# ON THE AVERAGES OF FACTORS OF AN INFINITE WORD ON A FINITE SET OF INTEGERS

**Tom C. Brown**
*Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada*
tbrown@sfu.ca

**Allen R. Freedman**
*Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada*
allenfreedman@gmail.com

*Received: 10/30/19, Accepted: 6/19/20, Published: 7/24/20*

**Abstract**

Given an infinite word $\omega = x_1x_2\cdots$ on a finite set of non-negative integers, and two adjacent *factors* or *blocks* of $\omega$,

$$A = x_{i+1}x_{i+2}\cdots x_{i+m} \text{ and } B = x_{i+m+1}x_{i+m+2}\cdots x_{i+m+n},$$

one can ask whether $A = B$, or whether $A$ is a permutation of $B$, or whether the *sum* of $A$, $x_{i+1} + x_{i+2} + \cdots + x_{i+m}$, equals the sum of $B$, or whether $A, B$ have the same sum *and* the same length. In this note, we are concerned with whether the *average* of $A$,

$$\frac{1}{m}(x_{i+1} + x_{i+2} + \cdots + x_{i+m}),$$

equals the average of $B$. (The word 5 4 1 0 1 3 5 1 2 4 2 2 0 3 5 on the alphabet $S = \{0, 1, 2, 3, 4, 5\}$ contains the three consecutive blocks 0 1 3 5 1, 2, 4 2 2 0, each block having average 2.) Let $\omega = x_1x_2\cdots$, an infinite word on some set of integers, be fixed, and color all the pairs $\{a < b\}$ of non-negative integers by setting

$$f_\omega(a, b) = \frac{1}{b-a}(x_{a+1} + x_{a+2} + \cdots + x_b).$$

Then, applying the canonical Ramsey's theorem to this coloring, we find that there are only two "canonical" colorings rather than the usual four, namely, just the "constant" and "$1-1$" colorings. We study this in detail, for various classes of words. We also give a new and self-contained proof that for every infinite word $\omega$ (on a finite set of integers), and every $k \in \mathbb{N}$, $\omega$ contains $k$ consecutive blocks all with the same average.

*– Dedicated to the memory of Ron Graham*

## 1. Introduction

Let $S \subset \mathbb{Z}$, and let $\omega = x_1x_2x_3\cdots,\ x_i \in S$. Thus $\omega$ is an infinite word on the alphabet $S$. Although a number of definitions and results make sense if $S$ is allowed to be infinite (in particular Definitions 1 and 2, and Theorems 1 and 2), in general we restrict ourselves to the case where $S$ is finite.

Here are some simple classes of such words $\omega$:

1. $\omega$ is *periodic* if there exist (finite) words $y$ and $u$ with $\omega = yuuu\cdots$.

2. $\omega$ is *abelian periodic* if there exist words $y, u, u_1, u_2, u_3, \ldots$ with $\omega = yu_1u_2u_3\cdots$, where each $u_i$ is a permutation of $u$.

3. $\omega$ is *sum periodic* if $\omega = yu_1u_2u_3\ldots$, where all the $u_i$ have the same sum and the same length.

4. $\omega$ is *bounded average periodic* if $\omega = yu_1u_2u_3\ldots$, and all the $u_i$ have the same average, and and all the lengths of the $u_i$ are bounded by some constant.

5. $\omega$ is *average periodic* if $\omega = yu_1u_2u_3\ldots$, and all the $u_i$ have the same average. (Here, the lengths of the $u_i$ are not necessarily bounded.)

6. $\omega$ has the *average property*, which means that for all $k$, $\omega$ has a factor $B_1B_2\cdots B_k$, where $B_1, B_2, \cdots, B_k$ all have the same average. (Here the lengths of $B_1, B_2, \cdots, B_k$ are not necessarily equal.)

It turns out that *every* word $\omega$ (on a finite set of integers) has the average property. This is the subject of Section 4.

It is fairly clear that $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5$ and that none of the reverse implications hold. A simple example showing that $5 \not\Rightarrow 4$ is the word $\omega = u_1u_2u_3\cdots$, where $u_n = 0^n1^n0^n$. Thus $\omega = 010001100000111000\cdots$.

The following definition and notation will be used throughout.

