. By symmetry, the same approach will also bound the expected number of hops the tame kangaroo requires to reach the location where it can trap the wild kangaroo.

First however, because the walk does not proceed long enough to approach its stationary distribution (obvious on $\mathbb{Z}$ and also true on $\mathbb{Z}/|G|\mathbb{Z}$ when $b - a \ll |G|$), alternate notions resembling mixing time and a stationary distribution will be required. Recall the heuristic that at each step the probability of a collision is roughly the inverse of the average step size. Our mixing time quantity will measure the number of steps required for this to become a rigorous statement:

DEFINITION 3.1. *Consider a Markov chain* P *on* $\mathbb{Z}$ *which is* increasing *(i.e.* $\mathsf{P}(u,v) > 0$ *only when* $v > u$*) and* transitive *(i.e.* $\forall u, v \in \mathbb{Z} : \mathsf{P}(u,v) = \mathsf{P}(0, v-u)$*). Define the* uniform intersection probability U *by*

$$U = \frac{1}{\sum_{k=1}^{\infty} k\mathsf{P}(0,k)} .$$

Let X_i *and* Y_j *denote independent walks starting from states* $(X_0, Y_0) \in \Omega$*, where*

$$\Omega = \{(x,y) \in \mathbb{Z} \times \mathbb{Z} : |x - y| < \max\{k : \mathsf{P}(0,k) > 0\}\} .$$

If $\epsilon \in [0,1]$ *the* uniform intersection time $T(\epsilon) \in \mathbb{N}$ *is*

$$T(\epsilon) = \min\left\{ T \;\middle|\; \begin{array}{c} \forall i \geq T, \forall (X_0, Y_0) \in \Omega, \\ \left| \frac{\mathsf{Pr}(\exists j : X_i = Y_j)}{U} - 1 \right| \leq \epsilon \end{array} \right\} .$$

If $gcd\{k : \mathsf{P}(0,k) > 0\} = 1$ then the uniform intersection time is finite. To avoid clutter we write T to denote $T(\epsilon)$ in the remainder.

More generally, our results will apply if P is an increasing Markov chain on a poset S, i.e. $\mathsf{P}(u,v) > 0$ only if $u \prec v$, and Ω is a binary relation on S which is reflexive, symmetric, and such that if $(X_0, Y_0) \in \Omega$ then

$$\forall X_i \exists Y_j : (X_i, Y_j) \in \Omega \wedge (Y_j \not\prec X_i) .$$

A uniform intersection probability would be any value U satisfying the definition for the uniform intersection time.

A natural approach to studying collisions is to consider an appropriate random variable counting the number of intersections of the two walks. Towards this, let S_N denote the number of times the X_i walk intersects the Y_j walk in the first N steps, i.e.

$$S_N = \sum_{i=1}^{N} \mathbf{1}_{\{\exists j : X_i = Y_j\}} .$$

The second moment method used will involve showing that $\mathsf{Pr}(S_N > 0)$ is non-trivial for some N.

Our collision bound will involve the quantity B_T, the expected number of collisions in the first T steps between two independent walks started at nearby states. To be precise, define:

$$B_T = \max_{(X_0,Y_0)\in\Omega} \sum_{i=1}^{T} \mathsf{Pr}(\exists j : X_i = Y_j) .$$

Then the expected number of steps until a collision can be bounded as follows.

THEOREM 3.2. *Given an increasing Markov chain* P *on* $\mathbb{Z}$*, if two independent walks have starting states* $(X_0, Y_0) \in \Omega$ *then*

$$\begin{aligned} \frac{1 - 2\sqrt{B_T}}{U(1+\epsilon)} - T &\leq \mathbb{E} \min\{i > 0 : \exists j,\, X_i = Y_j\} \\ &\leq (1 - 4\epsilon)^{-1} \left(\sqrt{T} + \sqrt{\frac{1 + 2B_T}{U}} \right)^2 \end{aligned}$$

If $B_T, \epsilon \approx 0$ and $U^{-1} \gg T$ then these bounds show that

$$\mathbb{E} \min\{i > 0 : \exists j,\, X_i = Y_j\} \sim \frac{1}{U} ,$$

which makes rigorous the heuristic that the expected number of steps needed until a collision is the average step size.

