2017

Standard deviation of recurrence times for piecewise linear transformations

Mimoon Ismael
University of Anbar, mii998@uowmail.edu.au

Rodney Nillsen
University of Wollongong, nillsen@uow.edu.au

Graham H. Williams
University of Wollongong, ghw@uow.edu.au

Standard deviation of recurrence times for piecewise linear transformations

Abstract

This paper is concerned with dynamical systems of the form (X,f) where X is a bounded interval and f comes from a class of measure-preserving, piecewise linear transformations on X. If A⊆X is a Borel set and x∈A, the Poincaré recurrence time of x relative to A is defined to be the minimum of $\{n : n \in \mathbb{N} \text{ and } f^n(x) \in A\}$, if the minimum exists, and ∞ otherwise. The mean of the recurrence time is finite and is given by Kac's recurrence formula. In general, the standard deviation of the recurrence times need not be finite but, for the systems considered here, a bound for the standard deviation is derived.

STANDARD DEVIATION OF RECURRENCE TIMES FOR PIECEWISE LINEAR TRANSFORMATIONS

MIMOON ISMAEL
PURE MATHEMATICS DEPARTMENT
COLLEGE OF EDUCATION AND THE PURE SCIENCES
UNIVERSITY OF ANBAR, AL-ANBAR
REPUBLIC OF IRAQ

AND

RODNEY NILLSEN AND GRAHAM WILLIAMS
SCHOOL OF MATHEMATICS AND APPLIED STATISTICS
UNIVERSITY OF WOLLONGONG
NORTHFIELDS AVENUE, WOLLONGONG
NEW SOUTH WALES, 2522
AUSTRALIA

ABSTRACT. Abstract. This paper is concerned with dynamical systems of the form (X, f), where X is a bounded interval and f comes from a class of measure-preserving, piecewise linear transformations on X. If $A \subseteq X$ is a Borel set and $x \in A$, the Poincaré recurrence time of x relative to A is defined to be the minimum of $\{n : n \in \mathbb{N} \text{ and } f^n(x) \in A\}$, if the minimum exists, and ∞ otherwise. The mean of the recurrence time is finite and is given by Kac's recurrence formula. In general, the standard deviation of the recurrence times need not be finite but, for the systems considered here, a bound for the standard deviation is derived.

1. INTRODUCTION

A *dynamical system* is a set S together with a transformation $f : S \longrightarrow S$. Letting $\circ$ denote the composition of functions, we use the notation that $f^1 = f$, $f^2 = f \circ f$, $f^3 = f \circ f \circ f$, and so on. We denote the set $\{1, 2, \ldots\}$ by $\mathbb{N}$. Now if we have a dynamical system (S, f), let $A \subseteq S$, let $x \in A$ and assume that $f^n(x) \in A$ for some $n \in \mathbb{N}$. Then, we put

$$\Theta_A(x) = \min\{n : n \in \mathbb{N} \text{ and } f^n(x) \in A\}.$$

If $x \in A$ it may happen that $f^n(x) \notin A$ for all $n \in \mathbb{N}$, in which case we put $\Theta_A(x) = \infty$. Thus, $\Theta_A : A \longrightarrow \mathbb{N} \cup \{\infty\}$ and we call $\Theta_A(x)$ the *recurrence time of x relative to A and (S, f)*, or simply the *recurrence time* if (S, f) and A are understood. See Brown [3] and Petersen [11], for example, for a discussion of this notion.

We consider the case when S is a bounded interval X of length 1. Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of X, and let μ denote the usual Lebesgue measure on $\mathcal{B}$. A formulation of the Recurrence Theorem of Henri Poincaré says that if μ preserves the measure of the Borel sets under the action of a transformation f on X, then $\Theta_A(x) < \infty$ for μ-almost all points of X. In ergodic systems, Kac [7] showed that the average value of the recurrence time is $1/\mu(A)$. The question arises as to the variation, predictability and uncertainty of the recurrence times. A formula for the moments of recurrence times in the context of stationary stochastic processes was derived by Blum and Rosenblatt [2], and a formula for the standard deviation of the recurrence times was derived by Kasteleyn [8]. A discussion of some of these matters is also in Ismael [5] and Chapter 4 of [10].