**Definition 1.** Let $\omega = x_1x_2\cdots$ be an infinite word on an alphabet consisting of a finite set of integers. For each such word $\omega$, we define a function $f_\omega$, whose domain is the set of all 2-element subsets of $\mathbb{N} \cup \{0\}$, which we denote in the usual way by $[\mathbb{N} \cup \{0\}]^2$. Given $a, b \in \mathbb{N} \cup \{0\}, a < b$, we write $f_\omega(a, b)$ instead of $f_\omega(\{a, b\})$, and we define, for $u = x_{a+1}x_{a+2}\cdots x_b,\ f_\omega(u) = f_\omega(a, b) =$ average of $\{x_{a+1}, x_{a+2}, \cdots, x_b\} = \frac{1}{b-a}(x_{a+1} + x_{a+2} + \cdots + a_b)$.

Given $\omega = x_1x_2x_3\cdots$, where each $x_i \in S,\ S$ a finite set of integers, we will often be concerned with whether or not

$$\lim_{j\to\infty} f_\omega(0, j) = \lim_{j\to\infty} \frac{1}{j}(x_1 + x_2 + x_3 + \cdots + x_j)$$

exists.

Let $\omega$ be a given infinite word, on a finite set of integers. In Section 2 we apply the canonical Ramsey's theorem to the coloring $f_\omega$ of $[\mathbb{N} \cup \{0\}]^2$, and we find that there is an infinite subset $I$ of $\mathbb{N} \cup \{0\}$ such that $f_\omega$ restricted to $[I]^2$ is either constant or $1-1$.

**Definition 2.** Let $\omega = x_1x_2x_3\cdots$, where each $x_i \in S$, $S$ a finite set of integers. If there is an infinite subset $I$ of $\mathbb{N} \cup \{0\}$ such that $f_\omega$ restricted to $[I]^2$ is constant, we say that $f_\omega$ has the *constant property.* If there is an infinite subset $I$ of $\mathbb{N} \cup \{0\}$ such that $f_\omega$ restricted to $[I]^2$ is $1-1$, we say that $f_\omega$ has the $1-1$ *property.*

We completely characterize $f_\omega$ (in terms of whether $f_\omega$ has the constant property or the $1-1$ property or both), whenever $\omega$ is not average periodic, or is average periodic but not bounded average periodic, or is bounded average periodic.

Somewhat surprisingly, it turns out that if $\omega$ is not average periodic, then the sequence $\{f_\omega(0, i)\}$ must converge as $i \to \infty$.

We summarize the results of Section 2 in Section 3.

Then, in Section 4, we show that *every* infinite word $\omega$ (on a finite set $S$ of integers) has the *average property* (#6 on the previous list). This fact was conjectured by the second author in the late 1970s. A proof did not appear until 2012 [1]. That proof relied on the existence of many collinear points in certain seqences of planar lattice points [7]. The present proof is self-contained and makes no reference to lattice points; however, the method was inspired by Peter L. Montgomery's proof [5] of the existence of many collinear points in certain sequences of lattice points. (Namely, any sequence $\{P_n\}$ where for all $n \geq 0, P_{n+1} - P_n \in \{(0, 1), (1, 0)\}$.)

Section 5 has a few remarks on collinear points in the plane.

In Section 6, we show that it would be enough, in Section 4, to prove the main result only for the case $S = \{0, 1\}$.

## 2. The Canonical Version of Ramsey's Theorem for the Coloring $f_\omega$

The following lemma is crucial.

**Lemma 1.** *Let $\omega = x_1x_2x_3\cdots, x_i \in \mathbb{Z}$, with $f_\omega(a, b)$ defined as in Definition 1. If $a, b, c \in \mathbb{N} \cup \{0\}, a < b < c$, and any two of $f_\omega(a, b), f_\omega(b, c), f_\omega(a, c)$ are equal, then all three are equal.*

*Proof.* The proof is a simple computation. □

**Theorem 1.** *Let $\omega = x_1x_2\cdots$ be an infinite word on an alphabet consisting of a finite set of integers. Then, referring to Definitions 1 and 2, $f_\omega$ has the constant property, or the $1-1$ property, or both.*

*Proof.* Applying the ordinary canonical Ramsey's theorem (see, for example, [4], Section 5) to this coloring tells us that there exists an infinite subset $I$ of $\mathbb{N} \cup \{0\}$ such that $f_\omega$ restricted to $[I]^2$ is one of:

1. $1-1$ or
2. "min": $f_\omega(a,b) = f_\omega(c,d)$ iff $a = c$ (for all $a,b,c,d \in I, a < b, c < d$) or
3. "max": $f_\omega(a,b) = f_\omega(c,d)$ iff $b = d$ (for all $a,b,c,d \in I, a < b, c < d$) or
4. constant.