It will prove easiest to study S_N by first considering the first and second moments of the number of intersections in steps $T+1$ to N, i.e.

$$\mathcal{R}_N = \sum_{i=T+1}^{N} \mathbf{1}_{\{\exists j : X_i = Y_j\}} ,$$

in terms of the uniform intersection time and probability:

LEMMA 3.3. *Under the conditions of Theorem 3.2, if* $N \geq T = T(\epsilon)$ *then*

$$(1-\epsilon)(N-T)U \leq E[\mathcal{R}_N] \leq (1+\epsilon)(N-T)U$$

$$E[\mathcal{R}_N^2]^{1/2} \leq (1+\epsilon)(N-T)U \left[1 + \frac{1 + 2B_T}{(N-T)U} \right]^{1/2}$$

PROOF. The expectation $E[\mathcal{R}_N]$ satisfies

$$\begin{aligned} E[\mathcal{R}_N] &= E \sum_{i=T+1}^{N} \mathbf{1}_{\{\exists j : X_i = Y_j\}} \\ &= \sum_{i=T+1}^{N} E[\mathbf{1}_{\{\exists j : X_i = Y_j\}}] \\ &\geq (N-T)U(1-\epsilon) \end{aligned} \tag{1}$$

The inequality follows from the relation $E[\mathbf{1}_{\{\exists j:\, X_i=Y_j\}}] = \mathsf{Pr}\,(\exists j:\, X_i = Y_j)$. The upper bound on $E[\mathcal{R}_N]$ follows by taking $(1+\epsilon)$ in place of $(1-\epsilon)$.

Now for $E[\mathcal{R}_N^2]$. Note that

$$\begin{aligned} E[\mathcal{R}_N^2] &= E\left[\sum_{i=T+1}^{N}\sum_{k=T+1}^{N} \mathbf{1}_{\{\exists j:\, X_i=Y_j\}}\mathbf{1}_{\{\exists \ell:\, X_k=Y_\ell\}}\right] \\ &= \sum_{i=T+1}^{N}\sum_{k=T+1}^{N} \mathsf{Pr}\,(\exists j,\ell:\, X_i=Y_j,\, X_k=Y_\ell)\,. \end{aligned}$$

By symmetry it suffices to consider the case that $k \geq i > T$. Then if $X_i = Y_j$ then because the X and Y walks are increasing then $X_k = Y_\ell$ is possible only if $\ell \geq j$.

When $k > i+T$ then $\mathsf{Pr}\,(\exists \ell:\, X_k = Y_\ell \mid X_i = Y_j) \leq U(1+\epsilon)$ by definition of T, and so

$$\begin{aligned} &\mathsf{Pr}\,(\exists j,\ell:\, X_i=Y_j,\, X_k=Y_\ell) \\ &\quad= \mathsf{Pr}\,(\exists j:\, X_i=Y_j)\,\mathsf{Pr}\,(\exists \ell:\, X_k=Y_\ell \mid X_i=Y_j) \\ &\quad\leq (1+\epsilon)^2U^2\,. \end{aligned}$$

When $k \leq i+T$ then

$$\begin{aligned} &\sum_{k=i+1}^{i+T} \mathsf{Pr}\,(\exists j,\ell:\, X_i=Y_j,\, X_k=Y_\ell) \\ &\quad\leq \mathsf{Pr}\,(\exists j:\, X_i=Y_j) \\ &\qquad\times \max_u \sum_{k=1}^{T} \mathsf{Pr}\,(\exists \ell:\, X_k=Y_\ell \mid X_0=Y_0=u) \\ &\quad\leq B_T\,U(1+\epsilon)\,, \end{aligned}$$

since $i \geq T$. It follows that

$$\begin{aligned} &E[\mathcal{R}_N^2] \\ &= \sum_{i=T+1}^{N}\Bigg(\mathsf{Pr}\,(\exists j:\, X_i=Y_j) \\ &\qquad +2\sum_{k=i+1}^{i+T} \mathsf{Pr}\,(\exists j,\ell:\, X_i=Y_j,\, X_k=Y_\ell)\Bigg) \\ &\quad +2\sum_{i=T+1}^{N}\sum_{k=i+T+1}^{N} \mathsf{Pr}\,(\exists j,\ell:\, X_i=Y_j,\, X_k=Y_\ell) \\ &\leq (N-T)U(1+\epsilon)(1+2B_T) \\ &\quad +2\frac{(N-2T)(N-2T+1)}{2}U^2(1+\epsilon)^2 \\ &\leq (1+\epsilon)^2U^2(N-T)^2\left[1+\frac{1+2B_T}{(1+\epsilon)U(N-T)}\right]. \end{aligned}$$