2. Estimating the standard deviation for the baker's transformations

We now introduce some further notations. Let X be an interval in $\mathbb{R}$ of length 1. Let $\mathcal{B}$ denote the σ-algebra of Borel subsets of X, and let μ denote the usual Lebesgue measure on $\mathcal{B}$. Let $I = \mathbb{N}$ or let I denote an interval in $\mathbb{N}$ of the form $\{1, 2, \ldots, r\}$ with $r \geq 2$. For each $j \in I$ let X_j be a subinterval of $[0, 1]$ of positive length such that the sets in the family $(X_j)_{j \in I}$ form a partition of X. That is, $X_i \cap X_j = \emptyset$ when $i \neq j$ and $\cup_{j \in I} X_j = X$. Such a partition we call a *standard partition* and may be denoted by $\mathcal{P}$. Note that a standard partition has at least two elements so, as $\sum_{j \in I} \mu(X_j) = 1$, for a standard partition the set $\{\mu(X_j) : j \in I\}$ has a maximum value and

$$(2.1) \qquad \max\{\mu(X_j) : j \in I\} < 1.$$

We call a transformation $f : X \longrightarrow X$ *piecewise linear* if there is a standard partition $\mathcal{P} = (X_j)_{j \in I}$ such that f is linear on each interval $X_j \in \mathcal{P}$, by which is meant that there are $a_j, b_j \in \mathbb{R}$ such that $f(x) = a_j x + b_j$ for all $x \in X_j$. If, in addition, the range of f on each interval X_j is either X or X less one or two endpoints, then f is called a (generalised) *baker's transformation* relative to $\mathcal{P}$. A particular type of baker's transformation is discussed in [12, page 77]. Note that a baker's transformation is a piecewise linear function which is not one-to-one, but is one-to-one when restricted to each interval X_j. A Baker's transformation is Lebesgue measure preserving. Note that a piecewise linear function will be either increasing or decreasing on each interval X_j in the associated partition.

Lemma 2.1. *let f be a baker's transformation on an interval X of length 1. Then, f is Lebesgue measure preserving on $\mathcal{B}$. That is, $\mu(f^{-1}(A)) = \mu(A)$, for all $A \in \mathcal{B}$.*

Proof. As f is linear on each interval in the standard partition associated with f, it is easy to see that $\mu(f^{-1}(A)) = \mu(A)$ for all subintervals A of X. Let $\mathcal{A} = \{A : A \in \mathcal{B} \text{ and } \mu(f^{-1}(A)) = \mu(A)\}$. Then $\mathcal{A}$ is a monotone class in the sense of [1, page 6] that contains all subintervals of X. By the Monotone Class Lemma, see [1, pages 6-7] for example, $\mathcal{A}$ contains the Borel sets. □

Lemma 2.2. *Let X be an interval of length 1 and let f be a baker's transformation relative to a standard partition $(X_j)_{j \in I}$ of X, as described above. Then, if $n_1, n_2, \ldots, n_k \in I$, the set $X_{n_1 n_2 \ldots n_k}$ is defined by*

$$X_{n_1 n_2 \ldots n_k} = \Big\{ x : x \in X_{n_1}, f(x) \in X_{n_2}, f^2(x) \in X_{n_3}, \ldots, f^{k-1}(x) \in X_{n_k} \Big\}.$$

Then, $X_{n_1 n_2 \ldots n_k}$ is a subinterval of X having positive length, and the restriction of f^k to $X_{n_1 n_2 \ldots n_k}$ is a linear function whose range is either X, or X less one or two endpoints.

Proof. When $k = 1$, the result is true from the definitions. So, we proceed by induction. Let $\text{int}(J)$ denote the interior of any interval J. Let $k \in \mathbb{N}$ and assume that the result holds for any choice of $n_1, n_2, \ldots, n_k \in I$. Now, let $n_1, \ldots, n_k, n_{k+1} \in I$ and observe that

$$X_{n_1 n_2 \ldots n_k n_{k+1}} = \Big\{ x : x \in X_{n_1 n_2 \ldots n_k} \text{ and } f^k(x) \in X_{n_{k+1}} \Big\}.$$