Now let $a,b,c \in I$, where $a < b < c$. If $f_\omega$ restricted to $[I]^2$ is the "min" coloring, then $f_\omega(a,b) = f_\omega(a,c)$. Then, by Lemma 1, $f_\omega(a,b) = f_\omega(b,c) = f_\omega(a,c)$. Since $f_\omega$ is the "min" coloring, $f_\omega(a,c) \neq f_\omega(b,c)$, a contradiction. Similarly, $f_\omega$ restricted to $[I]^2$ cannot be the "max" coloring. □

**Lemma 2.** *Given $\omega$, an infinite word on some finite set of integers, if*

$$\lim_{i\to\infty} f_\omega(0,i) = \alpha \in \mathbb{R},$$

*then, for all $m \in \mathbb{N}$, $\lim_{i\to\infty} f_\omega(m,i) = \alpha$.*

*Proof.* It's easy to see that for each fixed $m \in \mathbb{N}$, $\lim_{i\to\infty} |f_\omega(0,i) - f_\omega(m,i)| = 0$. □

**Theorem 2.** *If $\omega = x_1x_2x_3\cdots$, $x_i \in S \subset \mathbb{Z}$, $S$ finite, is not average periodic, then $f_\omega$ has the $1-1$ property, but does not have the constant property.*

*Proof.* (For the definition of "average periodic," see the beginning of the Introduction.) According to Theorem 1, $f_\omega$ has at least one of the two properties. But $f_\omega$ cannot have the constant property, since if $I = \{i_1 < i_2 < i_3 < \cdots\}$ and $f_\omega$ is constant on $[I]^2$, then in particular $f_\omega(i_1,i_2) = f_\omega(i_2,i_3) = \cdots$. Setting

$$y = x_1x_2\cdots x_{i_1}, u_1 = x_{i_1+1}x_{i_1+2}\cdots x_{i_2}, u_2 = x_{i_2+1}x_{i_2+2}\cdots x_{i_3}, \quad \cdots,$$

we have $\omega = yu_1u_2u_3\cdots$ and $u_1, u_2, u_3, \ldots$ all have the same average, i.e., $\omega$ is average periodic. □

**Remark 1.** If $\omega = x_1x_2x_3\cdots$, $x_i \in S$, $S$ finite, a simple calculation shows that

$$|f_\omega(0,n) - f_\omega(0,n+1)| \leq \frac{1}{n+1}(\max S - \min S).$$

**Remark 2.** For any such $\omega$, the sequence $\{f_\omega(0,i)\}_{i=1}^{\infty}$, being a bounded sequence, has one or more limit points.

**Theorem 3.** *Given $\omega = x_1x_2x_3\cdots$, $x_i \in S$, $S$ finite, assume that for some set $D = \{j_1 < j_2 < j_3 < \cdots\}$ and real number $\alpha$,*

$$f_\omega(0,j_i) \to \alpha \quad as \quad i \to \infty.$$

*Then the following two statements are equivalent.*

$(a)$ *There is an infinite set* $A \subset D$ *such that for all* $s, t \in A, s < t, f_\omega(s,t) \neq \alpha$.

$(b)$ *There is an infinite set* $B \subset D$ *such that* $f_\omega$ *is* $1-1$ *on* $[B]^2$.

*Proof.* First we show $(a) \Rightarrow (b)$. To simplify the notation, we might as well assume that $D = \mathbb{N}$, that is, we assume that $f_\omega(0,i) \to \alpha$ as $i \to \infty$ and, for all $s, t \in \mathbb{N}, s < t, f_\omega(s,t) \neq \alpha$. We now choose the set $B = \{k_1 < k_2 < k_3 < \cdots\}$ inductively as follows.

