Lower bounds for ergodic averages

A. Leibman

Abstract

We compute the exact lower bounds for some averages arising in ergodic theory. In particular, we prove that for any measure preserving system $(X,\mathcal{B},\mu,T)$ with $\mu(X)<\infty$, any $A\in\mathcal{B}$ and any $N\in\mathbb{N}$, $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A)\geq\sqrt{\mu(A)^2+(\mu(X)-\mu(A))^2}+\mu(A)-\mu(X)$.

1. Lower bound for the averages $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A)$

1.1. Let T be a measure preserving transformation of a probability measure space $(X,\mathcal{B},\mu)$. Let $0<a\leq 1$; it follows from the mean ergodic theorem that if A is a subset of X with $\mu(A)\geq a$, then the limit of the averages

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) \tag{1.1}$$

exists and satisfies $\lim\limits_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A)\geq a^2$ ([Kh]). This does not apriori guarantee that there is a uniform positive lower bound of the averages (1.1) for all A with $\mu(A)\geq a$, that is, that there is $c=c(a)>0$ such that for any X, T and A with $\mu(A)\geq a$ one has $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A)\geq b$ for all $N\in\mathbb{N}$. Indeed, for the more general expressions $\frac{1}{N-M}\sum_{n=M}^{N-1}\mu(A\cap T^{-n}A)$ one still has $\lim\limits_{N-M\to\infty}\frac{1}{N-M}\sum_{n=M}^{N-1}\mu(A\cap T^{-n}A)\geq a^2$ ([Kh]), while, if $a<\frac{1}{2}$, for arbitrarily large $N-M$ one may have $\frac{1}{N-M}\sum_{n=M}^{N-1}\mu(A\cap T^{-n}A)=0$ for appropriately chosen T, A and M. (For example, take $X=[0,1]$, $A=[0,a]$ with $a<\frac{1}{2}$ and $T(x)=(x+\alpha)\bmod 1$ with $\alpha\ll 1-2a$; then there are large intervals of n for which $\mu(A\cap T^{-n}A)=\emptyset$.)

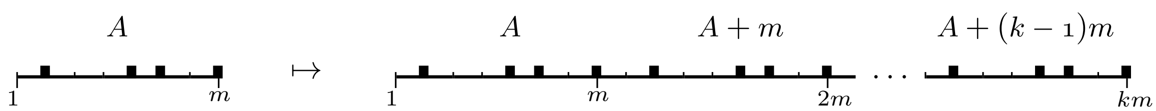
The existence of positive lower bound for averages of the form $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T_1^{-n}A\cap\ldots\cap T_k^{-n}A)$, where $T_1,\ldots,T_k$ are pairwise commuting measure preserving transformations of X, is proven in [BHMP]. We compute the exact lower bound of the averages (1.1):

This work was supported by NSF grant DMS-9706057, by the OSU Seed grant and by the Sloan Foundation grant BR-3969

1.2. Theorem. *Let* $0 \le a \le 1$*.*
(a) *For any probability measure preserving system* $(X, \mathcal{B}, \mu, T)$ *and any* $A \in \mathcal{B}$ *with* $\mu(A) \ge a$ *one has* $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A \cap T^{-n}A) \ge \sqrt{a^2 + (1-a)^2} + a - 1$ *for all* $N \in \mathbb{N}$*.*
(b) *For any* $\delta > 0$ *there exist a measure preserving system* $(X, \mathcal{B}, \mu, T)$*,* $A \in \mathcal{B}$ *with* $\mu(A) = a$ *and* $N \in \mathbb{N}$ *such that* $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A \cap T^{-n}A) < \sqrt{a^2 + (1-a)^2} + a - 1 + \delta$*.*

Proof. Passing, if needed, to the natural extension of $(X, \mathcal{B}, \mu, T)$ ([R]), we may assume that T is invertible. We may also assume that X is finite with $\mu(B) = |B|/|X|$, $B \in \mathcal{B}$. Indeed, given $A \in \mathcal{B}$, $\mu(A) = a$, for any $N \in \mathbb{N}$ and $\varepsilon > 0$ there exists a finite set $\hat{X}$, a permutation $\hat{T}$ of $\hat{X}$ and a set $\hat{A} \subseteq \hat{X}$ such that $\left|\frac{|\hat{A}|}{|\hat{X}|} - a\right| < \varepsilon$ and $\left|\frac{|\hat{A} \cap \hat{T}^{-n}\hat{A}|}{|\hat{X}|} - \mu(A \cap T^{-n}A)\right| < \varepsilon$ for all $n \le N$. (One can deduce this fact from the Rohlin lemma, or prove it directly.) Thus, we arrive at the following problem: given a permutation T of a finite set X, a subset A of X with $|A| = a|X|$ and $N \in \mathbb{N}$, we have to estimate $\frac{1}{N|X|}\sum_{n=0}^{N-1}|A \cap T^{-n}A|$.