It follows that $X_{n_1 n_2 \ldots n_k n_{k+1}}$ is an interval of positive length because it is assumed that f^k is linear on the interval $X_{n_1 n_2 \ldots n_k}$ of positive length and that the range of f^k on $X_{n_1 n_2 \ldots n_k}$ is X or X less one or two endpoints. It is immediate from the definition that $f : X_{n_1 n_2 \ldots n_k n_{k+1}} \longrightarrow X_{n_2 \ldots n_k n_{k+1}}$, and f^k is linear on $X_{n_2 \ldots n_k n_{k+1}}$ by the inductive assumption. Also, as f is linear on X_{n_1}, f is linear on the smaller set $X_{n_1 n_2 \ldots n_k n_{k+1}}$. So we have

$$f : X_{n_1 n_2 \ldots n_k n_{k+1}} \longrightarrow X_{n_2 \ldots n_k n_{k+1}} \quad \text{and} \quad f^k : X_{n_2 \ldots n_k n_{k+1}} \longrightarrow X,$$

where both f and f^k are linear. Since the composition of linear functions is linear, it follows that $f^{k+1} = f^k \circ f$ is linear on $X_{n_1 n_2 \ldots n_k n_{k+1}}$.

Now, as earlier, let $\text{int}(J)$ denote the interior of any sub-interval J of X, and let $z \in \text{int}(X)$. By the inductive assumption, there is y belonging to $\text{int}(X_{n_2 n_3 \ldots n_{k+1}})$ such that $f^k(y) = z$. Then, as f maps X_{n_1} onto either X or X less one or two endpoints, there is $x \in \text{int}(X_{n_1})$ such that $f(x) = y$ and it follows that $x \in X_{n_1 n_2 \cdots n_{k+1}}$. Also, $f^{k+1}(x) = f^k\big(f(x)\big) = f^k(y) = z$. We deduce that f^{k+1} maps $\text{int}\big(X_{n_1 n_2 \cdots n_{k+1}}\big)$ onto $\text{int}(X)$. As f^{k+1} is linear on $\text{int}\big(X_{n_1 n_2 \cdots n_{k+1}}\big)$, f^{k+1} must map $\text{int}\big(X_{n_1 n_2 \cdots n_{k+1}}\big)$ onto X, or onto X less one or two endpoints. Thus, assuming the Lemma holds for k, we have seen it also holds for $k + 1$, and the result follows by induction. $\qquad \square$

Lemma 2.3. *Let X be an interval of length 1 and let f be a baker's transformation relative to a standard partition $(X_j)_{j \in I}$ of X. Let $k \in \mathbb{N}$ and let $n_1, n_2, \ldots, n_k \in I$. Then,*

$$\mu(X_{n_1 n_2 \ldots n_k}) = \prod_{j=1}^{k} \mu(X_{n_j}).$$

Proof. If $k = 1$, the result is obviously true. So, assume that for some k, the result holds for any choice of $n_1, n_2, \ldots, n_k \in I$. Then

$$\mu(X_{n_1 n_2 \ldots n_k n_{k+1}}) = \mu\left(\big\{ x : x \in X_{n_1 n_2 \ldots n_k} \text{ and } f^k(x) \in X_{n_{k+1}} \big\} \right),$$

$$= \mu(X_{n_1 n_2 \ldots n_k}) \cdot \mu(X_{n_{k+1}}), \quad \text{by using Lemma 2.2,}$$

$$= \mu(X_{n_{k+1}}) \cdot \prod_{j=1}^{k} \mu(X_{n_j}),$$

as the result is assumed true for k,

$$= \prod_{j=1}^{k+1} \mu(X_{n_j}).$$

Thus, if the result is true for k, it is true for $k+1$. The result follows by induction.
$\square$

Lemma 2.4. *Let X be an interval of length 1 and let f be a baker's transformation relative to a standard partition $(X_j)_{j \in I}$ of X. Let $k, s \in \mathbb{N}$ and for each $j \in \{1, 2, \ldots, s\}$, suppose there are given $n_{j1}, n_{j2}, \ldots, n_{jk} \in I$ and put*

$$Y_j = X_{n_{j1} n_{j2} \cdots n_{jk}}.$$

Then,

$$\mu\Big(Y_1 \cap f^{-k}(Y_2) \cap f^{-2k}(Y_3) \cdots \cap f^{-(s-1)k}(Y_s)\Big) = \prod_{j=1}^{s} \mu(Y_j).$$

