MOMENTS OF RECURRENCE TIMES FOR MARKOV CHAINS

FRANK AURZADA
Technische Universität Berlin
e-mail: aurzada@math.tu-berlin.de

HANNA DÖRING
Technische Universität Berlin
e-mail: hdoering@math.tu-berlin.de

MARCEL ORTGIESE
Technische Universität Berlin
e-mail: ortgiese@math.tu-berlin.de

MICHAEL SCHEUTZOW
Technische Universität Berlin
e-mail: ms@math.tu-berlin.de

Submitted April 11, 2011, accepted in final form May 10, 2011

AMS 2000 Subject classification: 60J10
Keywords: Discrete time Markov chain, recurrence time, generalized moment.

Abstract
We consider moments of the return times (or first hitting times) in an irreducible discrete time
discrete space Markov chain. It is classical that the finiteness of the first moment of a return time
of one state implies the finiteness of the first moment of the first return time of any other state. We
extend this statement to moments with respect to a function $f$, where $f$ satisfies a certain, best
possible condition. This generalizes results of K. L. Chung (1954) who considered the functions
$f(n) = n^p$ and wondered "[...] what property of the power $n^p$ lies behind this theorem [...]" (see
Chung (1967), p. 70). We exhibit that exactly the functions that do not increase exponentially –
neither globally nor locally – fulfill the above statement.

1 Introduction

A classical result, see e.g. [Kol36], states that for any recurrent, irreducible Markov chain on a
countable state space the following holds: if for any state $i$ the first moment of the recurrence
time is finite then this also applies to any other state. A first generalization of this result appeared
in [HR53]. If we denote by $T_{ij}$ the first time that the Markov chain visits state $j$ if it is started in
$i$, then the result can be stated as follows: $Ef(T_{ii}) < \infty$ for some state $i$ implies that $Ef(T_{jj}) < \infty$
for any other state $j$, where $f(x) = x^n$ for some integer $n$. The authors state the result as a lemma
and refer to the proof to an (unpublished) note by K. L. Chung and R. N. Snow. A more general
result can be found in [Chu54, Theorem 1], where $f$ is allowed to be of the form $x^p$, for any
real \( p > 0 \). After stating the theorem, the author also comments that the concept of generalized moments defined in terms of a general function \( f \) was suggested to him by J. L. Doob, but Chung only mentions that his results can also be shown for functions \( f \) satisfying \( f(x + y) \leq f(x) + Af(y) \), for some constant \( A \).

Further related research considers recursive formulas for second moments in terms of first moments [Chu54, Sect. 2], factorial moments [Lam60], and also, more recently, explicit formulas for higher polynomial moments [Sze08].

These considerations naturally lead to the question, for which functions it is true that a finite generalized moment of the return time for one state \( i \) implies that the moment is also finite for the return time for any other state of the Markov chain. In this note, we characterize this class of functions. In the following, we only consider irreducible, recurrent discrete time Markov chains with a countable state space.

To formulate our results, we introduce the following notation. The candidate functions \( f \) are taken from the set \( \mathcal{F}_{\text{cal}} = \{ f : \mathbb{N} \to (0, \infty) \text{ is non-decreasing and } \lim_{n \to \infty} f(n) = \infty \} \).

Then our objective is to classify the collection \( \mathcal{G}_{\text{cal}} \) of all \( f \in \mathcal{F}_{\text{cal}} \) such that for each irreducible recurrent discrete time Markov chain with a finite or countably infinite state space \( E \), the following holds: if \( E f(T_{ii}) < \infty \) for some \( i \in E \) then \( E f(T_{jj}) < \infty \) for all \( j \in E \).

Following Chung [Chu54] we additionally introduce the class \( \mathcal{H}_{\text{cal}} \), by stating that \( f \in \mathcal{H}_{\text{cal}} \), if for any Markov chain the following holds: if there exist two states \( i \) and \( j \) such that \( E f(T_{ij}) < \infty \) and \( E f(T_{ji}) < \infty \) then \( E f(T_{kl}) < \infty \) for any pair \( k, l \).

