How long does it take to catch a wild kangaroo?

Ravi Montenegro ∗ Prasad Tetali †

December 3, 2008

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 ∈ G such that g^x = h. We give the first rigorous proof that Pollard’s Kangaroo method finds the discrete logarithm in expected time \((3 + o(1))\sqrt{b - a}\) when the logarithm \(x \in [a, b]\), and \((2 + o(1))\sqrt{b - a}\) when \(x \in \text{uar} [a, b]\). 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.

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 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 ∈ 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. A Birthday Attack is a common approach towards solving these problems, and although heuristics can be given for the time complexity of these methods, rigorous results are rare.

In [3] we examined one such method, namely Pollard’s Rho Algorithm to find the discrete logarithm on a cyclic group G, and verified the correctness of commonly held intuition. This work generated further interest among some of the experts in cryptography, and Dan Boneh [1] in particular encouraged us to analyze Pollard’s Kangaroo method [5], due to its very many applications. 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, for some cyclic groups the Kangaroo method is the most efficient means known for finding discrete logarithm over an interval, as Shanks baby-step giant-step method requires too much memory.

Among the cases in which this would be useful, Boneh and Boyen [2] give a signature scheme in which a shorter signature can be transmitted if the receiver uses the Kangaroo method to determine the missing information. Verification of the time complexity of the Kangaroo method (as we do

∗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.
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 by the kangaroos before the logarithm is determined. In particular, we find that for the Distinguished Points implementation of the Kangaroo method,

**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 Kangaroo method is

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

If $x \in_{un} [a, b]$ 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 rare among the analyses of algorithms based on Markov chains. Previously the first bound was known only by rough heuristic, while Pollard [6] 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 [5].

Although our approach borrows a few concepts from the study the Rho algorithm in [3], 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 randomized algorithms.

The paper proceeds as follows. In Section 2 we introduce the Kangaroo method. A general framework for analysing 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. The Appendix contains the proof of a technical lemma used in Section 3.

## 2 Preliminaries

We describe here the Kangaroo method, originally known as the Lambda method for catching Kangaroos. The Distinguished Points implementation of [4] is given, rather than the original implementation of [5], as the former is more efficient.

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

**Method:** Pollard’s Kangaroo method (with distinguished points).

**Preliminary Steps:**
Define a set $D$ 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 $d = \log_2 \sqrt{b-a} + \log_2 \log_2 \sqrt{b-a} - 2$ chosen so that elements of $S$ average $\sqrt{b-a}$.

Finally, a hash function $F : G \rightarrow S$.

The Algorithm:

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

The $X_i$ walk is called the “tame kangaroo” because its position is known, whereas the position $Y_j$ 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 $g^{X_i} = g^{Y_j}$ then they proceed along the same route, forming a $\lambda$ shape.

Theorem 1.1 makes rigorous the following commonly used rough heuristic: Suppose $X_0 \geq Y_0$. Run the tame kangaroo infinitely far. Since the kangaroos have an average step size $\sqrt{b-a}$, one expects the wild kangaroo requires $\frac{X_0 - Y_0}{\sqrt{b-a}/2}$ steps to reach $X_0$. 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 $Y_i$ are incremented simultaneously the total number of steps taken is then

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

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

We make only two assumptions in our analysis. First, that the hash $F : G \rightarrow S$ is a random function, i.e. if $g \in G$ then $F(g)$ is equally likely to be any value in $S$, independent of all other $F(g')$. Second, that the distinguished points are well distributed with $c \overset{(b-a)\rightarrow\infty}{\sim} \infty$; either they are chosen uniformly at random, or if $c = \Omega(d^2 \log^2 d)$ then roughly constant spacing between points will suffice. The assumption on distinguished points can be dropped if we instead analyze Pollard’s (slower) original algorithm, to which our methods also apply. Both assumptions are made in most discussions of the Kangaroo method [7 4 6], and so are quite acceptable.
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 by the assumption about the hash \( F : G \to S \) that \( X_i \) 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 “tame kangaroo” (described by \( X_i \)) takes until it meets the “wild kangaroo” (i.e. \( X_i = Y_j \) on \( \mathbb{Z} \)), it suffices to run the wild kangaroo infinitely long and only after this have the tame 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 some \( N \) steps until \( \exists j, X_N = Y_j \). Thus, rather than considering a specific probability \( \Pr[X_i = Y_j] \) it is better to look at \( \Pr[\exists j, X_i = Y_j] \). By symmetry, the same approach will also bound the expected number of hops the wild kangaroo requires to reach a location the tame kangaroo visits.