□

This shows that if the number of steps N is much bigger than U^{-1} then the standard deviation will be small relative to expectation. Thus $\mathcal{R}_N$ is concentrated around its mean. In particular we have

LEMMA 3.4. *Under the conditions of Theorem 3.2, if $N \geq T$ then*

$$\begin{aligned} \mathsf{Pr}\,(S_N>0) &\leq B_T + (N-T)U(1+\epsilon) \\ \mathsf{Pr}\,(S_N>0) &\geq (1-4\epsilon)\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1}. \end{aligned}$$

PROOF. Observe that $\mathsf{Pr}\,(S_N>0) \geq \mathsf{Pr}\,(\mathcal{R}_N>0)$, so for the lower bound it suffices to consider $\mathcal{R}_N$. Recall the standard second moment bound: using Cauchy-Schwartz, we have that

$$E[\mathcal{R}_N] = E[\mathcal{R}_N\mathbf{1}_{\{\mathcal{R}_N>0\}}] \leq E[\mathcal{R}_N^2]^{1/2}E[\mathbf{1}_{\{\mathcal{R}_N>0\}}]^{1/2}$$

and hence $\mathsf{Pr}\,(\mathcal{R}_N>0) \geq E[\mathcal{R}_N]^2/E[\mathcal{R}_N^2]$. By Lemma 3.3 then, independent of starting point,

$$\begin{aligned} \mathsf{Pr}\,(\mathcal{R}_N>0) &\geq \left(\frac{1-\epsilon}{1+\epsilon}\right)^2\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1} \\ &\geq (1-4\epsilon)\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1}, \end{aligned}$$

since $\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \geq 1-4\epsilon$, for $\epsilon \geq 0$.

Now to upper bound $\mathsf{Pr}\,(S_N>0)$. Since $S_N \in \mathbb{N}$ then

$$\mathsf{Pr}\,(S_N>0) = E[\mathbf{1}_{\{S_N>0\}}] \leq E[S_N]\,.$$

The expectation $E[S_N]$ satisfies

$$\begin{aligned} E[S_N] &= E\sum_{i=1}^{N}\mathbf{1}_{\{\exists j:\, X_i=Y_j\}} = \sum_{i=1}^{N}E[\mathbf{1}_{\{\exists j:\, X_i=Y_j\}}] \\ &= \sum_{i=1}^{T}\mathsf{Pr}\,(\exists j:\, X_i=Y_j) + \sum_{i=T+1}^{N}\mathsf{Pr}\,(\exists j:\, X_i=Y_j) \\ &\leq B_T + (N-T)U(1+\epsilon)\,. \end{aligned}$$

□

PROOF OF THEOREM 3.2. To start with, upper and lower bounds on $\mathsf{Pr}\,(S_{kN}=0)$ will be shown in terms of $k \geq 1$.

For $\ell \geq 1$, let

$$S_N^{(\ell)} = \sum_{i=1}^{N}\mathbf{1}_{\{\exists j:\, X_{(\ell-1)N+i}=Y_j\}}\,,$$

so that $S_N^{(1)} = S_N$. By taking $X_0 \leftarrow X_{(\ell-1)N}$ and $Y_0 \leftarrow Y_j$ where j is the smallest index such that $(X_{(\ell-1)N}, Y_j) \in \Omega$ and $Y_j \geq X_{(\ell-1)N}$, then by Lemma 3.4 we may bound:

$$\begin{aligned} &B_T + (N-T)U(1+\epsilon) \\ &\quad\geq 1 - \mathsf{Pr}\left(S_N^{(\ell)}=0 \mid S_{(\ell-1)N}=0\right) \\ &\quad\geq (1-4\epsilon)\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1}. \end{aligned}$$

But

$$\mathsf{Pr}\,(S_{kN}=0) = \prod_{\ell=1}^{k}\mathsf{Pr}\left(S_N^{(\ell)}=0 \mid S_{(\ell-1)N}=0\right)$$

and so

$$\begin{aligned} &\left(1-(1-4\epsilon)\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1}\right)^k \\ &\quad\geq \mathsf{Pr}\,(S_{kN}=0) \\ &\quad\geq (1-B_T-(N-T)U(1+\epsilon))^k\,. \end{aligned}$$

These upper and lower bounds will now be used to bound the collision time.