As the states \( i \) and \( j \) do not have to be distinct, it follows that \( \mathcal{H}_{\text{cal}} \) is contained in \( \mathcal{G}_{\text{cal}} \). The classical result due to Kolmogorov implies that the identity function belongs to \( \mathcal{H}_{\text{cal}} \) and [Chu54] shows that any \( f(x) = x^p \), for \( p > 0 \), belongs to \( \mathcal{H}_{\text{cal}} \).

Our main result states that the two classes \( \mathcal{G}_{\text{cal}} \) and \( \mathcal{H}_{\text{cal}} \) are in fact the same; and we also give a characterization for a function to be in this class.

**Theorem 1.1.** Let \( f \in \mathcal{F}_{\text{cal}} \). Then the following statements are equivalent:

(a) \( f \in \mathcal{H}_{\text{cal}} \);

(b) \( f \in \mathcal{G}_{\text{cal}} \);

(c) the following two conditions are satisfied:

(i) there exists \( K > 0 \) such that for any \( x, y > 0 \), \( f(x + y) \leq K f(x)f(y) \).

(ii) \( \limsup_{n \to \infty} \frac{1}{n} \log f(n) = 0 \).

Condition (c) has an easy interpretation: It ensures that the function \( f \) does not grow exponentially fast – neither globally nor locally. In fact, one can construct functions outside \( \mathcal{G}_{\text{cal}} \) that globally increase as slowly as one wishes, but locally have parts of exponential increase (cf. Example 3.3).

This note is structured as follows. In Section 2, we prove the implication \( (c) \Rightarrow (a) \) of our main theorem. In Section 3, we prove the implication \( (b) \Rightarrow (c) \) by showing that any function that violates either condition (i) or (ii) in (c) is not in the class \( \mathcal{G}_{\text{cal}} \). Together with the earlier observation that \( \mathcal{H}_{\text{cal}} \subseteq \mathcal{G}_{\text{cal}} \), so that \( (a) \Rightarrow (b) \), this shows the equivalence of the three statements.
2 Proof of \((c) \Rightarrow (a)\)

In this section, we prove that \((c) \Rightarrow (a)\) in our main theorem. We start with the following lemma, which collects some preliminary facts.

For this purpose, it is convenient to introduce the following notation: for states \(i\) and \(j\) of a Markov chain, we denote by \(U_{ij}\) the return time from \(i\) to \(i\) conditioned on not crossing \(j\) (if there is such a path with positive probability). Further, \(V_{ij}\) denotes the first hitting time of state \(j\) when started from \(i\) conditioned on not returning to \(i\) before hitting \(j\).

![Figure 1: Visualization of the relevant return/hitting times](image)

**Lemma 2.1.** Let \(f \in \mathcal{F}\).

(i) If for a state \(i\) of a Markov chain we have \(\mathbb{E}f(T_{ii}) < \infty\) then, for any state \(j\), \(\mathbb{E}f(U_{ij}) < \infty\) and \(\mathbb{E}f(V_{ij}) < \infty\).

(ii) If for a state \(i\) of a Markov chain we have \(\mathbb{E}f(T_{ii}) < \infty\) then, for any state \(j\), \(\mathbb{E}f(T_{ji}) < \infty\).

**Proof.** To see the first part of (i) note that if the probability \(p\) of going from \(i\) to \(i\) without crossing \(j\) is positive, we have

\[
\infty > \mathbb{E}f(T_{ii}) \geq \mathbb{E}[f(T_{ii}) \mathbf{1}_{\text{do not cross } j}] = \mathbb{E}[f(T_{ii}) \mid \text{do not cross } j] p = p \mathbb{E}f(U_{ij}).
\]

This shows \(\mathbb{E}f(U_{ij}) < \infty\).

The remaining parts are shown similarly.

We stress that, in the second part of Lemma 2.1, one cannot prove in the same way that \(\mathbb{E}f(T_{ij}) < \infty\), since a typical path from \(i\) to \(i\) does not necessarily contain a path from \(i\) to \(j\). However, as the next lemma shows this can be shown if we assume that condition \((c)\) holds. This lemma is also the main part of the argument for the proof of \((c) \Rightarrow (a)\) in our main theorem.

