A paradox for expected hitting times

Mark Holmes* and Peter Taylor†

February 20, 2019

Abstract

We prove a counterintuitive result concerning the expected hitting/absorption time for a class of Markov chains. The "paradox" already shows itself in the following elementary example that is suitable for undergraduate teaching:

Batman and the Joker perform independent discrete-time random walks on the vertices of a square until they meet, starting from opposite vertices. Batman always moves, while the Joker remains still on any given step with probability $q \in [0, 1]$, and clockwise and anticlockwise moves are equally likely for both. On average the Joker survives for twice as long by staying still with arbitrarily small but positive probability (i.e. $\lim_{q \downarrow 0}$) than by always moving (i.e. $q = 0$).

1 Introduction

sec:basic

There are various "paradoxes" or counterintuitive results that have entered the probability and statistics folklore. Elementary examples include the so-called *birthday paradox*, *Simpson's paradox*, *non-transitive dice*, the *inspection paradox* (biased sampling), the *St. Petersburg paradox* and when correctly posed, the *Monty Hall problem*. Each of these can be presented in a first course in probability theory, and can often be found in standard undergraduate probability texts GW,Ross,Feller1,Feller2 [3, 6, 4, 5]. Some paradoxes such as *two-envelope*

*School of Mathematics and Statistics, University of Melbourne. holmes.m@unimelb.edu.au

†School of Mathematics and Statistics, University of Melbourne. taylorpg@unimelb.edu.au

problems (see e.g. envelopes [2]) and *Bertrand's paradox* (see e.g. GW [3]) arise simply because the original problem as stated is not well-posed.

In this paper we describe a probabilistic paradox that we have not found elsewhere in the literature. The paradox concerns a kind of discontinuity in expected hitting times for certain finite discrete-time Markov chains, and it arose from the following simple question, set by the first author as an assignment problem in an undergraduate course at the University of Melbourne:

exa:bat_simple **Example 1** (Simple Batman and Joker on a square)**.** Batman and the Joker each walk at random on the vertices of a square. At each time step, Batman takes one step clockwise with probability $1/2$, and one step anticlockwise with probability $1/2$, while the Joker stays where he is with probability $q \in [0,1]$, steps clockwise with probability $(1-q)/2$ and anticlockwise with probability $(1-q)/2$. Batman catches the Joker if they are located at the same vertex at the same time. All steps are taken independently of previous steps and of the steps of the other character.

Problem: *Starting from opposite corners of the square, find the expected time until Batman catches the Joker (Batman catches the Joker when they occupy the same vertex at the same time).*

Solution: When $q=0$, the time T until Batman catches the Joker has a Geometric(1/2) distribution, so $\mathbb{E}_0[T]=2$. When $q>0$ it is natural and standard to consider the clockwise graph distance X_n from Batman to the Joker after n time steps. This sequence $\boldsymbol{X}=(X_n)_{n\geq 0}$ is a Markov chain with state space $\mathcal{S}=\{0,1,2,3\}$, initial state $X_0=2$, and single-step transition probabilities $p_{i,(i+1) \bmod 4}=q/2=p_{i,(i-1) \bmod 4}$, $p_{i,i}=(1-q)/2=p_{i,(i+2) \bmod 4}$. The solution $m_2=\mathbb{E}_q[T]$ is found by solving the system of linear equations (see e.g. GW [3, Theorem 12.57])

$$
\begin{aligned}
m_2 &= 1+\frac{q}{2}m_1+\frac{q}{2}m_3+\frac{1-q}{2}m_2\\
m_1 &= 1+\frac{q}{2}m_2+\frac{1-q}{2}m_1+\frac{1-q}{2}m_3\\
m_3 &= 1+\frac{q}{2}m_2+\frac{1-q}{2}m_3+\frac{1-q}{2}m_1, \qquad (1)
\end{aligned}
$$
m_eqs

where m_i is the expected hitting time of 0, starting from state i (note that $m_1=m_3$ by symmetry). Solving gives $m_2=4$ for each $q>0$.

The above example has two interesting features. Observe first that for $q>0$ the answer does not depend on q. This is an amusing consequence of the fact that the protagonists are walking on a square, and it will not hold more generally. The second observation is the discontinuity at $q = 0$, which we state as follows.