First, the upper bound.

$$\begin{aligned} & E\min\{i:\ S_i>0\} \\ &= E\sum_{i=0}^{\infty}\mathbf{1}_{\{S_i=0\}} = 1+\sum_{i=0}^{\infty}\mathsf{Pr}\left(S_i=0\right) \\ &\le \sum_{k=0}^{\infty}\mathsf{Pr}\left(S_{kN}=0\right)N \\ &\le N\sum_{k=0}^{\infty}\left(1-(1-4\epsilon)\left[1+\frac{1+2B_T}{(N-T)U}\right]^{-1}\right)^k \\ &= (1-4\epsilon)^{-1}N\left(1+\frac{1+2B_T}{(N-T)U}\right). \end{aligned}$$

This is minimized when $N = T+\sqrt{\frac{(1+2B_T)T}{U}}$, which gives the upper bound of the theorem.

To show the lower bound, take

$$\begin{aligned} & E\min\{i:\ S_i>0\} \\ &= \sum_{i=0}^{\infty}\mathsf{Pr}\left(S_i=0\right) \ge \sum_{k=1}^{\infty}\mathsf{Pr}\left(S_{kN}=0\right)N \\ &\ge N\sum_{k=1}^{\infty}\left(1-B_T-(N-T)U(1+\epsilon)\right)^k \\ &= N\left(\frac{1}{B_T+(N-T)U(1+\epsilon)}-1\right). \end{aligned}$$

If $B_T \ge 1$ then the bound stated in the theorem is trivial, so assume $B_T < 1$.

If $B_T(1-B_T) < TU(1+\epsilon)$ then the maximum of the above bound is at $N = T$. In this case the bound is

$$\begin{aligned} E\min\{i:\ S_i>0\} &\ge N\left(\frac{1}{B_T}-1\right) \\ &\ge \frac{1-B_T}{U(1+\epsilon)}-T. \end{aligned}$$

When $B_T(1-B_T) \ge TU(1+\epsilon)$ then the maximum is at $N = \frac{\gamma(1-\gamma)}{U(1+\epsilon)}$, where $\gamma = \sqrt{B_T - TU(1+\epsilon)}$. In this case the bound is

$$\begin{aligned} E\min\{i:\ S_i>0\} &\ge \frac{\left(1-\sqrt{B_T-TU(1+\epsilon)}\right)^2}{U(1+\epsilon)} \\ &\ge \frac{(1-\sqrt{B_T})^2}{U(1+\epsilon)}. \end{aligned}$$

□

To bound the value of B_T it will prove easier to consider those intersections that occur early in the Y_j walk separately from those that occur later.

LEMMA 3.5. *Let $\gamma \in [0,1]$ and $\tau \ge T$ be such that*

$$\forall (X_0,Y_0)\in\Omega:\ \mathsf{Pr}\left(Y_\tau < X_T\right) \le \gamma.$$

Then

$$B_T \le \gamma T + (\tau - T)U(1+\epsilon) + \sum_{i=1}^{T}(1+2i)\max_{u,v}\mathsf{P}^i(u,v).$$

PROOF. Recall that

$$B_T = \max_{(X_0,Y_0)\in\Omega}\sum_{i=1}^{T}\mathsf{Pr}\left(\exists j:\ X_i=Y_j\right).$$

When $j > \tau$ then

$$\begin{aligned} \sum_{i=1}^{T}\mathsf{Pr}\left(\exists j>\tau:\ X_i=Y_j\right) &\le T\ \mathsf{Pr}\left(Y_\tau < X_T\right) \\ &\le \gamma T. \end{aligned}$$

When $T < j \le \tau$ then

$$\begin{aligned} & \sum_{i=1}^{T}\mathsf{Pr}\left(\exists j\in(T,\tau]:\ X_i=Y_j\right) \\ &= \sum_{i=1}^{T}E\sum_{j=T+1}^{\tau}\mathbf{1}_{\{X_i=Y_j\}} \\ &= \sum_{j=T+1}^{\tau}E\sum_{i=1}^{T}\mathbf{1}_{\{X_i=Y_j\}} \\ &= \sum_{j=T+1}^{\tau}\mathsf{Pr}\left(\exists i\in[1,T]:\ X_i=Y_j\right) \\ &\le \sum_{j=T+1}^{\tau}\mathsf{Pr}\left(\exists i:\ X_i=Y_j\right) \\ &\le (\tau-T)U(1+\epsilon) \end{aligned}$$

The first equality is because the walks are increasing, so for fixed i there can be at most one j with $X_i = Y_j$. Likewise for the third equality but for fixed j. The final inequality is because Ω is symmetric, so the uniform intersection time also holds if the roles of the X and Y walks are reversed.