**Lemma 2.2.** Let \(f \in \mathcal{F}\) and assume that \((c)\) holds.

(i) If for two states \(i\) and \(j\) of a Markov chain we have \(\mathbb{E}f(T_{ij}) < \infty\) and \(\mathbb{E}f(T_{ji}) < \infty\) then \(\mathbb{E}f(T_{ii}) < \infty\).

(ii) If for a state \(i\) of a Markov chain we have \(\mathbb{E}f(T_{ii}) < \infty\) then, for any state \(j\), \(\mathbb{E}f(T_{ij}) < \infty\).
Proof. First we show (i). Clearly $T_{ij}$ is stochastically dominated by $T_{ij} + T_{ji}$, where $T_{ij}$ and $T_{ji}$ are independent. Using (c) and the monotonicity of $f$, we get

$$
Ef(T_{ij}) \leq Ef(T_{ij} + T_{ji}) \leq K Ef(T_{ij})Ef(T_{ji}) < \infty.
$$

Now we turn our attention to (ii). For the purpose of this proof, define $f(0) := 1/K$, where $K$ is as in (i) of (c). Note that $f(1) \leq f(2) \leq Kf(1)^2$, so that $f(0) = K^{-1} \leq f(1)$. This shows that in fact $f(x + y) \leq Kf(x)f(y)$ for all $x, y \geq 0$, since $f(x + 0) = f(x) = Kf(0)f(x)$.

The crucial observation (also cf. (3.3) below) is that

$$
T_{ij} = \sum_{r=1}^{M} U^{(r)} + V,
$$

(2.1)

where the random variables $U^{(r)}$ are i.i.d. copies of the random variable $U_{ij}$ as defined before Lemma 2.1 and $V$ is a copy of the random variable $V_{ij}$ as defined before Lemma 2.1, and all variables are independent. Further, $M$ (independent of the $U$’s and $V$) is a geometric random variable with mean $1/\pi - 1$, where $\pi > 0$ is the probability of first hitting $j$ before $i$ when started from $i$.

It may be that $\pi = 1$ – which is the case if and only if there is no path from $i$ to $i$ without crossing $j$ – in which case $T_{ij} = V$ and we are already done. Excluding this case, we derive from (c) and (2.1) that

$$
Ef(T_{ij}) \leq KEf \left( \sum_{r=1}^{M} U^{(r)} \right) Ef(V) = K \sum_{m=0}^{\infty} (1 - \pi)^m \pi Ef \left( \sum_{r=1}^{m} U^{(r)} \right) Ef(V).
$$

Since by Lemma 2.1(i) $Ef(V) < \infty$, it is clear that the last expression is finite provided that

$$
\limsup_{m \to \infty} \frac{1}{m} \log Ef \left( \sum_{r=1}^{m} U^{(r)} \right) = 0,
$$

(2.2)

where we know that $Ef(U^{(1)}) < \infty$ from Lemma 2.1(i).

To show (2.2) fix a large constant $A > 0$ and estimate using (c) as follows:

$$
Ef(U^{(1)} + \ldots + U^{(m)}) \leq Ef \left( \sum_{r=1}^{m} (U^{(r)} \vee A) \right)
$$

$$
\leq Ef \left( \sum_{r=1}^{m} (U^{(r)} 1_{\{U^{(r)} > A\}} + A) \right)
$$

$$
\leq Kf(Am)Ef \left( \sum_{r=1}^{m} U^{(r)} 1_{\{U^{(r)} > A\}} \right)
$$

$$
\leq Kf(Am)Ef \left( K^{m-1} \prod_{r=1}^{m} f \left( U^{(r)} 1_{\{U^{(r)} > A\}} \right) \right)
$$