First however, because the walk does not proceed long enough to approach its stationary distribution (true on both \( \mathbb{Z}/|G|\mathbb{Z} \) and more obviously on \( \mathbb{Z} \)), alternate notions resembling mixing time and a stationary distribution will be required.

**Definition 3.1.** Consider a Markov chain \( P \) on \( \mathbb{Z} \) which is non-decreasing, i.e. \( P(u, v) > 0 \) only when \( v \geq u \). Let \( X_i \) and \( Y_j \) denote independent walks starting at states \( (X_0, Y_0) \in \Omega \subseteq \mathbb{Z} \times \mathbb{Z} \), for some set of permitted initial states \( \Omega \supset \cup_{v \in \mathbb{Z}} \{(v, v)\} \). For fixed \( \epsilon \in [0, 1] \), the uniform intersection time \( T(\epsilon) \in \mathbb{N} \) and uniform intersection probability \( U \in \mathbb{R}^+ \) are such that

\[
\forall i \geq T(\epsilon) : (1 - \epsilon)U \leq \Pr[\exists j, X_i = Y_j] \leq (1 + \epsilon)U .
\]

We do not attempt to show a general existence result for uniform intersection time and probability, as our primary interest is in the Kangaroo method. Also, to avoid clutter we write \( T(\epsilon) \) in the remainder.

A natural approach 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 \( \Pr[S_N > 0] \) is non-trivial for some \( N \). Our collision bound will involve the quantity \( B_T \), an upper bound on the expected number of collisions in the first \( T \) steps between two independent walks. To be precise, define:

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

Then the expected number of steps until a collision can be bounded as follows.
Theorem 3.2. Consider a non-decreasing Markov chain on \( \mathbb{Z} \), two independent walks with starting states \((X_0, Y_0) \in \Omega\), and uniform intersection time and probability \(T = T(\epsilon)\) and \(U\) respectively. Then
\[
\frac{1 - 2\sqrt{B_T}}{U(1 + \epsilon)} - T \leq \mathbb{E}\min\{i > 0 : \exists j, X_i = Y_j\} \leq (1 - 4\epsilon)^{-1} \left( \sqrt{T} + \sqrt{\frac{1 + 2B_T}{U}} \right)^2.
\]
If \(B_T, \epsilon \approx 0\) and \(U^{-1} \gg T\) then these bounds show that
\[
\mathbb{E}\min\{i > 0 : \exists j, X_i = Y_j\} \sim \frac{1}{U}.
\]
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.
\[
S_N = \sum_{i=T+1}^{N} 1_{\{\exists j: X_i = Y_j\}},
\]
in terms of the uniform intersection time and probability:

Lemma 3.3. Under the conditions of Theorem 3.2, if \(N \geq T\) then
\[
(1 - \epsilon)(N - T)U \leq E[S_N] \leq (1 + \epsilon)(N - T)U,
\]
\[
E[S_N^2] \leq (1 + \epsilon)^2(N - T)^2U^2 \left[ 1 + \frac{1 + 2B_T}{(N - T)U} \right].
\]

This is a technical lemma and offers little insight into our proof, so it is left for the Appendix.

We now upper and lower bound the probability of an intersection in the first \(N\) steps:

Lemma 3.4. Under the conditions of Theorem 3.2, if \(N \geq T\) then
\[
B_T + (N - T)U(1 + \epsilon) \geq \text{Pr}[S_N > 0] \geq (1 - 4\epsilon) \left[ 1 + \frac{1 + 2B_T}{(N - T)U} \right]^{-1}.
\]

Proof. Observe that \(\text{Pr}[S_N > 0] \geq \text{Pr}[S_N > 0]\), so for the lower bound it suffices to consider \(S_N\).
Recall the standard second moment bound: using Cauchy-Schwartz, we have that
\[
E[S_N] = E[S_N 1_{\{S_N > 0\}}] \leq E[S_N^2]^{1/2} E[1_{\{S_N > 0\}}]^{1/2}
\]
and hence \(\text{Pr}[S_N > 0] \geq E[S_N^2]/E[S_N^2]\). By Lemma 3.3 then, independent of starting point,
\[
\text{Pr}[S_N > 0] \geq \left( \frac{1 - \epsilon}{1 + \epsilon} \right)^2 \left[ 1 + \frac{1 + 2B_T}{(N - T)U} \right]^{-1} \geq (1 - 4\epsilon) \left[ 1 + \frac{1 + 2B_T}{(N - T)U} \right]^{-1},
\]
since \(\left( \frac{1 - \epsilon}{1 + \epsilon} \right)^2 \geq 1 - 4\epsilon, \text{ for } \epsilon \geq 0\).