When $j \le T$ then

$$\begin{aligned} & \sum_{i=1}^{T}\mathsf{Pr}\left(\exists j\le T:\ X_i=Y_j\right) \\ &\le \sum_{i=1}^{T}\sum_{j=0}^{T}\sum_{w}\mathsf{P}^i(X_0,w)\mathsf{P}^j(Y_0,w) \\ &\le \sum_{i=1}^{T}\max_{u,v}\mathsf{P}^i(u,v)\sum_{j=0}^{i}(1+\mathbf{1}_{\{j<i\}})\max_{z}\sum_{w}\mathsf{P}^j(z,w) \\ &= \sum_{i=1}^{T}(1+2i)\max_{u,v}\mathsf{P}^i(u,v). \end{aligned}$$

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

4. CATCHING KANGAROOS

The collision results of the previous section will now be applied to the Kangaroo method. Recall that d is chosen so that the average step size is roughly $\frac{\sqrt{b-a}}{2}$, and so in particular the uniform intersection probability is

$$U = \frac{d+1}{2^{d+1}-1} \sim \frac{2}{\sqrt{b-a}}.$$

Throughout we take

$$\Omega = \{(X_0,Y_0)\in\mathbb{Z}\times\mathbb{Z}:\ |X_0-Y_0| < 2^d\}.$$

The first step in bounding collision time will be to bound the uniform intersection time. This will be done by selecting some d of the first T steps of the X_i walk (for suitable T), and using these to construct a uniformly random d-bit

binary string which is independent of the specific step sizes taken on other steps. This implies that if $i \geq T$ then the X_i walk is uniformly distributed over some interval of 2^d elements, and so the probability that some $X_i = Y_j$ will be exactly the expected number of times the Y_j walk visits this interval, divided by the interval size (i.e. 2^d).

LEMMA 4.1. *The Kangaroo walk has*

$$T\left(\frac{3}{d+1}\right) \leq (d+1)^2 \ln 2 + (d+1)\ln d\,.$$

That is, if $(X_0, Y_0) \in \Omega$ *and* $i \geq (d+1)^2 \ln 2 + (d+1)\ln d$ *then*

$$\left|\frac{\mathsf{Pr}\left(\exists j:\ X_i = Y_j\right)}{U} - 1\right| \leq \frac{3}{d+1} \sim \frac{3}{\log_2 \sqrt{b-a}}\,.$$

PROOF. The X_i walk will be implemented by choosing $k \in_{uar} \{0, 1, \ldots, d\}$ and then flipping a coin to decide whether to increment by 2^k or 2^{k+1} (if $k = d$ then increment by 2^d or 2^0). We say generator 2^k has been chosen if value k was chosen, even though the step size taken might not be 2^k.

We now decompose the X_i walk into a few components. For $k \in \{0, 1, \ldots, d-1\}$ let δ_k denote the step taken the first time generator 2^k is chosen, so that $\delta_k - 2^k \in_{uar} \{0, 2^k\}$. Also, let $\mathcal{T}$ be the first time all of the generators $\{2^k\}_{k=0}^{d-1}$ have been chosen (so ignore generator 2^d). Define

$$\delta = \sum_{k=0}^{d-1} (\delta_k - 2^k) \in_{uar} \{0, 1, \ldots, 2^d - 1\}$$

and let $\mathcal{I}_i$ denote the sum of all increments in the first i steps except those incorporated in a δ_k, so that if $i \geq \mathcal{T}$ then $X_i = X_0 + \mathcal{I}_i + 2^d - 1 + \delta$.