$$
= f(Am)Ef \left( Kf \left( U^{(1)} 1_{\{U^{(1)} > A\}} \right) \right)^m.
$$

Hence,

$$
\frac{1}{m} \log Ef(U^{(1)} + \ldots + U^{(m)}) \leq \frac{1}{m} \log f(Am) + \log(KEf(U^{(1)} 1_{\{U^{(1)} > A\}})).
$$
In this section, we will show the implication \( f \) we give abstract conditions on component-wise growth conditions satisfied by \( f \) and \( T \). By Lemma 2.2(i), we have
\[
\log \mathbb{E} f(U^{(1)}) \leq \log \mathbb{E} f(U^{(1)} 1_{U^{(1)} > A}).
\]
Therefore, using once again part (i) of (c), we have
\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{E} f(U^{(1)} + \ldots + U^{(m)}) \leq \log(K \mathbb{E} f(U^{(1)} 1_{U^{(1)} > A})).
\]
Letting now \( A \to \infty \) and using dominated convergence since \( \mathbb{E} f(U^{(1)}) < \infty \), we have
\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{E} f(U^{(1)} + \ldots + U^{(m)}) \leq \log(K f(0)) = 0,
\]
which proves (2.2). \( \square \)

Now we can prove that (c) \( \Rightarrow \) (a) in our main Theorem 1.1.

**Proof of (c) \( \Rightarrow \) (a) in Theorem 1.1.** Consider a Markov chain with state space \( E \), let \( f \) satisfy condition (c) of Theorem 1.1, and assume that \( i, j \in E \) satisfy \( \mathbb{E} f(T_{ij}) < \infty \) and \( \mathbb{E} f(T_{ij}) < \infty \). Let \( k, \ell \in E \). We want to show that \( \mathbb{E} f(T_{k\ell}) < \infty \).

By Lemma 2.2(i), we have \( \mathbb{E} f(T_{ii}) < \infty \) and \( \mathbb{E} f(T_{ij}) < \infty \). Thus, by Lemma 2.2 part (ii) \( \mathbb{E} f(T_{ij}) < \infty \). And by Lemma 2.1(ii) \( \mathbb{E} f(T_{ik}) < \infty \). Therefore, using once again part (i) of (c), we have
\[
\mathbb{E} f(T_{k\ell}) \leq \mathbb{E} f(T_{ki} + T_{ij} + T_{j\ell}) \leq K^2 \mathbb{E} f(T_{ki}) \mathbb{E} f(T_{ij}) \mathbb{E} f(T_{j\ell}) < \infty.
\]
\( \square \)

# 3 Proof of (b) \( \Rightarrow \) (c)

In this section, we will show the implication (b) \( \Rightarrow \) (c) in Theorem 1.1 by showing that the conditions of subexponential growth rate and submultiplicativity are in fact necessary. In Lemma 3.1, we give abstract conditions on \( f \), which imply that \( f \notin \mathcal{G} \), which we exploit in Lemma 3.2 to show that if \( f \) violates the submultiplicativity condition (i) of (c) in the main theorem, we have that \( f \notin \mathcal{G} \). Finally, in Lemma 3.4, we show that any function growing exponentially fast, i.e. does not satisfy condition (ii), does not belong to \( \mathcal{G} \).

**Lemma 3.1.** If for a given \( f \in \mathcal{G} \) one can construct two independent random variables \( U_1 \) and \( U_2 \) taking values in \( \mathbb{N} \) with infinite support such that \( \mathbb{E} f(U_i) < \infty \) for \( i = 1, 2 \), but \( \mathbb{E} f(U_1 + U_2) = \infty \), then \( f \notin \mathcal{G} \).

**Proof.** Given the two random variables, we will construct a Markov chain with two special states 0 and 1 with the property that \( \mathbb{E} f(T_{11}) < \infty \), whereas \( \mathbb{E} f(T_{00}) \) is infinite, which shows that \( f \notin \mathcal{G} \). The construction of the Markov chain is in some regards similar to [YK39], where the authors construct a Markov chain with \( T_{00} \) having any particular distribution.