> *The Joker survives for twice as long, on average, if at each time step he chooses to stay where he is with arbitrarily small (but positive) probability rather than zero probability.*

That we should observe a discontinuity at $q = 0$ in this problem is not surprising. When q is very small, it is extremely likely that the Joker will get caught before he ever "stays still", however in the very unlikely event that he does stay still before being caught, it takes an enormous amount of time thereafter for the Joker to be caught, since this can only happen after the Joker has stayed still for a second time. What is interesting is that the "twice as long" phenomenon turns out to be remarkably general. For example, if Batman and the Joker each move on a (finite) bipartite graph (starting at distinct vertices of the same parity), with their own vertex specific transition probabilities, then the expected time until the Joker is captured will depend on the graph, the transition probabilities, and the starting locations. Nevertheless, as long as these are fixed, it remains true that the Joker survives for twice as long (on average) by taking $q > 0$ but minuscule instead of $q = 0$.

In principle, for any given finite graph one could verify this by solving a model-specific system of linear equations similar to (1), but in practice (except for some special cases, such as cycles of length $2k$ with constant clockwise step probabilities, which reduce to variants of the *Gambler's ruin* hitting time problem) for large graphs this is intractable.

2 The main result

Our main result will apply to a much more general class of models. To state the result, let $\mathcal{S}^{(0)}$ be a finite set and $\mathcal{S}^{(1)}$ be a finite or countably infinite set disjoint from $\mathcal{S}^{(0)}$. Let $A \subsetneq \mathcal{S}^{(0)}$ be a non-empty subset. Let $P^{(0)} = (p^{(0)}_{i,j})_{i,j} \in \mathcal{S}^{(0)}$ be the transition probabilities of an irreducible Markov chain on $\mathcal{S}^{(0)}$. Let $P^{(1)} = (p^{(1)}_{i,j})_{i,j} \in \mathcal{S}^{(1)}$ be the transition probabilities of a Markov chain on $\mathcal{S}^{(0)}$. For $i \in \mathcal{S}^{(k)}$ $(k = 0, 1)$ let $\boldsymbol{p_i} = (p_i(j))_{j \in \mathcal{S}^{(1-k)}}$ be a probability distribution supported on $\mathcal{S}^{(1-k)}$. We write $i \sim j$ if $i, j \in \mathcal{S}^{(0)}$ or $i, j \in \mathcal{S}^{(1)}$, and $i \nsim j$ otherwise. Let $c_0, c_1 > 0$ be fixed constants. Let

$q \in [0, \max\{1/c_0, 1/c_1\})$, and let $\boldsymbol{X}(q)$ be the Markov chain on $\mathcal{S} = \mathcal{S}^{(0)} \cup \mathcal{S}^{(1)}$ starting at $s_0 \in \mathcal{S}_0$ with transition probabilities

$$p_{i,j} = \begin{cases} c_0 q p_i(j), & \text{if } i \in \mathcal{S}^{(0)}, j \in \mathcal{S}^{(1)} \\ c_1 q p_i(j), & \text{if } i \in \mathcal{S}^{(1)}, j \in \mathcal{S}^{(0)} \\ (1 - c_0 q) p^{(0)}_{i,j}, & \text{if } i, j \in \mathcal{S}^{(0)} \\ (1 - c_1 q) p^{(1)}_{i,j}, & \text{if } i, j \in \mathcal{S}^{(1)}. \end{cases} \tag{2}$$

Thus, our chain behaves like a $P^{(k)}$ chain while moving on $\mathcal{S}^{(k)}$ but switches from $\mathcal{S}^{(k)}$ to $\mathcal{S}^{(1-k)}$ at Geometric($c_k q$) times (independent of the past). If $c_0 = c_1 = 1$ and $q = 1$ we switch each time, and this reduces to a Markov chain on $\mathcal{S}$ with transition probabilities $(p_i(j))_{i \nsim j}$ starting from s_0. When $q = 0$ we never switch, and this reduces to a Markov chain on $\mathcal{S}^{(0)}$ with transition probabilities $P^{(0)}$.