Now to upper bound \(\text{Pr}[S_N > 0]\). Since \(S_N \in \mathbb{N}\) then
\[
\text{Pr}[S_N > 0] = E[1_{S_N > 0}] \leq E[S_N].
\]
The expectation $E[S_N]$ satisfies

$$E[S_N] = E\sum_{i=1}^{N}1_{\{3j, X_i=Y_j\}} = \sum_{i=1}^{N}E[1_{\{3j, X_i=Y_j\}}]$$

$$= \sum_{i=1}^{T}\Pr[\exists j, X_i=Y_j] + \sum_{i=T+1}^{N}\Pr[\exists j, X_i=Y_j]$$

$$\leq B_T + (N-T)U(1+\epsilon).$$

\[\square\]

**Proof of Theorem 3.2.** First, we upper and lower bound $\Pr[S_{kN} = 0]$ for every $k \geq 1$. For $\ell \geq 1$, let

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

so that $S_{N}^{(1)} = S_N$. Thus

$$\Pr[S_{kN} = 0] = \Pr[S_N = 0]Pr[S_{2N} = 0 \mid S_N = 0] \cdots Pr[S_{kN} = 0 \mid S_{(k-1)N} = 0]$$

$$= \prod_{\ell=1}^{k}Pr[S_{N}^{(\ell)} = 0 \mid S_{(\ell-1)N} = 0]$$

By taking $X_0 \leftarrow X_{(\ell-1)N}$ and $Y_0 \leftarrow \min\{Y_j : Y_j > X_{(\ell-1)N}\}$, we may bound:

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

Hence

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

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

$$E \min\{i : S_i > 0\} = E\sum_{i=0}^{\infty}1_{S_i=0} = 1 + \sum_{i=0}^{\infty}\Pr[S_i = 0] \leq \sum_{k=0}^{\infty}\Pr[S_{kN} = 0].N$$

$$\leq N\sum_{k=0}^{\infty} \left(1 - (1 - 4\epsilon) \left[1 + \frac{1+2B_T}{(N-T)U}\right]^{-1}\right)^k$$

$$= (1 - 4\epsilon)^{-1}N \left(1 + \frac{1+2B_T}{(N-T)U}\right).$$

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
\[ E \min\{i : S_i > 0\} = \sum_{i=0}^{\infty} \Pr[S_i = 0] \geq \sum_{k=1}^{\infty} \Pr[S_{kN} = 0]N \]
\[ \geq N \sum_{k=1}^{\infty} (1 - BT - (N - T)U(1 + \epsilon))^k \]
\[ = N \left( \frac{1}{BT + (N - T)U(1 + \epsilon)} - 1 \right). \]

If \( BT \geq 1 \) then the bound stated in the theorem is trivial, so assume \( BT < 1 \).
If \( BT(1 - BT) < TU(1 + \epsilon) \) then the maximum of the above bound is at \( N = T \). In this case the bound is
\[ E \min\{i : S_i > 0\} \geq N \left( \frac{1}{BT} - 1 \right) \geq \frac{1 - BT}{U(1 + \epsilon)} - T. \]

When \( BT(1 - BT) \geq TU(1 + \epsilon) \) then the maximum is at \( N = \frac{\gamma(1 - \gamma)}{U(1 + \epsilon)} \), where the quantity \( \gamma = \sqrt{BT - TU(1 + \epsilon)} \). In this case the bound is
\[ E \min\{i : S_i > 0\} \geq \left( \frac{1 - \sqrt{BT - TU(1 + \epsilon)}}{U(1 + \epsilon)} \right)^2 \geq \frac{(1 - \sqrt{BT})^2}{U(1 + \epsilon)}. \]