Set $k_1 = 1, k_2 = 2$. If $k_1 < k_2 < \cdots < k_n$ have been chosen so that $f_\omega$ is $1-1$ on $[\{k_1, k_2, \ldots, k_n\}]^2$, then choose $k_{n+1} > k_n$ so that for each $r,\ 1 \le r \le n$,

$$0 < |\alpha - f_\omega(k_r, k_{n+1})| < \min\{|\alpha - f_\omega(k_u, k_v)| : 1 \le u < v \le n\}.$$

(This is possible by Lemma 2: for each $r, 1 \le r \le n,\ f_\omega(k_r, i) \to \alpha$ as $i \to \infty$.)

Thus we automatically have that for all $1 \le u < v \le n, 1 \le r \le n$,

$$f_\omega(k_u, k_v) \neq f_\omega(k_r, k_{n+1}). \tag{1}$$

It remains to show that

$$1 \le r < s \le n \Rightarrow f_\omega(k_r, k_{n+1}) \neq f_\omega(k_s, k_{n+1}).$$

But if $1 \le r < s \le n$ (and hence $k_r < k_s < k_{n+1}$), the equality $f_\omega(k_r, k_{n+1}) = f_\omega(k_s, k_{n+1})$ implies by Lemma 1 that $f_\omega(k_r, k_s) = f_\omega(k_s, k_{n+1})$, contradicting (from the definition of $k_{n+1}$) the inequality (1) above.

To show that $(b) \Rightarrow (a)$, simply note that if $f_\omega$ is $1-1$ on $[B]^2$, then $f_\omega(s,t) = \alpha$ can hold for at most one pair $s, t$ in $B$. Hence $B - \{s, t\}$ can serve as $A$. □

**Corollary 1.** *Let* $A$ *be any infinite subset of* $\mathbb{N}$ *with asymptotic density 0, and let* $\omega$ *be the characteristic sequence of* $A$*. Then there is an infinite subset* $B$ *of* $\mathbb{N}$ *such that* $f_\omega$ *is* $1-1$ *on* $[B]^2$.

**Lemma 3.** *Given* $\omega = x_1x_2x_3\cdots,\ x_i \in S \subset \mathbb{Z},\ S$ *finite, such that* $\{f_\omega(0,i)\}$ *does not converge, then there are minumum and maximum limit points* $L_1 < L_2$*. Furthermore, any real number* $r,\ L_1 < r < L_2$*, is also a limit point of* $\{f_\omega(0,i)\}$.

*Proof.* Clearly the infimum $L_1$ and the supremum $L_2$ of the set of limit points are also limit points. Let $L_1 < r < L_2$. Let $\epsilon > 0$. For infinitely many $n$, we have $f_\omega(0,n) < r \le f_\omega(0,n+1)$. For large $n$, by Remark 1,

$$f_\omega(0,n+1) - f_\omega(0,n) < \epsilon.$$

Hence, $|r - f_\omega(0,n)| < \epsilon$ for infinitely many $n$. □

**Example.** Consider the binary word $\omega = 0^{1!}1^{2!}0^{3!}1^{4!}0^{5!}1^{6!}\cdots$. By induction,

$$2! + 4! + 6! + \cdots + (2n)! < 2 \cdot (2n)!,$$

thus

$$f_\omega(0, 1! + 2! + 3! + \cdots + (2n+1)!) = \frac{2! + 4! + 6! + \cdots + (2n)!}{1! + 2! + 3! + \cdots + (2n+1)!} < \frac{2}{2n+1} \to 0.$$

Similarly,

$$1! + 3! + 5! + \cdots + (2n-1)! < 2 \cdot (2n-1)!,$$

and thus

$$f_\omega(0, 1! + 2! + 3! + \cdots + (2n)!) = 1 - \frac{1! + 3! + 5! + \cdots + (2n-1)!}{1! + 2! + 3! + \cdots + (2n)!} > 1 - \frac{2}{2n} \to 1.$$

Thus every real number in $[0,1]$ is a limit point of $\{f_\omega(0,i)\}$.

**Theorem 4.** *Let* $\omega = x_1x_2x_3\cdots$ *(on a finite alphabet* $S \subset \mathbb{Z}$*) be such that* $\{f_\omega(0,i)\}$ *does not converge. Then* $f_\omega$ *is* $1-1$ *on* $[B]^2$ *for some infinite subset* $B$ *of* $\mathbb{N}$.