Let $T(q)$ denote the hitting time of A by the chain $\boldsymbol{X}(q)$. Our main result is the following.

thm **Theorem 1.** *For any* $s_0 \neq 0$,

$$\lim_{q \downarrow 0} \mathbb{E}[T(q)] = \left[1 + \frac{c_0}{c_1}\right] \mathbb{E}[T(0)]. \tag{3}$$

Before we discuss this further, let us give a sketch proof which demonstrates where this result comes from.

Sketch proof of Theorem thm *1.* Let $\mathcal{T}(q)$ be the time that the process first jumps from $\mathcal{S}^{(0)}$ to $\mathcal{S}^{(1)}$. Then $\mathcal{T}(q)$ has a Geometric($c_0 q$) distribution. Since when $q = 0$ the chain just moves on $\mathcal{S}^{(0)}$ according to $P^{(0)}$, we can couple the processes for all (small) values of q such that $\boldsymbol{X}(q)$ makes the same moves as $\boldsymbol{X}(0)$ up to time $\mathcal{T}(q)$. Thus $\{\mathcal{T}(q) \leq T(q)\} = \{\mathcal{T}(q) \leq T(0)\}$ and on the event $\{T(q) < \mathcal{T}(q)\}$ we have that $T(q) = T(0)$. It follows that

$$\begin{aligned}
\mathbb{E}[T(q)] &= \mathbb{E}[T(q)\mathbb{1}_{\{T(q)<\mathcal{T}(q)\}}] + \mathbb{E}[T(q)|T(q) \geq \mathcal{T}(q)]\mathbb{P}(T(q) \geq \mathcal{T}(q)) \\
&= \mathbb{E}[T(0)\mathbb{1}_{\{T(0)<\mathcal{T}(q)\}}] + \mathbb{E}[T(q)|T(q) \geq \mathcal{T}(q)]\mathbb{P}(T(0) \geq \mathcal{T}(q)) \\
&= \mathbb{E}[T(0)\mathbb{1}_{\{T(0)<\mathcal{T}(q)\}}] \qquad (4) \\
&\quad + \mathbb{E}[T(q)|T(q) \geq \mathcal{T}(q)] \sum_{n=1}^{\infty} c_0 q(1-c_0 q)^{n-1}\mathbb{P}(T(0) \geq n) \qquad (5) \\
&= \mathbb{E}[T(0)] + o(1) + \left(\frac{1}{c_1 q} + O(1)\right) c_0 q(\mathbb{E}[T(0)] + o(1)) \qquad (6) \\
&= \left[1 + \frac{c_0}{c_1}\right] \mathbb{E}[T(0)] + o(1), \qquad (7)
\end{aligned}$$

where $o(1)$ is as $q \downarrow 0$. What makes this proof a sketch is the claim that $\mathbb{E}[T(q)|T(q) \geq \mathcal{T}(q)] = \frac{1}{c_1 q} + O(1)$, where the $\frac{1}{c_1 q}$ term is the expected time for the chain to come back to $\mathcal{S}^{(0)}$. ■

The above sketch proof indicates that the assumptions on the sequences $\boldsymbol{X}(q)$ can be further relaxed. Roughly speaking, the claim of the theorem holds as long as $\boldsymbol{X}(q)$ behaves like a $P^{(0)}$ Markov chain on $\mathcal{S}^{(0)}$, jumps out of $\mathcal{S}^{(0)}$ at Geometric$(c_0 q)$ times (independent of the history), and then jumps back in at Geometric$(c_1 q)$ times (independent of the history). The exact details of what state the sequence re-enters $\mathcal{S}^{(0)}$ in each time is immaterial (hence, the details of what the process does in $\mathcal{S}^{(1)}$ is also immaterial). For example, the statement holds if we choose a deterministic sequence $(s_i)_{i\in\mathbb{Z}_+} \subset \mathcal{S}^{(0)}$ of re-entry states (s_i is the state in which we re-enter $\mathcal{S}^{(0)}$ for the i-th time).