Suppose $i \geq \mathcal{T}$. Then δ is independent of the value of $\mathcal{I}_i$, and so

$$X_i \in_{uar} [X_0 + \mathcal{I}_i + 2^d - 1, X_0 + \mathcal{I}_i + 2^{d+1} - 2]\,.$$

Observe that $X_0 + \mathcal{I}_i + 2^d - 1 \geq X_0 + 2^d - 1 \geq Y_0$. Since U^{-1} is the average step size for Y_j then

$$\begin{aligned}
&\mathsf{Pr}\left(\exists j:\ X_i = Y_j \mid i \geq \mathcal{T}\right)\\
&\quad = \frac{\mathbb{E}\left|\{Y_j\} \cap [X_0 + \mathcal{I}_i + 2^d - 1, X_0 + \mathcal{I}_i + 2^{d+1} - 2]\right|}{2^d}\\
&\quad \geq \frac{\lfloor 2^d / U^{-1} \rfloor}{2^d} \geq U - 2^{-d}
\end{aligned}$$

An upper bound of $U + 2^{-d}$ follows by taking ceiling instead of floor, and so

$$\left|\mathsf{Pr}\left(\exists j:\ X_i = Y_j \mid i \geq \mathcal{T}\right) - U\right| \leq 2^{-d}\,. \tag{2}$$

Next, consider $\mathcal{T}$. In T steps the probability that not all generators $\{2^k\}_{k=0}^{d-1}$ have been chosen is at most

$$\mathsf{Pr}\left(\mathcal{T} > T\right) \leq d\left(1 - \frac{1}{d+1}\right)^T \leq de^{-T/(d+1)}\,.$$

It follows that

$$\mathsf{Pr}\left(\mathcal{T} \geq (d+1)\ln\left(d\,2^{d+1}\right)\right) \leq 2^{-(d+1)}\,.$$

Thus, if $i \geq (d+1)^2 \ln 2 + (d+1)\ln d$ then

$$\begin{aligned}
&\mathsf{Pr}\left(\exists j:\ X_i = Y_j\right)\\
&\quad = (1 - \mathsf{Pr}\left(\mathcal{T} > i\right))\mathsf{Pr}\left(\exists j:\ X_i = Y_j \mid i \geq \mathcal{T}\right)\\
&\qquad + \mathsf{Pr}\left(\mathcal{T} > i\right)\ \mathsf{Pr}\left(\exists j:\ X_i = Y_j \mid i < \mathcal{T}\right)\,.
\end{aligned}$$

Since $0 \leq \mathsf{Pr}\left(\mathcal{T} > i\right) \leq 2^{-(d+1)}$ and the other probabilities are in $[0, 1]$, then

$$\left|\mathsf{Pr}\left(\exists j:\ X_i = Y_j\right) - \mathsf{Pr}\left(\exists j:\ X_i = Y_j \mid i \geq \mathcal{T}\right)\right| \leq 2^{-(d+1)} \tag{3}$$

By (2) and (3) then

$$\left|\mathsf{Pr}\left(\exists j:\ X_i = Y_j\right) - U\right| \leq 2^{-d} + 2^{-(d+1)} \leq \frac{3U}{d+1}\,.$$

□

It remains only to upper bound B_T.

LEMMA 4.2. *If* $T = (d+1)^2 \ln 2 + (d+1)\ln d$ *then* $B_T = o_d(1)$.

PROOF. This will be shown by applying Lemma 3.5.

We compute the first few values of $\max_v \mathsf{P}^i(u, v)$ directly. Observe that $\mathsf{P}^i(u, v)$ is exactly $\frac{c_i}{(d+1)^i}$ where c_i is the number of ways to write $v - u$ as the sum of i (non-distinct, ordered) elements of $\{2^k\}_{k=0}^d$. To determine c_i note that if the binary expansion of $v - u$ contains B non-zero bits, then any non-zero bit 2^ℓ came about as the sum of at most $i - B + 1$ terms of $\{2^k\}_{k=\ell-(i-B)}^{\ell}$, and so any string of more than $i - B$ consecutive zeros can be contracted to $i - B$ zeros without effecting the number of ways to write $v - u$, i.e. it suffices to consider all $\max_B B(i - B + 1) \leq \left(\frac{i+1}{2}\right)^2$ bit strings. This simplifies the problem into few enough cases to make it feasible by brute force, done either by hand or on a computer. Either way we find the following:

$\mathbf{i}$	$\mathbf{c_i}$	**max at** $\mathbf{v - u} =$
1	1	1_2
2	2!	11_2
3	3!	111_2
4	$\binom{3}{1}\binom{4}{2,1,1} = 36$	101010_2
5	$\binom{3}{2}\binom{5}{2,2,1} + \binom{3}{1}\binom{5}{2,1,1,1} = 270$	100100100_2
6	$\binom{4}{2}\binom{6}{2,2,1,1} + \binom{4}{1}\binom{6}{2,1,1,1,1} = 2520$	100100100100_2

If $i \geq 6$ then

$$\begin{aligned}
\max_v \mathsf{P}^i(u, v) &= \max_v \sum_w \mathsf{P}^{i-6}(u, w)\mathsf{P}^6(w, v)\\
&\leq \max_v \sum_w \mathsf{P}^{i-6}(u, w) \max_w \mathsf{P}^6(w, v)\\
&= \max_{w,v} \mathsf{P}^6(w, v) \leq \frac{2520}{(d+1)^6}
\end{aligned}$$

because $\sum_w \mathsf{P}^{i-6}(u, w) = 1$. Hence

$$\begin{aligned}
&\sum_{i=1}^{T} (1 + 2i) \max_{u,v} \mathsf{P}^i(u, v)\\
&\quad \leq \frac{3*1}{d+1} + \frac{5*2}{(d+1)^2} + \frac{7*6}{(d+1)^3} + \frac{9*36}{(d+1)^4}\\
&\qquad + \frac{11*270}{(d+1)^5} + \frac{(T-5)(T+7)2520}{(d+1)^6}\\
&\quad \sim \frac{2520(\ln 2)^2}{(d+1)^2} = o_d(1)\,.
\end{aligned}$$

It remains only to find values for τ and γ in Lemma 3.5. Suppose $(X_0, Y_0) \in \Omega$, i.e. $|X_0 - Y_0| < 2^d$. Then

$$X_T \leq X_0 + T2^d < Y_0 + (T+1)2^d \,.$$

Consider the Y_j walk. In the first $\tau = 2(d+1)(T+1)$ steps the expected number of steps of size 2^d is $\mu = 2(T+1)$, so that $E[Y_\tau - Y_0] \geq 2(T+1)2^d$. With $\delta = 1/2$ then a Chernoff bound implies that

$$\begin{aligned} \mathsf{Pr}\left(Y_\tau < X_T\right) &\leq e^{-\mu\delta^2/2} \\ &\leq e^{-2\ln(2)(d+1)^2/8} \\ &\leq 2^{-(d+1)^2/4} \,. \end{aligned}$$

It thus suffices to take $\tau = 2(d+1)(T+1) \sim (2\ln 2 + o(1))(d+1)^3$, with $\gamma = 2^{-(d+1)^2/4}$ and $\gamma T = o_d(1)$. $\square$

REMARK 4.3. *It is possible to avoid computing powers of* P *at the cost of a weaker bound on the rate of convergence of* B_T *to zero. When* $i \in [1, \sqrt[3]{d+1}]$ *use the bound*

$$\begin{aligned} \max_{u,v} \mathsf{P}^i(u,v) &= \max_{u,v} \sum_w \mathsf{P}^{i-1}(u,w)\mathsf{P}(w,v) \\ &\leq \max_{u,v} \sum_w \mathsf{P}^{i-1}(u,w)\frac{1}{d+1} = \frac{1}{d+1} \end{aligned}$$

When $i > \kappa = \sqrt[3]{d+1}$ *then consider the proof of Lemma 4.1. Let* $\mathcal{R}$ *denote the generators chosen in the first* i *steps, and* $\mathcal{I}_i$ *denote the sum of the increments in the first* i *steps except those the first time a generator was chosen. Then* $\sum_{k\in\mathcal{R}}(\delta_k - 2^k)$ *is a uniform random variable in a set of* $2^{|\mathcal{R}|}$ *possible values, so*

$$\mathsf{Pr}\left(X_i = v \mid \mathcal{R}\right) \leq \max_{\mathcal{I}_i} \mathsf{Pr}\left(X_i = v \mid \mathcal{R}, \mathcal{I}_i\right) \leq 2^{-|\mathcal{R}|} \,.$$