*Proof.* Using Lemma 3, choose an irrational number $\alpha$ and a sequence $D = \{j_1 < j_2 < j_3 < \cdots\}$ such that

$$f_\omega(0, j_i) \to \alpha \quad \text{as} \quad i \to \infty.$$

Since $f_\omega(r,s)$ is always rational, and hence unequal to $\alpha$, Theorem 3 gives us an infinite subset $B$ of $D$ such that $f_\omega$ is $1-1$ on $[B]^2$. $\square$

**Remark 3.** We will see in Theorem 6 that when $\{f_\omega(0,i)\}$ does not converge, then $f_\omega$ also has the constant property in a very strong sense.

Let $\omega = x_1x_2x_3\cdots$ be such that the sequence $\{f_\omega(0,i)\}$ converges. We now give a condition on such words $\omega$ which is equivalent to the statement that the coloring $f_\omega$ has the $1-1$ property.

**Definition 3.** Given $\omega$ (on a finite alphabet $S \subset \mathbb{Z}$) such that the sequence $f_\omega(0,i) \to \alpha$ as $i \to \infty$, the equivalence relation $\mathbf{E}$ on $\mathbb{N}$ is defined as follows.

$$a \cong b \text{ if and only if } a = b \;\; \textit{or} \;\; a \neq b \;\; \textit{and} \;\; f_\omega(\min\{a,b\}, \max\{a,b\}) = \alpha.$$

Note that the transitivity property of the relation $\cong$ follows from Lemma 1.

**Theorem 5.** *Let* $\omega = x_1x_2x_3\cdots$ *(on a finite alphabet* $S \subset \mathbb{Z}$*) and assume* $f_\omega(0,i) \to \alpha$ *as* $i \to \infty$*. Then the following two statements are equivalent.*

($a$) *The coloring,* $f_\omega$*, has the* $1-1$ *property.*

($b$) *The number of equivalence classes produced by* $\mathbf{E}$ *is infinite.*

*Proof.* First we show that $(b)$ implies $(a)$. Assume the number of equivalence classes of $\mathbf{E}$ is infinite. Let $A$ consist of one element from each equivalence class. Clearly $i, j \in A \Rightarrow f_\omega(i,j) \neq \alpha$, so by Theorem 3, $f_\omega$ is $1-1$ on $[B]^2$, where $B$ is some infinite subset of $A$.

Now we show that (not $b$) implies (not $a$). Let $A$ be any infinite subset of $\mathbb{N}$. Let $r, s, t \in A, r < s < t$, where $r \cong s \cong t$. Then $f_\omega(r,s) = \alpha = f_\omega(s,t)$, so $f_\omega$ is not $1-1$ on $[A]^2$. Thus $f_\omega$ does not have the $1-1$ property. □

**Corollary 2.** *If $\omega = yu_1u_2u_3\cdots$ (each $f_\omega(u_i) = \alpha$) is bounded average periodic, then $f_\omega$ has the constant property, but not the $1-1$ property.*

*Proof.* We must have $f_\omega(0,i) \to \alpha$ as $i \to \infty$. But the number of equivalence classes produced by $\mathbf{E}$ must be finite, for if $i$ and $j$ are in the same position in two equal $u_k$, then they will be in the same class. Hence an upper bound for the number of classes is $|y| + RT$, where $R$ is the number of different $u_k$ and $T$ is an upper bound for the lengths of the $u_k$. □

**Theorem 6.** *Given $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers), suppose $\{f_\omega(0,i)\}$ does not converge as $i \to \infty$. Then $\omega$ is average periodic. In fact, if $\alpha$ is* any *rational number with $L_1 < \alpha < L_2$ (where $L_1$ and $L_2$ are the minimum and maximum limit points of $\{f_\omega(0,i)\}$), then there is an infinite set $A \subset \mathbb{N}$ such that, if $i, j \in A, i < j$, then $f_\omega(i,j) = \alpha$.*

*Proof.* Recall that for $0 \leq i < j, f_\omega(i,j) = (x_{i+1} + x_{i+2} + x_{i+3} + \cdots + x_j)/(j-i)$. We will use the notation $g(n) = x_1 + x_2 + x_3 + \cdots + x_n = nf_\omega(0,n)$.