To bound the value of \( BT \) 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 \( \tau \geq T \) be such that whenever \( (X_0, Y_0) \in \Omega \) then
\[ \Pr[\{X_i\}_{i=1}^{T} \cap \{Y_j\}_{j=\tau} \neq \emptyset] \leq \gamma. \]

Then
\[ BT \leq \gamma T + 2 \sum_{j=1}^{\tau} (1 + j) \max_{u,v} P^j(u, v). \]

**Proof.** Recall that
\[ BT = \max_{(X_0, Y_0) \in \Omega} \sum_{i=1}^{T} \Pr[\exists j, X_i = Y_j]. \]

When \( j > \tau \) then
\[ \sum_{i=1}^{T} \Pr[\exists j > \tau : X_i = Y_j] \leq T \Pr[\{X_i\}_{i=1}^{T} \cap \{Y_j\}_{j=\tau} \neq \emptyset] \leq \gamma T. \]

When \( j \leq \tau \) then
\[ \sum_{i=1}^{T} \Pr[\exists j \leq \tau : X_i = Y_j] \leq \sum_{i=1}^{T} \sum_{j=0}^{\tau} \sum_{v} P^i(X_0, v)P^j(Y_0, v) \]
\[ \leq 2 \sum_{j=1}^{\tau} \sum_{i=0}^{j} \max_{w,x} P^j(w, x) \max_{u} \sum_{v} P^i(u, v) \]
\[ = 2 \sum_{j=1}^{\tau} (1 + j) \max_{w,x} P^j(w, x). \]
The second inequality follows by letting \( j \) denote the larger of the two indices and \( i \) the smaller. The final equality is because \( \sum_v P^i(u,v) = 1 \).

4 Catching Kangaroos

The collision results of the previous section will now be applied to the Kangaroo method. The first step in bounding collision time will be to bound the uniform intersection time and probability. This will be done by selecting some \( d \) of the first \( T \) steps of the \( X_i \) walk (for suitable \( i \)), 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 the \( X_i \) walk is uniformly distributed over some interval of \( 2^d \) elements, and so the probability that some \( Y_j = X_i \) will be exactly the expected number of times the \( Y_j \) walk visits this interval, divided by the interval size (i.e. \( 2^d \)).

Throughout we take

\[
\Omega = \{ (X_0, Y_0) : X_0 \leq Y_0 < X_0 + 2^d \}.
\]

**Lemma 4.1.** If \((X_0, Y_0) \in \Omega \) and \( i \geq T = 2(d+1)^2(1 + \log_2(d+1)) \) then

\[
\left| \frac{\Pr(\exists j, X_i = Y_j)}{2/\sqrt{b-a}} - 1 \right| \leq \frac{3}{\log_2 \sqrt{b-a}} \sim \frac{3}{d},
\]

i.e. when \( \epsilon = \frac{3}{d} \) then one may take \( T(\epsilon) = T \) as above and \( U = \frac{2}{\sqrt{b-a}} \).

**Proof.** The tame kangaroo will be implemented by choosing \( k \in \text{uar} \{0,1,\ldots,d\} \) 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 may not be \( 2^k \).

Consider the tame kangaroo. For \( k \in \{0,1,\ldots,d-1\} \) let \( \delta_k \) denote the step taken the first time generator \( 2^k \) is chosen, so that \( \delta_k - 2^k \in \text{uar} \{0,2^k\} \). Also, let \( T \) be the first time all of the generators have been chosen (including \( 2^d \)). Define \( \delta = \sum_{k=0}^{d-1}(\delta_k - 2^k) \in \text{uar} \{0,1,\ldots,2^d - 1\} \) and let \( I_i \) denote the sum of all increments except those incorporated in a \( \delta_k \), so that if \( i \geq T \) then \( X_i = X_0 + I_i + 2^d - 1 + \delta \).

Suppose \( i \geq T \). Then \( \delta \) is independent of the value of \( I_i \), and so

\[
X_i \in \text{uar} \{X_0 + I_i + 2^d - 1, X_0 + I_i + 2^{d+1} - 1\}.
\]

Observe that \( X_0 + I_i + 2^d - 1 \geq X_0 + 2^d - 1 \geq Y_0 \). Since the average non-zero step size for \( Y_j \) is \( \frac{\sqrt{b-a}}{2} \) (recall \( d \) was chosen to guarantee this) then

\[
\Pr(\exists j, X_i = Y_j \mid i \geq T)
= \frac{\mathbb{E}[\{Y_j\} \cap \{X_0 + I_i + 2^d - 1, X_0 + I_i + 2^{d+1} - 1\}]}{2^d}
\geq \frac{[2^d/(\frac{1}{2}\sqrt{b-a})]}{2^d}
= \frac{2}{\sqrt{b-a}} - 2^{-d} = \frac{2}{\sqrt{b-a}} \left(1 - \frac{2}{\log_2 \sqrt{b-a}} \right).
\]

Similarly, an upper bound of \( \frac{2}{\sqrt{b-a}} + 2^{-d} \) follows by taking ceiling instead of floor.