First, let us assume that T is a cyclic permutation: $X = \{1, \dots, m\}$ and $Tx = (x \bmod m) + 1$. Let $A \subseteq \{1, \dots, m\}$ with $|A| = b = ma$. For any $k \in \mathbb{N}$, if we replace X by $\{1, \dots, km\}$ and A by $A \cup (A + m) \cup \ldots \cup (A + (k-1)m)$:

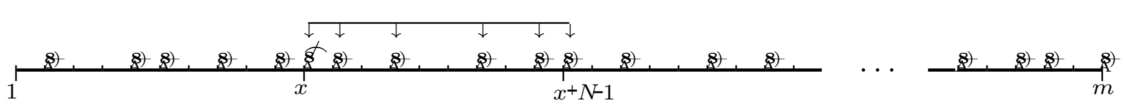


then the quantities $|A|/|X|$ and $|A \cap T^{-n}A|/|X|$, $n \in \mathbb{Z}$, do not change. Hence, we may assume that m is arbitrarily large. Fix $\varepsilon > 0$ and assume that $N/m < \varepsilon$. Under this assumption, we will estimate from below the sum

$$S = \sum_{n=0}^{N-1}\left|A \cap (A-n)\right| = \sum_{x \in A}\left|A \cap [x, x+N-1]\right|,$$

which does not exceed $\sum_{n=0}^{N-1}|A \cap T^{-n}A|$.

To make the argument more transparent, let us reformulate the problem in combinatorial language. Assume that b *archers* are positioned at the points $1, 2, \dots, m$ of the real line, no more than one archer at a point: there is an archer at x iff $x \in A$. Every archer *threatens* himself and all other archers positioned at his right at the distance $< N$. (That is, the archer located at a point x threatens the archers located in the interval $[x, x+N-1]$.)

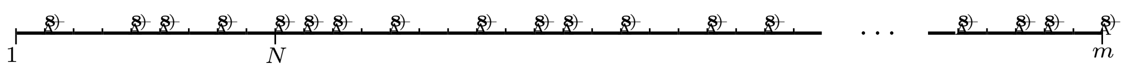


The question is: how should one position the archers in order to minimize "the total number of threats"

$$S = \sum_{R \text{ is an archer}} \text{the number of archers threatened by } R,$$

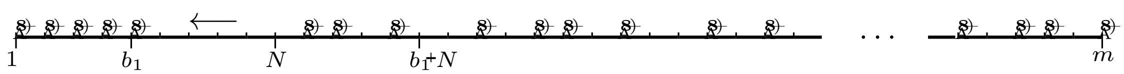
and what is the minimal value of S?

We start with an arbitrary positioning of archers

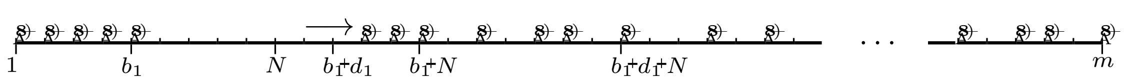


and will "improve" it by moving the archers in such a way that S will not increase.

Step 1. Assume that b_1 archers are located at the points of the interval $[1, N]$. If $b_1 > 0$, we move these archers to the left end of the interval $[1, N]$; clearly, this does not increase S. As a result, all (integer) points in the interval $[1, b_1]$ become occupied (we will say that $[1, b_1]$ *is full*), while all points in the interval $[b_1 + 1, N]$ become free (we will say that $[b_1 + 1, N]$ *is empty*):

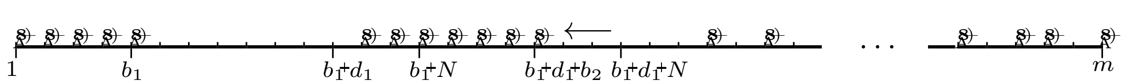


Step 2. Now, if an archer R is located at a point $x \in [N, b_1 + N - 1]$ and the point $x + 1$ is not occupied, then R can be moved to $x + 1$. Indeed, after this relocation R is no longer threatened by the archer located at $x - N + 1 \in [1, b_1]$ and so, the number of archers threatening R decreases by 1. On the other hand, the number of archers threatened by R increases by at most 1 and, hence, the total number of threats S does not increase. This allows us to move all archers located in $[N + 1, b_1 + N]$ to the right end of this interval:



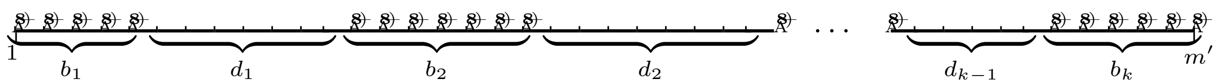
Assume that there are c_1 archers in $[N + 1, b_1 + N]$ (possibly, $c_1 = 0$) and put $d_1 = N - c_1$; then after this rearrangement the interval $[N + 1, b_1 + d_1]$ becomes empty and the interval $[b_1 + d_1 + 1, b_1 + N]$ becomes full. Note that $c_1 \le b_1$ and so, $b_1 + d_1 \ge N$.