Let $\alpha = P/Q$. We know that there are infinitely many $n$ such that

$$f_\omega(0,n) \leq P/Q < f_\omega(0,n+1).$$

Using $nf_\omega(0,n) = g(n)$, the left and right hand inequalities become, respectively, $Qg(n) \leq Pn$ and $(n+1)P < Qg(n+1)$.

Now $g(n+1) = g(n) + x_{n+1}$, so the right hand side becomes $P + Pn < Qx_{n+1} + Qg(n)$. Since $Qg(n) \leq Pn$, we have

$$P + Pn < Qx_{n+1} + Qg(n) \leq Qx_{n+1} + Pn,$$

or, subtracting $P + Pn$,

$$0 < (Qx_{n+1} - P) + (Qg(n) - Pn) \leq Qx_{n+1} - P \leq Q(\max S) - P. \qquad (2)$$

The $x_{n+1}$ vary with $n$, but since $S$ is finite, there will be infinitely many $n$ such that (2) holds with each $x_{n+1}$ being the same (say each $x_{n+1} = a$). Among these values of $n$, since the integers $(Qa - P) + (Qg(n) - Pn)$ are bounded above and below, there is an infinite set $A$ on which they are all equal. For any two elements of $A$, say $i < j$, subtracting gives

$$0 = (Qa - P) + Qg(j) - Pj - ((Qa - P) + Qg(i) - Pi) = Q(g(j) - g(i)) - P(j - i),$$

or $P(j - i) = Q(g(j) - g(i))$, and finally $\alpha = P/Q = \frac{g(j)-g(i)}{j-i} =$

$$\frac{(x_1 + x_2 + \cdots + x_j) - (x_1 + x_2 + \cdots + x_i)}{j - i} = \frac{x_{i+1} + x_{i+2} + \cdots + x_j}{j - i} = f_\omega(i, j).$$

$\square$

## 3. Summary of Section 2

Here is a summary of our results so far. According to Theorem 2 and Theorem 6, if $\omega$ is not average periodic then $\{f_\omega(0, i)\}$ converges and $f_\omega$ has only the $1-1$ property. According to Corollary 2, if $\omega$ is bounded average periodic, then $\{f_\omega(0, i)\}$ converges and $f_\omega$ has only the constant property. Theorems 4, 5, 6 fill in the remaining parts of the diagram below.

$\omega$ NOT AVERAGE PERIODIC
ALL $f_\omega$ HAVE THE 1-1 PROPERTY AND
DO NOT HAVE THE CONSTANT PROPERTY
NOTE: ALL $\{f_\omega(0, n)\}$ ARE CONVERGENT HERE

$\omega$ AVERAGE PERIODIC
BUT NOT BOUNDED AVERAGE PERIODIC
ALL $f_\omega$ HAVE THE CONSTANT PROPERTY
PLUS THE FOLLOWING:

| | | |
|---|---|---|
| IF $\{f_\omega(0, n)\}$ IS NOT CONVERGENT, THEN $f_\omega$ HAS THE 1–1 PROPERTY NOTE: ALL $\omega$ WITH NON-CONVERGENT $\{f_\omega(0, n)\}$ ARE AVERAGE PERIODIC | IF $\{f_\omega(0, n)\}$ IS CONVERGENT AND E HAS INFINITELY MANY EQUIVALENCE CLASSES, THEN $f_\omega$ HAS THE 1–1 PROPERTY | IF $\{f_\omega(0, n)\}$ IS CONVERGENT AND E HAS ONLY FINITELY MANY EQUIVALENCE CLASSES, THEN $f_\omega$ DOES NOT HAVE THE 1–1 PROPERTY |

$\omega$ BOUNDED AVERAGE PERIODIC
ALL $f_\omega$ HAVE THE CONSTANT PROPERTY AND
DO NOT HAVE THE 1–1 PROPERTY
NOTE: ALL $\{f_\omega(0, n)\}$ ARE CONVERGENT HERE

## 4. The Average Property

We now show (Theorem 7 below) that every infinite word $\omega = x_1x_2x_3\cdots$, where each $x_i \in S, S$ a finite set of integers, has the average property. That is, $\omega$ has arbitrarily long sequences of consecutive blocks, all with the same average.