Proof. Observe from the definitions that

$$Y_1 \cap f^{-k}(Y_2) \cap f^{-2k}(Y_3) \cdots \cap f^{-(s-1)k}(Y_s)$$
$$= X_{n_{11} \cdots n_{1k} n_{21} \cdots n_{2k} n_{31} \cdots n_{s-1\,k} n_{s1} \cdots n_{sk}}.$$

Then, by Lemma 2.3,

$$\mu(Y_1 \cap f^{-k}(Y_2) \cap \cdots \cap f^{-(s-1)k}(Y_s))$$
$$= \mu(X_{n_{11}}) \cdots \mu(X_{n_{1k}}) \cdots \mu(X_{n_{s1}}) \cdots \mu(X_{n_{sk}})$$
$$= \prod_{j=1}^{s} \mu(Y_j).$$

$\square$

Let $\mathcal{P} = (X_j)_{j \in I}$ be a standard partition of an interval X of length 1. Given $k \in \mathbb{N}$, make the definition that

$$\mathcal{P}_k = \Big\{ X_{n_1 n_2 \ldots n_k} : n_1, n_2, \ldots, n_k \in I \Big\}.$$

Then $\mathcal{P}_k$ is a partition of X in that distinct sets of $\mathcal{P}_k$ are disjoint and the union of the sets in $\mathcal{P}_k$ is X. Lemma 2.2 shows that the sets in $\mathcal{P}_k$ are intervals of positive length. In the case when I has r elements, $\mathcal{P}_k$ consists of r^k disjoint sets. If $\ell > k$, each set in $\mathcal{P}_k$ is the union of $2^{\ell-k}$ sets in $\mathcal{P}_\ell$, as expressed by the identity

$$X_{n_1 n_2 \ldots n_k} = \bigcup \Big\{ X_{n_1 n_2 \ldots n_k n_{k+1} n_{k+2} \ldots n_\ell} : n_{k+1}, n_{k+2}, \ldots, n_\ell \in I \Big\}.$$

Lemma 2.5. *Let X be an interval of length 1 and let f be a baker's transformation relative to a standard partition $\mathcal{P} = (X_j)_{j \in I}$ of X. Let $k \in \mathbb{N}$ and let Z be a set that is a union of some family of sets in the partition $\mathcal{P}_k$ of X. Then, for $\ell = 1, 2, \ldots,$*

$$\mu\Big(Z \cap f^{-k}(Z) \cap f^{-2k}(Z) \cap \cdots \cap f^{-(\ell-1)k}(Z)\Big) = \mu(Z)^\ell.$$

Proof. First, assume that Z is a *finite* union of sets in $\mathcal{P}_k$. Then, there are disjoint sets $Y_1, Y_2, \ldots, Y_s$ in $\mathcal{P}_k$ such that $Z = \cup_{j=1}^{s} Y_j$. We have

$$\mu\Big(Z \cap f^{-k}(Z) \cap f^{-2k}(Z) \cap \cdots \cap f^{-(\ell-1)k}(Z)\Big)$$

$$= \mu\Big((\cup_{j=1}^{s} Y_j) \cap f^{-k}(\cup_{j=1}^{s} Y_j) \cap f^{-2k}(\cup_{j=1}^{s} Y_j) \cap \cdots \cap f^{-(\ell-1)k}(\cup_{j=1}^{s} Y_j)\Big)$$

$$= \mu\Big((\cup_{j=1}^{s} Y_j) \cap (\cup_{j=1}^{s} f^{-k}(Y_j)) \cap \cdots \cap (\cup_{j=1}^{s} f^{-(\ell-1)k}(Y_j))\Big)$$

$$= \mu\left(\bigcup_{j_0,j_1,\ldots,j_{\ell-1}\in\{1,\ldots,s\}} Y_{j_0} \cap f^{-k}(Y_{j_1}) \cap \cdots \cap f^{-(\ell-1)k}(Y_{j_{\ell-1}})\right)$$

$$= \sum_{j_0,j_1,\ldots,j_{\ell-1}\in\{1,\ldots,s\}} \mu\left(Y_{j_0} \cap f^{-k}(Y_{j_1}) \cap \cdots \cap f^{-(\ell-1)k}(Y_{j_{\ell-1}})\right)$$