Step 3. We shift the archers located in $[b_1 + N + 1, b_1 + N + d_1]$ to the left end of this interval; we can do this since, at any position, these archers are threatened by all archers from the interval $[b_1 + d_1 + 1, b_1 + N]$ and are not threatened by the archers from $[1, b_1]$:



Assume that the interval $[b_1 + N + 1, b_1 + N + d_1]$ contains e_1 archers and put $b_2 = c_1 + e_1$. Then, after this rearrangement, the interval $[b_1 + d_1 + 1, b_1 + d_1 + b_2]$ becomes full and the interval $[b_1 + d_1 + b_2 + 1, b_1 + d_1 + N]$ becomes empty. Note that $b_2 \ge c_1$ and so, $d_1 + b_2 \ge d_1 + c_1 = N$.

We repeat Steps 2 and 3 starting at the point b_1+d_1+1 instead of 1, and obtain an empty interval $[b_1+d_1+b_2+1, b_1+d_1+b_2+d_2]$ and a full interval $[b_1+d_1+b_2+d_2+1, b_1+d_1+b_2+d_2+b_3]$. And so on, until we reach the last archer. In the process of the last application of Step 2 some archers will possibly be forced to cross the boundary of the interval $[1,m]$ and move to the interval $[m+1,m']$ with $m' \le m+N$. The resulting configuration will represent an alternating sequence of full/empty intervals of lengths, respectively, $b_1, d_1, \dots, b_{k-1}, d_{k-1}, b_k$, where b_i, d_i satisfy $0 \le b_i \le N$ for $i = 1, \dots, k$; $0 \le d_i \le N$, $b_i + d_i \ge N$ and $d_i + b_{i+1} \ge N$ for $i = 1, \dots, k-1$; $b_1 + \ldots + b_k = b$ and $d_1 + \ldots + d_{k-1} = m' - b$.



In this situation, the first (from the left) archer of the i-th group of archers threatens all b_i members of this group, the next one threatens $b_i - 1$ archers, and so on. In addition, the last archer of the i-th group threatens $N - d_i - 1$ members of the $(i+1)$-st group, the next-to-last archer threatens $N - d_i - 2$ archers of the $(i+1)$-st group, and so on. Hence, the number of threats coming from the members of the i-th group is

$$\big(b_i + (b_i - 1) + \ldots + 1\big) + \big((N-d_i-1) + (N-d_i-2) + \ldots + 1\big) = \tfrac{b_i(b_i+1)}{2} + \tfrac{(N-d_i)(N-d_i-1)}{2}.$$

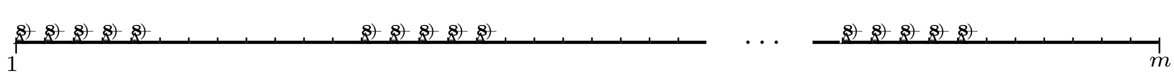
The total number of threats S is, therefore,

$$\begin{aligned} S &= \sum_{i=1}^{k} \frac{b_i(b_i+1)}{2} + \sum_{i=1}^{k-1} \frac{(N-d_i)(N-d_i-1)}{2} \\ &= \frac{1}{2}\sum_{i=1}^{k} b_i^2 + \frac{1}{2}\sum_{i=1}^{k-1}(N-d_i)^2 + \frac{1}{2}\sum_{i=1}^{k} b_i - \frac{1}{2}\sum_{i=1}^{k-1}(N-d_i) \\ &\ge \frac{1}{2k}\Big(\sum_{i=1}^{k} b_i\Big)^2 + \frac{1}{2(k-1)}\Big(\sum_{i=1}^{k-1}(N-d_i)\Big)^2 + \frac{1}{2}\sum_{i=1}^{k-1}(b_i + d_i - N) \\ &\ge \frac{1}{2k}b^2 + \frac{1}{2k}\big((k-1)N - m' + b\big)^2 = \frac{1}{2k}\big(b^2 + (kN - m'' + b)^2\big), \end{aligned} \tag{1.2}$$

where $m'' = M' + N$. Considering the right hand part of (1.2) as a function of k, one finds that its minimum is reached when $k = \frac{\sqrt{b^2+(m''-b)^2}}{N}$ and equals

$$N\sqrt{b^2 + (m''-b)^2} - N(m''-b) = mN\Big(\sqrt{a^2 + (\tfrac{m''}{m} - a)^2} + a - \tfrac{m''}{m}\Big).$$