Denote by \( \{x_1, x_2, \ldots\} \), \( \{y_1, y_2, \ldots\} \) the support of \( U_1 \) and \( U_2 \), respectively. Formally, we can write the state space \( E \) of our Markov chain as
\[
\{0\} \cup \{1\} \cup \{(L, n, m) : n \in \mathbb{N}, m = 1, \ldots, x_n - 1\} \cup \{(R, n, m) : n \in \mathbb{N}, m = 1, \ldots, y_n - 1\}.
\]

The state 0 is connected only to 1, and if in state 0, the chain always moves to 1 next, i.e. \( p_{01} = 1 \). If in state 1, the chain has three possibilities. The first one is that the chain moves to 0 with probability \( p \), for some parameter \( p \in (0, 1) \). Then, conditionally on not going to 0, with equal probability it either moves “left” (i.e. to a state \( (L, n, 1) \)) or “right” (i.e. to a state \( (R, n, 1) \)).
Moments of recurrence times

\[ \cdots \]

\[ \text{(L, n, 2)} \quad \text{(L, n, 1)} \quad \text{(R, n, 1)} \quad \text{(R, n, 2)} \]

\[ \text{\cdots} \]

\[ \text{(L, n, x_n - 1)} \quad \text{(R, n, y_n - 1)} \]

\[ \text{\cdots} \]

\[ 0 \]

Figure 2: Transition graph of the Markov chain

Conditionally on the next move going to the “left”, we want \( T_{11} \) to have distribution \( U_1 \), therefore we set

\[
P_{1,(L,n,1)} = \frac{1-p}{2} P(U_1 = x_n),
\]

\[
P_{(L,n,m-1),(L,n,m)} = 1, \quad n \in \mathbb{N}, 1 < m \leq x_n,
\]

where we identify \((L, n, x_n)\) with \( 1 \). Similarly, conditionally on the next move going “right”, we would like \( T_{11} \) to have the distribution \( U_2 \), and set

\[
P_{1,(R,n,1)} = \frac{1-p}{2} P(U_2 = y_n),
\]

\[
P_{(R,n,m-1),(R,n,m)} = 1, \quad n \in \mathbb{N}, 1 < m \leq y_n,
\]

where we again identify \((R, n, y_n)\) with \( 1 \).

Now, we can calculate the generalized moments of \( T_{00} \) and \( T_{11} \). Firstly, we find for \( T_{11} \) by conditioning on the three different possibilities

\[
Ef(T_{11}) = f(2)P\{1 \rightarrow 0\} + Ef(U_1)P\{1 \rightarrow \text{“left”}\} + Ef(U_2)P\{1 \rightarrow \text{“right”}\},
\]

which is finite by our assumptions on \( U_1 \) and \( U_2 \).

However, for \( T_{00} \), we obtain

\[
Ef(T_{00}) = Ef\left( \sum_{i=1}^{M} U_{10}^{(i)} + 2 \right),
\]

(3.3)

where \( M \) is a geometric random variable with parameter \( p \) and \( U_{10}^{(i)} \) are independent random variables that have the same distribution as \( T_{11} \) conditioned on not going to 0 in the first step. (Recall (2.1).) In particular we obtain a lower bound by considering the following strategy: first the Markov chain jumps from 0 to 1, then it takes a tour to the “left” and after that it takes a tour of “right”, before it finally returns to 0. Thus, we obtain using that \( f \) is non-decreasing

\[
Ef(T_{00}) \geq Ef(2 + U_{10}^{(1)} + U_{10}^{(2)}) \mid \text{first left, then right } \] P\{ first left, then right \}
\]

\[
\geq Ef(U_1 + U_2) p \left( \frac{1-p}{2} \right)^2,
\]
where the latter is infinite by our assumptions on \( U_1 \) and \( U_2 \); and thus as claimed \( \mathbb{E} f(T_{00}) \) is infinite.

The next lemma uses the construction in Lemma 3.1 to show that any function not satisfying the submultiplicativity condition (i) in (c) is not in \( \mathcal{G} \).

**Lemma 3.2.** Suppose that \( f \in \mathcal{F} \) is such that for any \( C > 0 \), there exist \( x_C \) and \( y_C \) such that

\[
  f(x_C + y_C) > Cf(x_C)f(y_C),
\]

then \( f \notin \mathcal{G} \).

**Proof.** By Lemma 3.1 it suffices to construct two random variables \( U_1 \) and \( U_2 \) such that \( \mathbb{E} f(U_i) < \infty \) for \( i = 1, 2 \) and \( \mathbb{E} f(U_1 + U_2) = \infty \).