$$= \sum_{j_0,j_1,\ldots,j_{\ell-1}\in\{1,\ldots,s\}} \prod_{t=0}^{\ell-1} \mu(Y_{j_t}), \text{ by Lemma 2.4,}$$

$$= \left(\sum_{j=1}^{s} \mu(Y_j)\right)^{\ell}$$

$$(2.2) \quad = \mu(Z)^{\ell}.$$

Now, when Z is an *infinite* union of sets in $\mathcal{P}_k$, there is an increasing sequence (Z_n) of sets, each set Z_n is a finite union of sets in $\mathcal{P}_k$, and $Z = \cup_{n=1}^{\infty} Z_n$. Using (2.2) for the sets Z_n, and the fact that μ is σ-additive on the Borel sets, if $k, \ell \in \mathbb{N}$ are given we now have

$$\mu\Big(Z \cap f^{-k}(Z) \cap f^{-2k}(Z) \cap \cdots \cap f^{-(\ell-1)k}(Z)\Big)$$

$$= \lim_{n\to\infty} \mu\Big(Z_n \cap f^{-k}(Z_n) \cap f^{-2k}(Z_n) \cap \cdots \cap f^{-(\ell-1)k}(Z_n)\Big)$$

$$= \lim_{n\to\infty} \mu(Z_n)^{\ell}$$

$$= \mu(Z)^{\ell}.$$

Lemma 2.6. *Let X be an interval of length 1, let $\mathcal{P} = (X_j)_{j\in I}$ be a standard partition of X and put $\rho = \max\{\mu(X_j) : j \in I\}$. Then, for each $k \in \mathbb{N}$, $\mathcal{P}_k$ is a standard partition of X, and we have*

$$\max\{\mu(Y) : Y \in \mathcal{P}_k\} \le \rho^k \text{ and } \lim_{k\to\infty} (\max\{\mu(Y) : Y \in \mathcal{P}_k\}) = 0.$$

Proof. This is immediate from (2.1), and Lemmas 2.2 and 2.3. $\qquad\square$

The following argument extends an idea in [9, pages 150-151] to prove ergodicity for the dynamical systems considered here. We use A^c to denote the complement of a set A.

Lemma 2.7. *Let X be an interval of length 1 and let f be a baker's transformation relative to a standard partition $(X_j)_{j\in I}$ of X. Then if A is a Borel subset of X such that $f^{-1}(A) = A$, then $\mu(A) = 0$ or $\mu(A) = 1$. That is, (X, f) is an ergodic system.*

Proof. Let A be a Borel subset of X with $f^{-1}(A) = A$ and $0 \le \mu(A) < 1$. Then, $0 < \mu(A^c) \le 1$. Let $\varepsilon > 0$. Then, there is a closed subinterval U of X such that $\mu(U) > 0$ and $\mu(U \cap A^c) > (1 - \varepsilon/2)\mu(U)$. Now by Lemma 2.6 there is $k \in \mathbb{N}$ and a sequence $V_1, V_2, \ldots$ of disjoint sets in $\mathcal{P}_k$ such that

$$(2.3) \qquad \bigcup_{n=1}^{\infty} V_n \subseteq U \text{ and } \mu\left(U \cap \left(\bigcup_{n=1}^{\infty} V_n\right)^c\right) < \frac{\varepsilon\,\mu(U)}{2}.$$

Note that if the partition $(X_j)_{j \in I}$ is finite, the sequence (V_n) will terminate, but when this is allowed for, the ensuing argument remains valid. Using (2.3) we now have

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu(A^c \cap V_n) &= \mu\left(A^c \cap \left(\bigcup_{n=1}^{\infty} V_n\right)\right) \\
&= \mu\left(U \cap A^c \cap \left(\bigcup_{n=1}^{\infty} V_n\right)\right) \\
&= \mu(U \cap A^c) - \mu\left(U \cap A^c \cap \left(\bigcup_{n=1}^{\infty} V_n\right)^c\right) \\
&\ge \mu(U \cap A^c) - \mu\left(U \cap \left(\bigcup_{n=1}^{\infty} V_n\right)^c\right) \\
&> \left(1 - \frac{\varepsilon}{2}\right)\mu(U) - \frac{\varepsilon\,\mu(U)}{2} \\
&= (1 - \varepsilon)\,\mu(U).
\end{aligned}$$