Since $1 < \frac{m''}{m} \le \frac{m+2N}{m} < 1 + 2\varepsilon$ and ε can be taken arbitrarily small, we have $S \ge mN\big(\sqrt{a^2 + (1-a)^2} + a - 1\big)$. (Returning to the archers, we see that, if we ignore the fact that k, b/k and m/k must be integers, the "safest" configuration is the following one: the b archers form $k = \frac{\sqrt{b^2+(m-b)^2}}{N}$ equal groups with equal distances between the groups:



For this configuration $S = N\sqrt{b^2+(m-b)^2} - N(m-b)$.)

We obtain, therefore, that in the case T is a cyclic permutation,

$$\frac{1}{N}\sum_{n=0}^{N-1}|A\cap T^{-n}A| \geq \frac{1}{N}\sum_{n=0}^{N-1}|A\cap(A-n)| = \tfrac{1}{N}S \geq m\big(\sqrt{a^2+(1-a)^2}+a-1\big).$$

Now let T be an arbitrary permutation of an m-element set X. Let $X = X_1\cup\ldots\cup X_l$ be the partition of X into the union of disjoint cycles of T and let $m_j = |X_j|$, $j = 1,\ldots,l$. Let $A\subseteq X$, $|A| = b$, $A_j = A\cap X_j$ and $a_j = |A_j|/|X_j|$, $j = 1,\ldots,l$. Then for any $N\in\mathbb{N}$ we have

$$\frac{1}{N}\sum_{n=0}^{N-1}|A\cap T^{-n}A| = \frac{1}{N}\sum_{j=1}^{l}\sum_{n=0}^{N-1}|A_j\cap T^{-n}A_j| \geq \sum_{j=1}^{l}m_j\big(\sqrt{a_j^2+(1-a_j)^2}+a_j-1\big).$$

Since the function $\varphi(a) = \sqrt{a^2+(1-a)^2}+a-1$ is convex, the conditions $m_1+\ldots+m_l = m$ and $\frac{1}{m}(a_1m_1+\ldots+a_lm_l) = a$ imply $\sum_{j=1}^{l}m_j\varphi(a_j) \geq m\varphi(a)$. Hence,

$$\frac{1}{N}\sum_{n=0}^{N-1}|A\cap T^{-n}A| \geq m\big(\sqrt{a^2+(1-a)^2}+a-1\big)$$

and

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) = \frac{1}{mN}\sum_{n=0}^{N-1}|A\cap T^{-n}A| \geq \sqrt{a^2+(1-a)^2}+a-1.$$

To prove part (b) of the theorem, we take $X = [0,1]$, $A = [0,a]$ and $T(x) = (x+\frac{1}{m}) \bmod 1$ with m to be specified later. We may assume that a is rational and, moreover, that $a = \frac{b}{m}$, $b\in\mathbb{N}$. Then for $m-b\leq N\leq m$ we have

$$\begin{aligned}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) &= \frac{a}{mN}\Big(\frac{b(b+1)}{2}+\frac{(N-m+b)(N-m+b-1)}{2}\Big)\\ &= \frac{1}{2y}\big(a(a+\tfrac{1}{m})+(y+a-1)(y+a-1-\tfrac{1}{m})\big),\end{aligned}\tag{1.3}$$

where we put $y = N/m$. By taking m large enough we may make (1.3) to be less then $\frac{1}{2y}\big(a^2+(y+a-1)^2\big)+\frac{\delta}{2}$ for all $y\in[0,1]$. For $y = \sqrt{a^2+(1-a)^2}$ one has $\frac{1}{2y}\big(a^2+(y+a-1)^2\big) = \sqrt{a^2+(1-a)^2}+a-1$. Therefore, choosing N and m so that $y = \frac{N}{m}$ is sufficiently close to $\sqrt{a^2+(1-a)^2}$, we get $\frac{1}{2y}\big(a^2+(y+a-1)^2\big) < \sqrt{a^2+(1-a)^2}+a-1+\frac{\delta}{2}$ and so, $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) < \sqrt{a^2+(1-a)^2}+a-1+\delta$. ∎

1.3. Given $a>0$, a positive lower bound also exists for the averages $\frac{1}{N}\sum_{n=0}^{N-1}\int fT^nf\,d\mu$ where f is a nonnegative function with $\int f\,d\mu=a$:

Theorem. *Let $a>0$.*
(a) *For any probability measure preserving system $(X,\mathcal{B},\mu,T)$ and any nonnegative integrable function f on X with $\int f\,d\mu\geq a$ one has $\frac{1}{N}\sum_{n=0}^{N-1}\int fT^nf\,d\mu\geq\frac{a^2}{2}$ for all $N\in\mathbb{N}$.*
(b) *For any $\delta>0$ there exist a measure preserving system $(X,\mathcal{B},\mu,T)$, a measurable function f on X with $\int f\,d\mu=a$ and $N\in\mathbb{N}$ such that $\frac{1}{N}\sum_{n=0}^{N-1}\int fT^nf\,d\mu<\frac{a^2}{2}+\delta$.*

Proof. Fix $N\in\mathbb{N}$. Again, we may replace our system by a finite one and assume that T is a permutation of a finite set X, $|X|=m$, and that f takes on only integer values. We have to estimate the sum $\sum_{n=0}^{N-1}\sum_{x\in X}f(x)f(T^nx)$, where $f(x)$, $x\in X$, are nonnegative integers satisfying $\sum_{x\in X}f(x)=am$.

First, let T be a cyclic permutation: $X=\{1,\ldots,m\}$, $Tx=(x\bmod m)+1$. Then the problem is equivalent to the following one: $b=am$ archers are positioned at the points $1,\ldots,m$, $f(x)$ archers at a point x. An archer located at x threatens the archers located in the interval $[x,x+N-1]$, totally $\sum_{n=0}^{N-1}f(x+n)$ archers. We have to estimate

$$S=\sum_{n=0}^{N-1}\sum_{x=1}^{m}f(x)f(x+n)=\sum_{x=1}^{m}f(x)\sum_{n=0}^{N-1}f(x+n)=\sum_{x=1}^{m}\sum_{r=1}^{f(x)}\Bigl(\sum_{n=0}^{N-1}f(x+n)\Bigr)$$
$$=\sum_{R\text{ is an archer}}(\text{the number of archers threatened by }R).$$

Having replaced $X=\{1,\ldots,m\}$ by $\{1,\ldots,Nm\}$ and extended f to $\{1,\ldots,Nm\}$ by $f(x)=f(x-m)$ for $x>m$, we may assume that m is divisible by N. Let us subdivide $\{1,\ldots,m\}$ into $\frac{m}{N}$ intervals of length N. Let b_i, $i=1,\ldots,\frac{m}{N}$, be the number of archers located in the i-th interval. Fix i and enumerate the archers of the i-th interval in succession from the left to the right. Then the first archer threatens all b_i archers in the interval, the second archer threatens at least b_i-1 archers, etc. The total number of threats coming from the archers located in the i-th interval (to the archers in the same interval) is $\geq\frac{b_i(b_i+1)}{2}\geq\frac{b_i^2}{2}$. Hence, the total number of threats S satisfies

$$S\geq\sum_{i=1}^{m/N}\frac{b_i^2}{2}\geq\frac{N}{2m}\Bigl(\sum_{i=1}^{m/N}b_i\Bigr)^2=\frac{Nb^2}{2m}.$$

We therefore have $\frac{1}{N}\sum_{n=0}^{N-1}\sum_{x\in X}f(x)f(T^nx)\geq\frac{1}{N}S\geq\frac{b^2}{2m}$.

Now let T be an arbitrary permutation of an m-element set X. Let $X=X_1\cup\ldots\cup X_l$ be the partition of X into the union of disjoint cycles of T, let $m_j=|X_j|$ and $b_j=\sum_{x\in X_j}f(x)$, $j=1,\ldots,l$. We have

$$\frac{1}{N}\sum_{n=0}^{N-1}\sum_{x\in X}f(x)f(T^nx)=\frac{1}{N}\sum_{j=1}^{l}\sum_{n=0}^{N-1}\sum_{x\in X_j}f(x)f(T^nx)\geq\sum_{j=1}^{l}\frac{b_j^2}{2m_j}.\qquad(1.4)$$

Under the conditions $m_1 + \ldots + m_l = m$ and $b_1 + \ldots + b_l = ma$, the minimal value of the right hand side of (1.4) is reached when $\frac{b_1}{m_1} = \ldots = \frac{b_l}{m_l} = a$ and equals $\frac{1}{2}ma^2$. Hence,

$$\frac{1}{N}\sum_{n=0}^{N-1}\int fT^n f\,d\mu = \frac{1}{Nm}\sum_{n=0}^{N-1}\sum_{x\in X} f(x)f(T^n x) \geq \frac{a^2}{2}.$$

To prove part (b) of the theorem, take f to be $\frac{a}{c}1_A$, where A is a set of measure $c > 0$ in X. By Theorem 1.2, for appropriately chosen X, A, T and N we have