Next, consider \( T \). By the Coupon Collector’s problem \( E(T) = (d+1)H_{d+1} \) where \( H_n = \sum_{i=1}^{n} \ell^{-1} \) is the \( n \)-th harmonic number. By Markov’s Inequality \( \Pr(T \geq 2(d+1)H_{d+1}) \leq 1/2 \) and
so if $\alpha = d + 1$ then

\[
\Pr [T \geq \alpha 2(d + 1)H_{d+1}] = \prod_{\ell=1}^{\alpha} \Pr [T \geq \ell 2(d + 1)H_{d+1} | T \geq (\ell - 1)2(d + 1)H_{d+1}] \\
\leq 2^{-\alpha} = 2^{-(d+1)}.
\]

Since $H_n \leq \ln n + 0.6 + \frac{1}{2n}$, we get in turn $\Pr[T > T] < 2^{-(d+1)}$. To finish,

\[
\Pr[\exists j, X_i = Y_j] = (1 - \Pr[T > T])\Pr[\exists j, X_i = Y_j | T \leq T] + \Pr[T > T] \Pr[\exists j, X_i = Y_j | T > T].
\]

Since all probabilities are in $[0, 1]$, and $0 \leq \Pr[T > T] < 2^{-(d+1)}$ then

\[
|\Pr[\exists j, X_i = Y_j] - \Pr[\exists j, X_i = Y_j | T \leq T]| < 2^{-(d+1)}.
\]

It remains only to upper bound $B_T$. This will be shown by breaking up the sum of Lemma 3.5 into two parts. Let $\kappa = \sqrt{d+1}$. When $j \leq 2\kappa$ then it will be shown that with high probability every step size taken was distinct, in which case the sum of the step sizes is a random $(d+1)$ bit binary string containing exactly $j$ ones, i.e. uniform over $\binom{d+1}{j}$ possibilities. When $j > 2\kappa$ then with high probability at least $\kappa$ distinct step sizes have been chosen, in which case a random $\kappa$-bit binary string is extracted as in the proof of Lemma 4.1 and used to show the maximum probability of a state is at most $2^{-\kappa}$.

**Lemma 4.2.** If $T = 2(d + 1)^2(1 + \log_2(d + 1))$ then $B_T = o_d(1)$.

**Proof.** This will be shown by applying Lemma 3.5. To bound $\mathbb{P}^j(u, v)$ we set $X_0 = u$ and consider $X_j$ for $j \in \{1, 2, \ldots, \tau\}$, where $\tau$ is to be determined later. Recall that $\kappa = \sqrt{d+1}$.

First suppose $1 \leq j \leq 2\kappa$. Assume $d \geq 2$ so that $j < d + 1$. Implement the kangaroo walk in the obvious way, i.e. choose $k \in \{0, 1, \ldots, d\}$ and increment by $2^k$. Let $\mathcal{E}$ denote the event that all $j$ increment sizes were distinct. Then

\[
\Pr[\mathcal{E}] = \frac{(d+1)d \cdots (d + 2 - j)}{(d+1)^j} \geq \left(\frac{d+2-j}{d+1}\right)^j \geq 1 - \frac{j(j-1)}{d+1},
\]

because $(1 - x)^n \geq 1 - nx$ if $x \in [0, 1]$ and $n \in \mathbb{N}$. Then

\[
\max_v \mathbb{P}^j(u, v) = \max_v \Pr[X_j = v] = \max_v \Pr[\mathcal{E}] \Pr[X_j = v | \mathcal{E}] + \Pr[\overline{\mathcal{E}}] \Pr[X_j = v | \overline{\mathcal{E}}] \\
\leq 1 + \frac{1}{d+1} + (\Pr[\overline{\mathcal{E}}]) * 1 \\
\leq \frac{1}{d+1} + \frac{j(j-1)}{d+1} \leq 4(d + 1)^{-3/5}
\]