Hence, we have

$$\sum_{n=1}^{\infty} \mu(V_n \cap A^c) > (1 - \varepsilon)\mu\left(\bigcup_{n=1}^{\infty} V_n\right) = (1 - \varepsilon)\sum_{n=1}^{\infty} \mu(V_n).$$

By choosing a suitable V_n, and putting $W = V_n$, we deduce that there is $W \in \mathcal{P}_k$ such that

$$(2.4) \qquad \mu(W \cap A^c) > (1 - \varepsilon)\mu(W).$$

Now, by Lemma 2.2, $f^k : W \longrightarrow [0,1]$ is onto except maybe for one or two endpoints. Let $B = A^c$ except that the endpoints of X that are in A^c, if any, are omitted. Put $C = \{x : x \in W \text{ and } f^k(x) \in B\}$. Note that

$$(2.5) \qquad \mu(B) = \mu(A^c) \text{ and } f^k(C) = B.$$

Also, as $f^{-1}(A) = A$, $f^{-1}(A^c) = A^c$ and so $f^{-k}(A^c) = A^c$. It follows that if $x \in C$, $f^k(x) \in B \subseteq A^c$ and $x \in f^{-k}(A^c) = A^c$. Hence, $C \subseteq A^c \cap W$.

Now as f is one-to-one on W and maps W onto X except perhaps for the endpoints, if it exists we let $w_0 \in W$ be the point in W such that $f^k(w_0)$ is the left endpoint of X and, if it exists, we let $w_1 \in W$ be the point in W such that $f^k(w_1)$ is the right endpoint of X. Put $D = \emptyset, \{w_0\}, \{w_1\}$ or $\{w_0, w_1\}$, depending on the existence of w_0, w_1. Then,

$$(2.6) \qquad A^c \cap W \cap D^c \subseteq C \subseteq A^c \cap W.$$

We now have, using (2.4), (2.5) and (2.6),

$$
\begin{aligned}
\mu(A^c) &= \mu(B) \\
&= \mu(f^k(C)) \\
&= \frac{\mu(C)}{\mu(W)}, \text{ as } f^k \text{ is linear on } W, \\
&\geq \frac{\mu(A^c \cap W)}{\mu(W)} \\
&> 1 - \varepsilon.
\end{aligned}
$$

As this holds for all $\varepsilon > 0$, we see that $\mu(A^c) = 1$, and so $\mu(A) = 0$. Thus, (X, f) is an ergodic system, by definition. $\qquad\square$

Definitions. Let f be a baker's transformation on an interval X of length 1 relative to a standard partition $(X_j)_{j \in I}$ of X. Let U be a subinterval of X of positive length and let $\Theta_U(x)$ denote the corresponding recurrence time for $x \in U$. Then, the *average* or *expectation* $E(\Theta_U)$ of Θ_U over U is defined as

$$
E(\Theta_U) = \frac{1}{\mu(U)} \sum_{n=1}^{\infty} n \mu \left(\{x : \Theta_U(x) = n\} \right).
$$

Also, the *standard deviation* of Θ_U over U is defined as

$$
\sigma(\Theta_U) = \sqrt{\frac{1}{\mu(U)} \sum_{n=1}^{\infty} |n - E(\Theta_U)|^2 \, \mu \left(\{x : \Theta_U(x) = n\} \right)}.
$$

Theorem 2.8. *Let f be a baker's transformation on an interval X of length 1 relative to a standard partition $\mathcal{P} = (X_j)_{j \in I}$ of X. Let $\rho = \max\{\mu(X_j) : j \in I\}$, as in Lemma 2.6. Let U be a subset of X that is a union of r disjoint subintervals $U_1, U_2, \ldots, U_r$ of positive length, and put $\tau = \min\{\mu(U_j) : 1 \leq j \leq r\}$. Then, the following hold.*

(i) The average of the recurrence function θ_U of U relative to f is $\mu(U)^{-1}$.

(ii) The standard deviation $\sigma(\theta_U)$ of θ_U is finite.