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) < \sqrt{c^2+(1-c)^2} + c - 1 + \frac{\delta c^2}{2a^2},$$

and so,

$$\frac{1}{N}\sum_{n=0}^{N-1}\int fT^n f\,d\mu = \frac{1}{N}\sum_{n=0}^{N-1}\Big(\frac{a}{c}\Big)^2\mu(A\cap T^{-n}A) < \Big(\frac{a}{c}\Big)^2\big(\sqrt{c^2+(1-c)^2} + c - 1\big) + \frac{\delta}{2}.$$

Since $\lim\limits_{c\to 0}\frac{a^2}{c^2}\big(\sqrt{c^2+(1-c)^2}+c-1\big) = \frac{a^2}{2}$, we have $\frac{1}{N}\sum_{n=0}^{N-1}\int fT^n f\,d\mu < \frac{a^2}{2} + \delta$ when c is small enough. ■

2. Lower bounds for some non-conventional ergodic averages

2.1. Let $T_1, \ldots, T_k$ be pairwise commuting measure preserving transformations of a probability measure space $(X, \mathcal{B}, \mu)$ and let A be a set of positive measure in X. Let us consider the averages

$$\begin{aligned}\frac{1}{N_1\ldots N_k}&\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_k=0}^{N_k-1}\mu\Big(\bigcap_{S\subseteq\{1,\ldots,k\}}\big(\prod_{i\in S}T_i^{-n_i}A\big)\Big)\\ &= \frac{1}{N_1\ldots N_k}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_k=0}^{N_k-1}\mu\big(A\cap T_1^{-n_1}A\cap T_2^{-n_2}A\cap\ldots\cap T_1^{-n_1}\cdots T_k^{-n_k}A\big).\end{aligned} \tag{2.1}$$

The convergence of (2.1) as $N_1, \ldots, N_k \to \infty$ is known only in the case $T_1 = \ldots = T_k$ for $k = 2$ (due to V. Bergelson) and $k = 3$ (B. Host and B. Kra).

2.2. If $T_1, \ldots, T_k$ do not commute the limit of the averages (2.1) may not exist:

Example. Let a measure preserving transformation P of a probability measure space $(Y,\mathcal{D},\nu)$ and a set $B\in\mathcal{D}$ with $\nu(B)=a$, $a\neq 0,1$, be such that $\nu\bigl(B\cap P^{-n}(B)\bigr)=a^2$ for all $n>0$. Let $S\subseteq\mathbb{N}$ with $1\notin S$; define $P_n=P$ if $n\in S$ and $P_n=\mathrm{Id}_Y$ otherwise. Take $(X,\mathcal{B},\mu)=(Y,\mathcal{D},\nu)^{\mathbb{N}}$, $A=B\times Y\times Y\times\ldots$ and define $T_1,T_2\colon X\longrightarrow X$ by $T_1(y_1,y_2,\ldots)=(P_1y_1,P_2y_2,\ldots)$ and $T_2(y_1,y_2,\ldots)=(y_2,y_3,\ldots)$. Then for any $n_1,n_2\geq 1$ one has $\mu\bigl(A\cap T_1^{-n_1}A\cap T_2^{-n_2}A\cap T_1^{-n_1}T_2^{-n_2}A\bigr)=a^3$ if $n_2\in S$ and $=a^2$ if $n_2\notin S$. Therefore, if S is such that the density $d(S)=\lim\limits_{N\to\infty}\frac{1}{N}\bigl|S\cap[1,N]\bigr|$ is not defined, then $\lim\limits_{N_1,N_2\to\infty}\frac{1}{N_1N_2}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\mu\bigl(A\cap T_1^{-n_1}A\cap T_2^{-n_2}A\cap T_1^{-n_1}T_2^{-n_2}A\bigr)$ does not exist.

2.3. Nevertheless, a positive lower bound of the averages (2.1) exists even for noncommuting $T_1,\ldots,T_k$. Put $\varphi(a)=\sqrt{a^2+(1-a)^2}+a-1$, $\varphi_1=\varphi$ and $\varphi_k(a)=\varphi\bigl(\varphi_{k-1}(a)\bigr)$, $k=2,3,\ldots$.

Theorem. *Let $T_1,\ldots,T_k$ be measure preserving transformations of a probability measure space $(X,\mathcal{B},\mu)$ and let $A\in\mathcal{B}$, $\mu(A)=a$. Then for any $N_1,\ldots,N_k\in\mathbb{N}$*

$$\frac{1}{N_1\ldots N_k}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_k=0}^{N_k-1}\mu\Bigl(\bigcap_{S\subseteq\{1,\ldots,k\}}\bigl(\prod_{i\in S}T_i^{-n_i}A\bigr)\Bigr)\geq\varphi_k(a). \tag{2.2}$$

Proof. We use induction on k; the case $k=1$ is Theorem 1.2. For all $n_1,\ldots,n_{k-1}\in\mathbb{Z}_+$ define $A_{n_1,\ldots,n_{k-1}}=\bigcap\limits_{S\subseteq\{1,\ldots,k-1\}}\bigl(\prod_{i\in S}T_i^{-n_i}A\bigr)$ and $a_{n_1,\ldots,n_{k-1}}=\mu(A_{n_1,\ldots,n_{k-1}})$.