If $d = 1$ then trivially $\mathbb{P}^j(u, v) \leq 4(d + 1)^{-3/5} \approx 2.64$. Then

\[
\sum_{j=1}^{2\kappa} (j + 1) \max_v \mathbb{P}^j(u, v) \leq (1 + 2\kappa)^2 4(d + 1)^{-3/5} = o_d(1).
\]
Before calculating the remaining terms in the sum, a value for $\tau$ in Lemma 3.5 is needed. Note that trivially
$$X_T \leq X_0 + T 2^d.$$ 
Let $\Delta Y$ be a random increment of the $Y$ walk. Then
$$\Pr \left[ \Delta Y \geq \frac{2^d}{d+1} \right] = \frac{1 + \log_2(d+1)}{d+1}.$$ 
A Chernoff bound can be used here. If $j \geq \frac{2T(d+1)^2}{1 + \log_2(d+1)}$ the expected number of steps of size at least $\frac{2^d}{d+1}$ is $\mu = 2T(d+1)$ so that $E[Y_j - X_0] \geq E[Y_j - Y_0] \geq 2T 2^d$. With $\delta = 1/2$ then
$$\Pr[Y_j \leq X_0 + T 2^d] \leq e^{-\mu \delta^2/2} \leq e^{-4(d+1)^3(1+\log_2(d+1))/8} \leq 2^{-(d+1)}.$$ 
It thus suffices to take $\tau = \frac{2T(d+1)^2}{1 + \log_2(d+1)} = 4(d+1)^4$, with $\gamma = 2^{-(d+1)}$ and $\gamma T = o_d(1)$.

Finally, suppose $2\kappa < j \leq \tau$. Implement the kangaroo walk as in the proof of Lemma 4.1, and likewise assume the same terminology. Let $S$ denote the set of distinct generators that have been chosen excluding $2^d$, so that $|S| \leq d$, and observe that $\sum_{k \in S} (\delta_k - 2^k) \in \text{var} \{2^{|S|} \text{ elements}\}$, so that if $T_j$ is the sum of all increments except those used the first time an element of $S$ was chosen then $\Pr[X_j = v \mid S, T_j] \leq 2^{-|S|}$. It follows that $\Pr[X_j = v \mid |S|] \leq 2^{-|S|}$. Hence, if $E$ denotes the event that $\kappa$ or fewer distinct generators have been chosen, so that $E$ implies $|S| \geq \kappa$, then
$$\max_{u,v} P^j(u,v) \leq \Pr[E] * 1 + \Pr[E] \frac{1}{2\kappa}$$ 
$$\leq \left( \frac{d+1}{\kappa} \right) \left( \frac{\kappa}{d+1} \right)^j * 1 + 1 * \frac{1}{2\kappa}$$ 
$$\leq (d+1)^\kappa (d+1)^{-4j/5} + 2^{-\kappa}$$ 
$$\leq (d+1)^{-\frac{2}{5}\kappa} + 2^{-\kappa}$$ 
It follows that
$$\sum_{j=2\kappa+1}^{\tau} (1 + j) \max_{u,v} P^j(u,v) \leq (1 + \tau)^2 \left( (d+1)^{-\frac{2}{5}\kappa} + 2^{-\kappa} \right) = o_d(1).$$

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 $\frac{|X_0 - Y_0|}{\sqrt{b-a}/2}$ steps are needed to put the smaller of the starting states (e.g. $X_0 < Y_0$) within $2^d$ of the one that started ahead. If the Distinguished Points are randomly distributed then the heuristic for these points is again correct. If instead they are roughly constantly spaced and $c = \Omega(d^2 \log^2 d)$ then observe that in the proof of Lemma 4.1 it was established that after $T = T(\epsilon) = 2(d+1)^2(1 + \log_2(d+1))$ steps the kangaroos will be nearly uniformly random over some interval of length $2^{d+1} = \frac{1}{2} \sqrt{b-a} \log_2 \sqrt{b-a}$, so if the Distinguished Points are uniformly distributed and cover a $\frac{1}{\sqrt{b-a} c}$ fraction of vertices then an average of $\frac{\sqrt{b-a}}{c}$ such samples are needed, which amounts to $T \frac{\sqrt{b-a}}{c} = o(1) * \sqrt{b-a}$ extra steps.
It remains to make rigorous the claim regarding $p^{-1}$. In the remainder we may thus assume that $|X_0 - Y_0| \leq 2^d$. Take $\epsilon = \frac{3}{\log_2 \sqrt{b-a}} \sim \frac{3}{d}$. By Lemma 4.1 the uniform intersection time is $T = T(\epsilon) = 2(d + 1)^2(1 + \log_2(d + 1))$ with uniform intersection probability $U = \frac{2}{\sqrt{b-a}}$, while by Lemma 4.2 also $B_T = o(1)$. The upper bound of Theorem 3.2 is then $(\frac{1}{2} + o(1)) \sqrt{b-a}$. The lower bound of Theorem 3.2 is then $(\frac{1}{2} - o(1)) \sqrt{b-a}$.