(iii) Let $\lfloor x \rfloor$ denote the integer part of $x \in \mathbb{R}$, and put

$$
\nu(U) = 1 + \left\lfloor \frac{\log(\tau/2)}{\log \rho} \right\rfloor.
$$

Then, $2r\rho^{\nu(U)} < \mu(U)$ and

$$
(2.7) \qquad \sigma(\theta_U) \leq \sqrt{-\frac{1}{\mu(U)^2} + \frac{3}{\mu(U)} - 2 + \frac{2\nu(U)}{\mu(U)} \left(\frac{1}{\mu(U) - 2r\rho^{\nu(U)}} - 1 \right)}.
$$

Proof. (i) By Lemma 2.7, the system (X, f) is ergodic, so that the fact that the average of Θ_U over U is is $1/\mu(U)$ is immediate from Kac's formula [3], [7]. Alternatively, it can be deduced as a consequence of Theorem 4.15 in [10], by applying part of the proof of that Theorem to show the assumptions in Theorem 4.11 in [10] are satisfied.

(ii) and (iii). If $U = X$ except possibly for end points, then $\sigma(\Theta_U) = 0$, so the result is true in this case. So, we may assume that $0 < \mu(U) < 1$. By Lemma 2.6, there is $k \in \mathbb{N}$ and a set Y that is a non-empty union of sets in $\mathcal{P}_k$ such that $Y \subseteq U$. Note that there may be many such k and sets Y such that this holds; we

simply choose any such pair k, Y. Put $Z = Y^c$. Then, Z is also a union of sets in $\mathcal{P}_k$, $U^c \subseteq Z$ and $0 < \mu(Z) < 1$. Now, using any set Z obtained in this way we have

$$\sum_{n=1}^{\infty} \mu(U^c \cap f^{-1}(U^c) \cap \cdots \cap f^{-n}(U^c))$$

$$= \sum_{\ell=1}^{\infty} \left(\sum_{n=(\ell-1)k+1}^{\ell k} \mu(U^c \cap f^{-1}(U^c) \cap \cdots \cap f^{-n}(U^c)) \right)$$

$$\leq \sum_{\ell=1}^{\infty} \left(\sum_{n=(\ell-1)k+1}^{\ell k} \mu(Z \cap f^{-1}(Z) \cap \cdots \cap f^{-n}(Z)) \right)$$

$$\leq \sum_{\ell=1}^{\infty} \left(\sum_{n=(\ell-1)k+1}^{\ell k} \mu(Z \cap f^{-k}(Z) \cap \cdots \cap f^{-(\ell-1)k}(Z)) \right)$$

$$= \sum_{\ell=1}^{\infty} k\mu(Z)^{\ell}, \text{ by Lemma 2.5,}$$

$$= \frac{k\mu(Z)}{1 - \mu(Z)}$$

$$(2.8) \qquad = k \left(\frac{1}{\mu(Y)} - 1 \right)$$

$$< \infty.$$

By Theorem 4.15 of [10, page 275], (2.8) implies that the standard deviation $\sigma(\theta_U)$ is finite and, what is more, it implies that formula (4.39) in [10] holds for the standard deviation of Θ_U in the present case.

To get an upper bound for $\sigma(\Theta_U)$, we choose k and Y in (2.8) in a particular way, noting that k and Y are inter-related. We have from the definition of $\nu(U)$ that $2\rho^{\nu(U)} < \tau$, so for each $j = 1, 2, \ldots, r$ we have $2\rho^{\nu(U)} < \mu(U_j)$. Thus,

$$(2.9) \qquad 2r\rho^{\nu(U)} < \sum_{j=1}^{r} \mu(U_j) = \mu(U).$$

Also, using Lemma 2.6, we now see that if we put Y_j for the union of *all* the sets in $\mathcal{P}_{\nu(U)}$ that are subsets of U_j, then for each $j = 1, 2, \ldots, r$ we have

$$0 < \mu(Y_j) \leq \mu(U_j) < \mu(Y_j) + 2\rho^{\nu(U)}.$$

So, if we put $Y = \cup_{j=1}^{r} Y_j$, this union is disjoint and we have

$$(2.10) \qquad \mu(U) = \sum_{j=1}^{r} \mu(U_j) < \sum_{j=1}^{r} \mu(Y_j) + 2r\rho^{\nu(U)} = \mu(Y) + 2r\rho^{\nu(U)}.$$