Fix $N_1,\ldots,N_k$. By induction hypothesis we have

$$\frac{1}{N_1\ldots N_{k-1}}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_{k-1}=0}^{N_{k-1}-1}a_{n_1,\ldots,n_{k-1}}\geq\varphi_{k-1}(a). \tag{2.3}$$

The left hand part of (2.2) equals

$$\frac{1}{N_1\ldots N_{k-1}}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_{k-1}=0}^{N_{k-1}-1}\Bigl(\frac{1}{N_k}\sum_{n_k=0}^{N_k-1}\mu(A_{n_1,\ldots,n_{k-1}}\cap T^{-n_k}A_{n_1,\ldots,n_{k-1}})\Bigr)$$

By Theorem 1.2, for any $n_1,\ldots,n_{k-1}$ one has $\frac{1}{N_k}\sum_{n_k=0}^{N_k-1}\mu(A_{n_1,\ldots,n_{k-1}}\cap T^{-n_k}A_{n_1,\ldots,n_{k-1}})\geq\varphi(a_{n_1,\ldots,n_{k-1}})$. Since φ is a convex increasing function on $[0,1]$, taking into account (2.3) we get

$$\frac{1}{N_1\ldots N_{k-1}}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_{k-1}=0}^{N_{k-1}-1}\varphi(a_{n_1,\ldots,n_{k-1}})\geq\varphi\Bigl(\frac{1}{N_1\ldots N_{k-1}}\sum_{n_1=0}^{N_1-1}\cdots\sum_{n_{k-1}=0}^{N_{k-1}-1}a_{n_1,\ldots,n_{k-1}}\Bigr)$$
$$\geq\varphi\bigl(\varphi_{k-1}(a)\bigr)=\varphi_k(a).$$

■

2.4. We now pass to the averages

$$\frac{1}{N_1 \dots N_k} \sum_{n_1=0}^{N_1-1} \dots \sum_{n_k=0}^{N_k-1} \mu(A \cap T_1^{-n_1} A \cap \dots \cap T_k^{-n_k} A) \tag{2.4}$$

Theorem. *Let $T_1, \dots, T_k$ be (not necessarily commuting) measure preserving transformations of a probability measure space $(X, \mathcal{B}, \mu)$. For any $A \in \mathcal{B}$, $\mu(A) = a$,*

$$\lim_{N_1,\dots,N_k \to \infty} \frac{1}{N_1 \dots N_k} \sum_{n_1=0}^{N_1-1} \dots \sum_{n_k=0}^{N_k-1} \mu(A \cap T_1^{-n_1} A \cap \dots \cap T_k^{-n_k} A)$$

exists and is not less than a^{k+1}.

Proof. We have

$$\begin{aligned} &\lim_{N_1,\dots,N_k \to \infty} \frac{1}{N_1 \dots N_k} \sum_{n_1=0}^{N_1-1} \dots \sum_{n_k=0}^{N_k-1} \mu(A \cap T_1^{-n_1} A \cap \dots \cap T_k^{-n_k} A) \\ &\quad = \lim_{N_1,\dots,N_k \to \infty} \frac{1}{N_1 \dots N_k} \sum_{n_1=0}^{N_1-1} \dots \sum_{n_k=0}^{N_k-1} \int_X 1_A \cdot T_1^{n_1}(1_A) \cdot \ldots \cdot T_k^{n_k}(1_A)\, d\mu \\ &\quad = \int_X 1_A \cdot \big(\lim_{N_1 \to \infty} \tfrac{1}{N_1} \textstyle\sum_{n=0}^{N_1-1} T_1^n(1_A) \big) \cdot \ldots \cdot \big(\lim_{N_k \to \infty} \tfrac{1}{N_k} \textstyle\sum_{n=0}^{N_k-1} T_k^n(1_A) \big)\, d\mu = \int_A f_1 \dots f_k \, d\mu, \end{aligned}$$

where $f_i = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_i^n(1_A)$, $i = 1, \dots, k$.