Acknowledgments

The authors thank Dan Boneh for encouraging them to study the Kangaroo method.

References

[1] D. Boneh, Private Communication.

[2] D. Boneh and X. Boyen, “Short Signatures Without Random Oracles,” Proc. of Eurocrypt 2004, LNCS 3027, pp. 56-73 (2004).

[3] 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 Algorithmic Number Theory Symposium (ANTS-VIII), Springer LNCS 5011, pp. 402-415 (2008).

[4] 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).

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

[6] J. Pollard, “Kangaroos, Monopoly and Discrete Logarithms," Journal of Cryptology, vol. 13 no. 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).

Appendix

The proof of Lemma 3.3 was left to the Appendix:

**Proof.** The expectation $E[S_N]$ satisfies

$$E[S_N] = E \sum_{i=T+1}^{N} 1_{\exists j, X_i = Y_j} = \sum_{i=T+1}^{N} E[1_{\exists j, X_i = Y_j}] \geq (N - T)U(1 - \epsilon). \quad (1)$$

The inequality is because $E[1_{\exists j, X_i = Y_j}] = \Pr[\exists j, X_i = Y_j]$. The upper upper bound follows by taking $(1 + \epsilon)$ in place of $(1 - \epsilon)$.
Now for $E[S^2_N]$. Note that

$$E[S^2_N] = \sum_{i=T+1}^{N} \sum_{k=T+1}^{N} \mathbb{1}_{\exists j, X_i = Y_j, 1 \exists \ell, X_k = Y_\ell}$$

$$= \sum_{i=T+1}^{N} \sum_{k=T+1}^{N} \Pr(\exists j, \ell : X_i = Y_j, X_k = X_\ell).$$

By symmetry it suffices to consider the case that $k \geq i > T$. Also, observe that if $X_i = Y_j$ then $X_k = Y_\ell$ is possible only if $\ell \geq j$, because the $X$ and $Y$ walks proceed in the positive direction on the integer line.

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

$$\Pr(\exists j, \ell : X_i = Y_j, X_k = Y_\ell) = \Pr(\exists j : X_i = Y_j) \Pr(\exists \ell, X_k = Y_\ell \mid X_i = Y_j) \leq (1 + \epsilon)^2 U.$$ 

When $k \leq i + T$ then

$$\sum_{k=i+1}^{i+T} \Pr(\exists j, \ell : X_i = Y_j, X_k = Y_\ell) \leq \Pr(\exists j : X_i = Y_j) \max_u \sum_{k=1}^{T} \Pr(\exists \ell, X_k = Y_\ell \mid X_0 = Y_0 = u) \leq BT U(1 + \epsilon),$$

since $i \geq T$. It follows that

$$E[S^2_N] = \sum_{i=T+1}^{N} \left( \Pr(\exists j : X_i = Y_j) + 2 \sum_{k=i+1}^{i+T} \Pr(\exists j, \ell : X_i = Y_j, X_k = X_\ell) \right)$$

$$+ 2 \sum_{i=T+1}^{N} \sum_{k=i+T+1}^{N} \Pr(\exists j, \ell : X_i = Y_j, X_k = X_\ell) \leq 2(1 + \epsilon)U(N - T)(1/2 + BT) + 2(1 + \epsilon)^2 U^2 (N - 2T + 1)$$

$$\leq (1 + \epsilon)^2 U^2 (N - T)^2 \left[ 1 + \frac{1 + 2BT}{(1 + \epsilon)U(N - T)} \right].$$

\[\square\]