It is obvious that the result applies to Example exa:bat_simple 1. The result also applies to the general finite bipartite graph setting (mentioned earlier) as we will now describe explicitly.

exa:Batgen

Example 2 (Generalised Batman and Joker chain)**.** *Let $G = (V, E)$ be a finite bipartite with vertex set V and edge set E. Let $(p_*^B(v, v'))_{v,v'\in V}$ be transition probabilities of an irreducible Markov chain on V such that $p_*^B(v, v') = 0$ if $\{v, v'\} \notin E$ (in particular $p_*^B(v, v) = 0$ for each v) and similarly let $(p_*^J(v, v'))_{v,v'\in V}$ be transition probabilities of an irreducible Markov chain on V with the same properties.*

Batman and the Joker independently walk on G each as an irreducible discrete-time Markov chain with state space V (with vertex- and character-specific transition probabilities $p^B(v,v')$ and $p^J(v,v')$ for $v,v' \in V$) such that $p^B(v,v) = c_B q'$ and $p^J(v,v) = c_J q'$ for every $v \in V$ and $p^B(v,v') = (1 - c_B q')p_*^B(v,v')$ and $p^J(v,v') = (1 - c_J q')p_*^J(v,v')$.

Readers interested in the somewhat different setting where Batman and the Joker both move with purpose (each knowing the location of the other, with the former attempting to capture and the latter attempting to evade capture) can read about the so-called Cops and Robbers game [1] on a graph.

cor:bipartite

Corollary 3. *For a generalised Batman and Joker chain on a finite bipartite graph $G = (V,E)$, if Batman and the Joker start at distinct vertices of the same parity, on average the Joker remains free twice as long in the limit as $q' \downarrow 0$ as he does with $q' = 0$.*

Proof. Let $\mathcal{S} = \{(x,y) : x,y \in V\}$. Write $x \sim y$ if x and y have the same parity. Let $q = (c_B + c_J)q' - c_B c_J (q')^2$, which is the probability that exactly one of the characters makes a move. Apply Theorem 1 with the limit as $q \downarrow 0$ with the choices $\mathcal{S}^{(0)} = \{(x,y) \in \mathcal{S} : x \sim y\}$, $\mathcal{S}^{(0)} = \mathcal{S} \setminus \mathcal{S}^{(0)}$, $A = \{(x,x) : x \in V\} \subset \mathcal{S}^{(0)}$, and $c_0 = c_1 = 1$. ∎

Let us now prove our main result.

Proof of Theorem 1. As indicated in the sketch proof, we define a particular coupling of the processes $\boldsymbol{X}(q)$ for all q sufficiently small. Let $(\boldsymbol{X}^{(i)}(j))_{i\in\mathbb{Z}_+, j\in\mathcal{S}}$ be an independent sequence of Markov chains such that if $j \in \mathcal{S}^{(k)}$ then $\boldsymbol{X}^{(i)}(j)$ has the law of a $P^{(k)}$-Markov chain starting from j. Let $(U_i)_{i\in\mathbb{N}}$ be iid uniform random variables on $[0,1]$. Let $\mathcal{T}^{(0)}(q) = 0$, for odd j, let $\mathcal{T}^{(j)}(q) = \inf\{i > \mathcal{T}^{(j-1)}(q) : U_i < c_0 q\}$ and for even j let $\mathcal{T}^{(j)}(q) = \inf\{i > \mathcal{T}^{(j-1)}(q) : U_i < c_1 q\}$. Clearly the $\mathcal{T}^{(i)}(q)$ are decreasing in q and, for each $i \in \mathbb{N}$, the differences $\tau_j(q) = \mathcal{T}^{(j)}(q) - \mathcal{T}^{(j-1)}(q)$ are independent random variables that are Geometric($c_0 q$) when j is odd and Geometric($c_1 q$) when j is even. Let $(Z^{(m)}(i))_{m\in\mathbb{N}, i\in\mathcal{S}}$ be independent random variables with $Z^{(m)}(i)$ having distribution $\boldsymbol{p_i}$.

Define $\boldsymbol{X}(q)$ by setting $X_0(q) = s_0$ and (for each $i \in \mathbb{Z}_+$)

- $X_t(q) = \boldsymbol{X}^{(i)}_{t-\mathcal{T}^{(i)}(q)}$ for $t \in [\mathcal{T}^{(i)}(q), \mathcal{T}^{(i+1)}(q))$
- $X_{\mathcal{T}^{(i+1)}(q)} = Z^{(i+1)}(X_{\mathcal{T}^{(i+1)}(q)-1})$.