2.5. Lemma. *Let T be a measure preserving transformation of a probability measure space $(X, \mathcal{B}, \mu)$, let $A \in \mathcal{B}$, $\mu(A) > 0$, and let $f = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n(1_A)$. Then $0 \le f \le 1$, $f(x) \ne 0$ for almost all $x \in A$ and $\int_A \frac{d\mu}{f} \le 1$.*

Proof. Without loss of generality we may assume that $(X, \mathcal{B}, \mu)$ is a Lebesgue space. Let $\pi\colon X \longrightarrow Y$, $\mu = \int_Y \mu_y \, d\nu$ be the ergodic decomposition of μ and let $B = \big\{ y \in Y \bigm| \mu_y(A) > 0 \big\}$. For almost every $y \in Y$ we have $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n(1_A) = \mu_y(A)$ in $L^1(X, \mu_y)$ and so, $f|_{\pi^{-1}(y)} = \mu_y(A)$. Therefore,

$$\mu\big(\{x \in A \bigm| f(x) = 0\}\big) = \mu\big(A \setminus \pi^{-1}(B)\big) \le \int_{Y \setminus B} \mu_y(A)\, d\nu = 0$$

and

$$\int_A \frac{d\mu}{f} = \int_B \Big(\int_A \frac{d\mu_y}{f} \Big) d\nu = \int_B \frac{\mu_y(A)}{\mu_y(A)}\, d\nu = \nu(B) \le 1.$$

■

2.6. We have, therefore, to determine the minimum of $F = \int_A f_1 \dots f_k \, d\mu$ under the conditions $f_i|_A > 0$ and $\int_A \frac{d\mu}{f_i} = 1$, $i = 1, \dots, k$. We pass to a finite model: $A = \{1, \dots, m\}$ and $f_i(j) = x_{i,j} > 0$, $j = 1, \dots, m$, with $\sum_{j=1}^m \frac{1}{x_{i,j}} = 1$, $i = 1, \dots, k$. We have to minimize the function $F(x_{1,1}, \dots, x_{k,m}) = \sum_{j=1}^m x_{1,j} \dots x_{k,j}$. At a point of extremum of F it must be $\operatorname{grad} F \in \operatorname{Span}\big\{\operatorname{grad}\big(\sum_{j=1}^m \frac{1}{x_{1,j}}\big), \dots, \operatorname{grad}\big(\sum_{j=1}^m \frac{1}{x_{k,j}}\big)\big\}$, that is, for some $c_1, \dots, c_k \in \mathbb{R}$, $\frac{x_{1,j} \dots x_{k,j}}{x_{i,j}} = \frac{c_i}{x_{i,j}^2}$ for $i = 1, \dots, k$, $j = 1, \dots, m$. This implies $x_{i,1} = \dots = x_{i,m}$, $i = 1, \dots, k$, that is, $f_1, \dots, f_k$ are constant on A. Hence, the minimum of F is attained when $f_1|_A = \dots = f_k|_A = a$ and equals $a^k \mu(A) = a^{k+1}$. ■

2.7. The same proof works for the uniform version of Theorem 2.4:

Theorem. *For any measure preserving transformations $T_1, \dots, T_k$ of a probability measure space $(X, \mathcal{B}, \mu)$ and any $A \in \mathcal{B}$, $\mu(A) = a$,*

$$\lim_{N_1 - M_1, \dots, N_k - M_k \to \infty} \frac{1}{(N_1 - M_1) \dots (N_k - M_k)} \sum_{n_1 = M_1}^{N_1 - 1} \dots \sum_{n_k = M_k}^{N_k - 1} \mu(A \cap T_1^{-n_1} A \cap \dots \cap T_k^{-n_k} A)$$

exists and is not less than a^{k+1}.

2.8. A lower bound for the averages (2.4) (which is not exact, of course) can be taken from Theorem 2.3:

Corollary of Theorem 2.3. *Let $T_1, \dots, T_k$ be measure preserving transformations of a measure space $(X, \mathcal{B}, \mu)$ and let $A \in \mathcal{B}$, $\mu(A) = a$. Then for any $N_1, \dots, N_k \in \mathbb{N}$*

$$\frac{1}{N_1 \dots N_k} \sum_{n_1 = 0}^{N_1 - 1} \dots \sum_{n_k = 0}^{N_k - 1} \mu\big(A \cap T_1^{-n_1} A \cap \dots \cap T_k^{-n_k} A\big) \geq \varphi_k(a).$$

2.9. Acknowledgment. I thank V. Bergelson for stimulating discussions and for constructive criticism.

Bibliography

[BHMP] V. Bergelson, B. Host, R. McCutcheon and F. Parreau, Aspects of uniformity in recurrence, *Colloquim Mathematicum* **84/85,** part 2 (2000), 549-576.

[Kh] A. Y. Khintchine, Eine Verschärfung des Poincaréschen "Wiederkehrsatzes", *Comp. Math.* **1** (1934), 177-179.

[R] V. A. Rohlin, Exact endomorphisms of a Lebesgue space, *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 499-530.