Using (2.9) and (2.10), and taking $\nu(U)$ for k in (2.8), we now have

$$\sum_{n=1}^{\infty} \mu(U^c \cap f^{-1}(U^c) \cap \cdots \cap f^{-n}(U^c))$$

$$\leq \nu(U) \left(\frac{1}{\mu(Y)} - 1 \right)$$

$$\leq \nu(U) \left(\frac{1}{\mu(U) - 2r\rho^{\nu(U)}} - 1 \right).$$

It now follows from (2.8) and the formula for the standard deviation in Theorem 4.15 in [10] that the expression inside the square root sign in (2.7) is non-negative and that (2.7) holds. $\qquad \square$

The estimate in (2.7) can be simplified in the special case when the partition $\mathcal{P}$ is finite and there is $k \in \mathbb{N}$ such that the intervals $U_1, U_2, \ldots, U_r$ are in $\mathcal{P}_k$. This means that U is a finite union of sets in $\mathcal{P}_k$. Using the notations in the above proof we have, by Lemma 2.6, that $\mu(U) = \sum_{j=1}^{r} \mu(U_j) \leq r\rho^k$. Consequently,

$$k \leq \frac{\log\left(\dfrac{\mu(U)}{r}\right)}{\log \rho}.$$

Now putting this value of k and putting U for Y in (2.8), we have

$$\sum_{n=1}^{\infty} \mu(U^c \cap f^{-1}(U^c) \cap \cdots \cap f^{-n}(U^c)) \leq \frac{\log\left(\dfrac{\mu(U)}{r}\right)}{\log \rho} \left(\frac{1}{\mu(U)} - 1 \right),$$

and instead of (2.7) we have the simpler-looking estimate

$$\sigma(\theta_U) \leq \sqrt{-\frac{1}{\mu(U)^2} + \frac{3}{\mu(U)} - 2 + \frac{2\log\left(\dfrac{\mu(U)}{r}\right)}{\mu(U) \log \rho} \left(\frac{1}{\mu(U)} - 1 \right)}.$$

Note that the estimate for $\sigma(\Theta_U)$ in (2.7) cannot be expected to be the best possible. One reason is the fact that the inequality (2.8), upon which (2.7) depends, was obtained by replacing smaller terms by larger ones. However, note that in a general dynamical system the standard deviation of the recurrence times need not be finite [6].

References

[1] S. K. Berberian, *Measure and Integration* (Macmillan, New York, 1965).
[2] J. R. Blum and J. I. Rosenblatt, On the moments of recurrence time, *Journ. Math. Sci. (Delhi)* **2** (1967) 1-6.
[3] J. R. Brown, *Ergodic Theory and Topological Dynamics* (Academic Press, New York 1976).
[4] K. Dajani and C. Kraiikamp, *Ergodic Theory of Numbers* (Carus Monographs series volume 29, Mathematical Association of America, Washington DC, 2002).
[5] M. Ismael, *Recurrence times in stochastic processes and dynamical systems*, master's thesis (University of Wollongong, 2014).
[6] M. Ismael, R. Nillsen and G. Williams, Variation of recurrence times in discrete dynamical systems, *Advances and Applications in Discrete Mathematics.* **14** (2014) 95-123.
[7] M. Kac, On the notion of recurrence in discrete stochastic processes, *Bull. Amer. Math. Soc.* **53** (1947) 1002-1010.
[8] P. W. Kasteleyn, Variations on a theme of Mark Kac, *Jour. Stat. Phys.* **46** (1987) 811-827.

[9] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Cambridge University Press, Cambridge, 1995).

[10] R. Nillsen, *Randomness and Recurrence in Dynamical Systems* (Carus Monographs series volume 31, Mathematical Association of America, 2010).

[11] K. Petersen, *Ergodic Theory* (Cambridge Studies in Advanced Mathematics 2, Cambridge University Press, 1983).

[12] A. G. Postnikov, *Ergodic Problems in the Theory of Congruences and of Diophantine Approximations, translated by B. Volkmann*, Proceedings of the Steklov Institute of Mathematics, **82** (1966) (American Mathematical Society, Washington DC, 1967).