This generates a family of sequences $\boldsymbol{X}(q)$ for $q \in [0,1]$ with the correct law for each q and such that if $q_1 < q_2$ then $X_t(q_1) = X_t(q_2)$ for $t < \mathcal{T}^{(1)}(q_2)$. Thus, if $T(q) < \mathcal{T}^{(1)}(q)$ then $T(q) = T(0)$. Furthermore, since $\boldsymbol{X}(q)$ and $\boldsymbol{X}(0)$ agree up to time $\mathcal{T}^{(1)}(q) - 1$, we see that $\{T(q) \geq \mathcal{T}^{(1)}(q)\} = \{T(0) \geq \mathcal{T}^{(1)}(q)\}$. It follows that, for every $q \in [0,1]$,

$$\begin{aligned} \mathbb{E}[T(q)] &= \mathbb{E}[T(q)\mathbb{1}_{\{T(q)<\mathcal{T}^{(1)}(q)\}}] + \mathbb{E}[T(q)\mathbb{1}_{\{T(q)\geq\mathcal{T}^{(1)}(q)\}}] && (8)\\ &= \mathbb{E}[T(0)\mathbb{1}_{\{T(0)<\mathcal{T}^{(1)}(q)\}}] + \mathbb{E}[T(q)\mathbb{1}_{\{T(q)\geq\mathcal{T}^{(1)}(q)\}}]. && (9) \end{aligned}$$

As $q \downarrow 0$, we see that $\mathcal{T}^{(1)}(q) \uparrow \infty$ and the first expectation increases to $\mathbb{E}[T(0)]$ by monotone convergence. For the second expectation we may write $T'(q) = T(q) - (\tau^{(1)}(q) + \tau^{(2)}(q))$ which (since we cannot hit $0 \in \mathcal{S}^{(0)}$ on the interval $[\tau^{(1)}(q), \tau^{(1)}(q) + \tau^{(2)}(q)))$ is non-negative on the event $\{T(0) \geq \mathcal{T}^{(1)}(q)\}$. Thus the second expectation can be written as

$$\mathbb{E}[(\tau^{(1)}(q) + \tau^{(2)}(q) + T'(q))\mathbb{1}_{\{T(0)\geq\tau^{(1)}(q)\}}]. \qquad (10)$$ eq:expression5

The random variables $T(0)$ and $\tau^{(1)}(q)$ and $\tau^{(2)}(q)$ are independent. Therefore eq:expression5 (10) can be written as

$$\mathbb{E}[\tau^{(1)}(q)\mathbb{1}_{\{T(0)\geq\tau^{(1)}(q)\}}] + \frac{1}{c_1 q}\mathbb{P}(T(0) \geq \tau^{(1)}(q)) + \mathbb{E}[T'(q)\mathbb{1}_{\{T(0)\geq\tau^{(1)}(q)\}}]. \qquad (11)$$

The first term is at most $\mathbb{E}[T(0)\mathbb{1}_{\{T(0)\geq\tau^{(1)}(q)\}}]$ which goes to 0 as $q \to 0$ by dominated convergence, since $\mathbb{E}[T(0)] < \infty$ (this follows from the fact that $\mathbb{P}^{(0)}$ is irreducible with a finite state-space). Next, by independence,

$$\begin{aligned} \mathbb{P}(T(0) \geq \tau^{(1)}(q)) &= \sum_{n=1}^{\infty} \mathbb{P}(T(0) \geq n)\mathbb{P}(\mathcal{T}^{(1)}(q) = n) && (12)\\ &= c_0 q \sum_{n=1}^{\infty} \mathbb{P}(T(0) \geq n)(1 - c_0 q)^{n-1}. && (13) \end{aligned}$$

It follows that as $q \downarrow 0$,

$$\frac{1}{c_1 q}\mathbb{P}(T(0) \geq \tau^{(1)}(q)) \to \frac{c_0}{c_1}\sum_{n=1}^{\infty} \mathbb{P}(T(0) \geq n) = \frac{c_0}{c_1}\mathbb{E}[T(0)]. \qquad (14)$$