Hence,

$$\begin{aligned} \max_{u,v} \mathsf{P}^i(u,v) &\leq \mathsf{Pr}\left(|\mathcal{R}| \leq \kappa\right) * 1 + \mathsf{Pr}\left(|\mathcal{R}| > \kappa\right) 2^{-\kappa} \\ &\leq \binom{d+1}{\kappa}\left(\frac{\kappa}{d+1}\right)^i * 1 + 1 * \frac{1}{2^\kappa} \\ &\leq (d+1)^\kappa (d+1)^{-2i/3} + 2^{-\kappa} \\ &\leq (d+1)^{-\kappa/3} + 2^{-\kappa} \end{aligned}$$

Then

$$\begin{aligned} B_T &\leq \gamma T + (\tau - T)U(1+\epsilon) \\ &\quad + \sum_{i=1}^{\kappa} \frac{1+2i}{d+1} + \sum_{\kappa+1}^{T} (1+2i)\left((d+1)^{-\kappa/3} + 2^{-\kappa}\right) \\ &= O\left((d+1)^{-1/3}\right) = o_d(1) \end{aligned}$$

We can now prove the main result of the paper.

PROOF OF THEOREM 1.1. Note that the group elements $g^{(2^k)}$ can be pre-computed, so that each step of a kangaroo requires only a single group multiplication.

As discussed in the heuristic argument of Section 2, an average of $\frac{|Y_0 - X_0|}{\sqrt{b-a}/2}$ steps are needed to put the smaller of the starting states (e.g. $Y_0 < X_0$) within 2^d of the one that started ahead. If the Distinguished Points are uniformly randomly distributed then the heuristic for these points is again correct. If instead they are roughly constantly spaced and $c = \Omega(d^2 \log d)$ then observe that, in the proof of Lemma 4.1, it was established that after $T = T(\epsilon) = (d+1)^2 \ln 2 + (d+1)\ln d$ steps the kangaroos will be nearly uniformly random over some interval of length $2^d \sim \frac{1}{4}\sqrt{b-a}\log_2\sqrt{b-a}$; so if the Distinguished Points cover a $\frac{c}{\sqrt{b-a}}$ fraction of vertices, then an average of $\frac{\sqrt{b-a}}{c}$ such samples are needed, which amounts to $T\frac{\sqrt{b-a}}{c} = o_d(1) * \sqrt{b-a}$ extra steps.

It remains to make rigorous the claim regarding p^{-1}. In the remainder we may thus assume that $|Y_0 - X_0| < 2^d$. Take $\epsilon = \frac{3}{d+1} \sim \frac{3}{\log_2\sqrt{b-a}}$. By Lemma 4.1, the uniform intersection time is $T = T(\epsilon) = (d+1)^2\ln 2 + (d+1)\ln d$ with uniform intersection probability $U \sim \frac{2}{\sqrt{b-a}}$, while by Lemma 4.2 also $B_T = o(1)$. The upper bound of Theorem 3.2 is then $\left(\frac{1}{2} + o(1)\right)\sqrt{b-a}$. The lower bound of Theorem 3.2 is then $\left(\frac{1}{2} - o(1)\right)\sqrt{b-a}$. $\square$

Acknowledgments

The authors thank Dan Boneh for encouraging them to study the Kangaroo method, and John Pollard for several helpful comments.

5. REFERENCES

[1] D. Boneh and X. Boyen, "Short Signatures Without Random Oracles," *Proc. of Eurocrypt 2004*, LNCS 3027, pp. 56–73 (2004).

[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," *Proc. of the 8th Algorithmic Number Theory Symposium (ANTS-VIII)*, Springer LNCS vol. 5011, pp. 402–415 (2008).

[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] S. Miller and R. Venkatesan, "Spectral Analysis of Pollard Rho Collisions," *Proc. of the 7th Algorithmic Number Theory Symposium (ANTS-VII)*, Springer LNCS vol. 4076, pp. 573–581 (2006).

[5] P.C. van Oorschot and M.J. Wiener, "Parallel collision search with cryptanalytic applications," *Journal of Cryptology*, vol. 12 no. 1, pp. 1–28 (1999).

[6] J. Pollard, "Monte Carlo methods for index computation mod p," *Mathematics of Computation*, vol. 32 no. 143, pp. 918–924 (1978).

[7] J. Pollard, "Kangaroos, Monopoly and Discrete Logarithms," *Journal of Cryptology*, vol. 13 no. 4, pp. 437–447 (2000).

[8] 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).