By our assumption on \( f \), we can find increasing sequences \( (x_k)_{k \geq 1} \) and \( (y_k)_{k \geq 1} \) such that

\[
  f(x_k + y_k) > k^6 f(x_k)f(y_k).
\]

Indeed, to see the existence of such sequences assume that for all \( x \geq x_1 + 1 \) and \( y \geq y_1 + 1 \) the following inequality holds

\[
  f(x + y) \leq (k + 1)^6 f(x)f(y).
\]

A short calculation implies that

\[
  f(x + y) \leq (k + 1)^6 \left( \frac{f(x_k \vee y_k + 1)}{f(0)} \right)^2 f(x)f(y) \quad \text{for all } x, y \geq 0,
\]

in contradiction to the assumption of the lemma.

Then, for definiteness, let \( U_1 \) be a random variable taking value \( x_k \) with probability \( p_k := c_1 f(x_k)^{-1} k^{-2} \) (with a suitable normalizing constant \( c_1 \)) and similarly, \( U_2 \) takes value \( y_k \) with probability \( q_k := c_2 f(y_k)^{-1} k^{-2} \) (with a suitable normalizing constant \( c_2 \)). In particular, we find that \( \mathbb{E} f(U_i) \) is finite for \( i = 1, 2 \). However,

\[
  \mathbb{E} f(U_1 + U_2) \geq \sum_{k \geq 1} f(x_k + y_k)p_kq_k \geq c_1 c_2 \sum_{k \geq 1} k^6 f(x_k)f(y_k)f(x_k)^{-1} k^{-2} f(y_k)^{-1} k^{-2},
\]

which is infinite. \( \square \)

The following example exhibits a “typical” function that can be chosen to satisfy (ii) of (c), while not obeying (i) of (c).

**Example 3.3.** We now construct a function \( f \) where we can choose the parameters in such a way that the condition in Lemma 3.2 is satisfied and we have that \( \limsup_n \frac{1}{n} \log f(n) = 0 \). We first describe the function \( g = \log f \). Take two sequences \( (s_i)_{i \geq 1} \) and \( (u_i)_{i \geq 1} \) such that \( u_i \leq s_{i+1} - s_i \) and \( u_n \to \infty \) as \( n \to \infty \). Then, by setting \( g(0) = 0, s_0 = u_0 = 0 \) and for \( i \geq 0 \), define \( g \) to be constant on the interval \( (s_i + u_i, s_{i+1}) \) and for \( i \geq 1 \), assume that on the interval \( (s_i, s_i + u_i) \) the function \( g \) grows linearly with slope 1. Then, by adjusting the parameters \( u_i, s_i \) (in such a way that \( \lim_{n \to \infty} g(n)/n = 0 \)) one can make sure that the condition in Lemma 3.2 is fulfilled for \( f = e^g \), while at the same time by making the differences \( s_i - s_{i-1} \) large enough, one can let \( f \) grow as slowly as desired.

The following lemma shows that the condition on the subexponential growth rate of \( f \) – (ii) of (c) – is really necessary.
Lemma 3.4. Suppose that $f \in \mathcal{F}$ satisfies

$$
\limsup_{n \to \infty} \frac{1}{n} \log f(n) > 0,
$$

then $f \notin \mathcal{G}$.

Proof. By our assumption on $f$ we can find an increasing sequence $x_i$ and $\delta > 0$, such that $f(x_i) \geq e^{\delta x_i}$. Then, consider the Markov chain with two states 0 and 1 and transition probabilities $p_{01} = p$ and $p_{10} = 1$ for some parameter $p \in (0, 1)$. Now, $Ef(T_{00})$ is finite, while for any $i \geq 1$,

$$
Ef(T_{11}) \geq e^{\delta x_i}P\{T_{11} \geq x_i\} = e^{\delta x_i}(1-p)^{x_i-2},
$$

which tends to infinity as $i \to \infty$ provided that $p$ is sufficiently small. □

Acknowledgements: Frank Aurzada was supported by the DFG Emmy Noether programme. Hanna Döring and Marcel Ortgiese were supported by the international research training group 1339 of the DFG.

References