To prove the theorem, it therefore remains to show that

$$\lim_{q\downarrow 0} \mathbb{E}[T'(q)\mathbb{1}_{\{T(0)\geq\mathcal{T}^{(1)}(q)\}}] = 0. \qquad (15)$$

Conditioning on the value of $X_{\mathcal{T}^{(2)}(q)}(q)$ and using the Markov property, this expectation is equal to

$$\sum_{k\in\mathcal{S}^{(0)}} \mathbb{E}[T'(q)\mathbb{1}_{\{X_{\mathcal{T}^{(2)}(q)}(q)=k\}}\mathbb{1}_{\{T(0)\geq\mathcal{T}^{(1)}(q)\}}] \tag{16}$$

$$= \sum_{k\in\mathcal{S}^{(0)}} \mathbb{E}_k[T(q)]\mathbb{P}_{s_0}\big(X_{\mathcal{T}^{(2)}(q)}(q)=k, T(0)\geq\mathcal{T}^{(1)}(q)\big), \tag{17}$$

where the subscript in the measure now indicates the starting point of the chain. This is at most

$$\sum_{k\in\mathcal{S}^{(0)}} \mathbb{E}_k[T(q)]\mathbb{P}_{s_0}(T(0)\geq\mathcal{T}^{(1)}(q)\big). \tag{18}$$

Since $\mathbb{P}_{s_0}(T(0)\geq T^{(1)}(q))\downarrow 0$ as $q\downarrow 0$ and $\mathcal{S}^{(0)}$ is finite, it is sufficient to prove that for some $q_0>0$, $\sup_{q\in(0,q_0]}\mathbb{E}_k[T(q)]<\infty$ for every $k\in\mathcal{S}^{(0)}\setminus A$.

Let $r_0=|\mathcal{S}^{(0)}|-1$. Then there exists $p_0>0$ depending only on $P^{(0)}$ (in particular not on $q\leq 1/2$ and k), such that $\mathbb{P}_k(T(q)<\mathcal{T}^{(1)}(q)|\mathcal{T}^{(1)}(q)> r_0)\geq p_0^{r_0}>0$. On the other hand, letting $\boldsymbol{\mathcal{T}}(q)=\{\mathcal{T}^{(i)}(q): i\in\mathbb{N}\}$, we have that $\mathbb{P}(\boldsymbol{\mathcal{T}}(q)\cap(\ell r_0,(\ell+1)r_0]=\varnothing)\geq(1-\bar{c}q)^{r_0}\geq 2^{-r_0}$ (for $q\leq 1/\bar{c}$), where $\bar{c}=\max\{c_0,c_1\}$. This shows that $\mathbb{P}_k(T(q)\leq r_0)\geq\alpha:=(p_0/2)^{r_0}$ for all $k\in\mathcal{S}^{(0)}\setminus A$ and $q\leq 1/2$. Note that $\alpha<1/2$.

In preparation for a stochastic domination argument, let $(G_j(q))_{j\in\mathbb{N}}$ be iid Geometric(c_1q) random variables, and let $((Y_j^{[1]}(q),Y_j^{[2]}(q)))_{j\in\mathbb{N}}$ be iid random vectors with

$$\mathbb{P}\Big((Y_j^{[1]}(q),Y_j^{[2]}(q))=(1,r_0)\Big)=\alpha \tag{19}$$

$$\mathbb{P}\Big((Y_j^{[1]}(q),Y_j^{[2]}(q))=(0,r_0+G_j(q))\Big)=\beta(q) \tag{20}$$

$$\mathbb{P}\Big((Y_j^{[1]}(q),Y_j^{[2]})(q)=(0,r_0)\Big)=1-\alpha-\beta(q), \tag{21}$$

where $\beta(q):=1-(1-c_0q)^{r_0}$. Then $M:=\inf\{j: Y_j^{[1]}=1\}$ has a Geometric(α)

distribution and

$$\mathbb{E}\left[\sum_{j=1}^{M} Y_j^{[2]}(q)\right] = r_0 + \mathbb{E}[M-1]\mathbb{E}[Y_1^{[2]}(q)|1 < M] \tag{22}$$

$$\leq r_0 + \frac{1}{1-\alpha}\left[(r_0 + \frac{1}{c_1 q})\beta(q) + r_0(1-\alpha-\beta(q))\right] \tag{23}$$

$$= 2r_0 + \frac{1}{1-\alpha}\frac{\beta(q)}{c_1 q}, \tag{24}$$

which is bounded in $q \leq q_0$, say less than γ_0, since $\beta(q) \leq c_{r_0} q$ for some $c_{r_0} < \infty$. We will show that, starting from any $k \in \mathcal{S}^{(0)} \setminus A$, for all $q \leq q_0$, $T(q)$ is stochastically dominated by $\sum_{j=1}^{M} Y_j^{[2]}$.

Let $S_0(q) = 0$. Given $S_i(q)$, we define $S_{i+1}(q)$ by

$$S_{i+1}(q) = \begin{cases} S_i(q) + r_0, & \text{if } \boldsymbol{\mathcal{T}}(q) \cap (S_i(q), S_i(q) + r_0] = \varnothing \\ \mathcal{T}^{(j)}(q), & \text{if } \inf\{\ell : \mathcal{T}^{(\ell)}(q) \in (S_i(q), S_i(q) + r_0]\} = j - 1. \end{cases} \tag{25}$$

By construction, j above will always be even. Let $\boldsymbol{T}(q)$ denote the set of hitting times of A by the chain $\boldsymbol{X}(q)$ and let $N(q)$ be the index of the first interval $(S_{i-1}(q), S_i(q)]$ containing no element of $\boldsymbol{\mathcal{T}}(q)$ but some element of $\boldsymbol{T}(q)$. Then $T(q) \leq r_0 + S_{N-1}(q)$, since the time in the interval $(S_{N-1}(q), S_N(q)]$ that $\boldsymbol{X}(q)$ hits A is at least $T(q)$, and the length of this last interval is r_0. Let $D_i = \{(S_{i-1}(q), S_i(q)] \cap \boldsymbol{T}(q) = \varnothing\}$ and $R_i = \{\boldsymbol{\mathcal{T}}(q) \cap (S_i, S_i + r] = \varnothing\}$. Now, given the history of the process up to time $S_{i-1}(q)$, the probability of the event $A_i = D_i^c \cap R_i$ is at least α (thus N is dominated by a Geometric(α) random variable), and the probability of R_i^c is $\beta(q)$. On the event R_i^c, the length of the interval $(S_{i-1}(q), S_i(q)]$ is at most the sum of two independent Geometric(q) random variables (that are independent of the history) with the first one conditioned to be at most r_0, and the second having parameter c_1. Hence on the event R_i^c, the length of the interval $(S_{i-1}(q), S_i(q)]$ is at most r_0 plus a Geometric($c_1 q$) random variable (independent of the past). On R_i, the length of the interval is r_0.

Thus, $r_0 + S_{N-1}(q)$ is stochastically dominated by $\sum_{j=1}^{M} Y_j^{[2]}$, and the result follows. ■

Acknowledgements. The work of MH was supported by Future Fellowship FT160100166, from the Australian Research Council (ARC). Peter Taylor's research is supported by ARC Laureate Fellowship FL130100039 and the ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS). MH thanks Victor Kleptsyn for suggesting a particular generalisation of an earlier formulation of the theorem.

References

[1] A. Bonato and R.J. Nowakowski *The Game of Cops and Robbers on Graphs.* AMS Student Mathematical Library, Vol. 61, (2011).

[2] D.J. Chalmers The St. Petersburg two-envelope paradox. *Analysis*, Volume 62, Issue 2, 1 April 2002, Pages 155?157.

[3] G. Grimmett and D. Welsh. *Probability: An Introduction.* Second Edition, Oxford University Press, (2014).

[4] W. Feller. *An Introduction to Probability Theory and its Applications, Volume 1* - 3rd edition Wiley, (1968).

[5] W. Feller. *An Introduction to Probability Theory and its Applications, Volume 2* - 2nd edition Wiley, (1971).

[6] S.M. Ross *A first course in probability* - 5th edition. Prentice-Hall, (1998).