**Lemma 4.** *Given $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers), let $k \geq 2$, and suppose that $\omega$ does not contain $k$ consecutive blocks with equal averages. Let $g(n) = x_1 + x_2 + \cdots + x_n = nf_\omega(0, n)$ and $h(n) = Qg(n) - Pn$, where $P \geq 0$ and $Q > 0$ are integers. Let $c$ be any integer. Then $h(n) = c$ for at most $k$ positive integers $n$.*

*Proof.* Suppose $h(n_i) = c$ for $i = 1, 2, 3, \ldots, k+1$. Then $g(n_i) = (h(n_i)+Pn_i)/Q = (c + Pn_i)/Q$ and

$$f_\omega(n_i, n_{i+1}) = \frac{g(n_{i+1}) - g(n_i)}{n_{i+1} - n_i} = \frac{c + Pn_{i+1} - c - Pn_i}{Q(n_{i+1} - n_i)} = P/Q.$$

Thus $\omega$ has $k$ consecutive blocks with equal averages, a contradiction. □

**Lemma 5.** *Given $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers), let $k \geq 2$, and suppose that $\omega$ does not contain $k$ consecutive blocks with equal averages. Let $g(n) = x_1 + x_2 + \cdots + x_n = nf_\omega(0, n)$ and $h(n) = Qg(n) - Pn$, where $P \geq 0$ and $Q > 0$ are integers. Then, for any $m > 0$, there exists $n \in [1, (2m + 1)k + 1]$ such that*

$$|f_\omega(0, n) - P/Q| > \frac{1}{2kQ}.$$

*Proof.* For any $c \in [-m, m]$, by the previous Lemma, at most $k$ values of $n$ are such that $h(n) = c$. Hence there are at most $(2m + 1)k$ values of $n$ such that $h(n) \in [-m, m]$. Hence there must be $n \in [1, (2m+1)k+1]$ such that $h(n) \notin [-m, m]$ and for this $n$ we must have $|h(n)| > m$. Hence, $|f_\omega(0, n)-P/Q| = |g(n)/n-P/Q|=$

$$|\frac{Qg(n) - Pn}{Qn}| = |\frac{h(n)}{Qn}| \geq \frac{m + 1}{Q((2m + 1)k + 1)} > \frac{1}{2Qk}.$$

□

**Lemma 6.** *Given $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers), let $k \geq 2$, and suppose that $\omega$ does not contain $k$ consecutive blocks with equal averages. Let $P \geq 0$ and $Q > 0$ be integers. Then there exists an ascending sequence $n_1 < n_2 < n_3 < \cdots$ such that for each $i$,*

$$|f_\omega(0, n_i) - P/Q| > 1/2kQ. \tag{3}$$

*Proof.* Let $m > 0$. By the previous lemma, there is an $n_1 \in [1, (2m+1)k+1]$ such that $|h(n_1)| > m$ such that (3) holds with $i = 1$. Now let

$$m' = \max\{|h(t)| : t \in [1, (2m+1)k+1]\}.$$

Clearly $m' > m$ and there exists an $n_2 \in [1, (2m'+1)k+1]$ such that $|h(n_2)| > m'$ and (3) holds with $i = 2$. We must have $n_2 \in [1, (2m'+1)k+1] \backslash [1, (2m+1)k+1]$, so that

$$n_1 < n_2.$$

The argument can be repeated with $m'' = \max\{|h(t)| : t \in [1, (2m'+1)k+1\}$ to obtain $n_3 > n_2$, etc., with each $n_i$ satisfying (3). □

In Theorem 7 below, we will use the well known result from approximation theory that, if $L$ is any real and $M > 0$, then there exits a rational number $P/Q$ such that $Q > M$ and

$$|L - \frac{P}{Q}| < \frac{1}{2Q^2}. \tag{4}$$

**Theorem 7.** *Every infinite word $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers) contains $M$ consecutive equal average blocks for any $M > 0$.*

*Proof.* If $\{f_\omega(0, i)\}$ does not converge, then Theorem 6 shows that $\omega$ has, in fact, an infinite sequence of equal average consecutive blocks.

If $f_\omega(0, i) \to L$ as $i \to \infty$, assume $\omega$ does not have $k$ consecutive blocks of equal average. We choose $P \geq 0$ and $Q > k$ such that (4) holds. Lemma 6 gives us an infinite set of indices,

$$n_1 < n_2 < n_3 < \cdots$$

such that (3) holds for each $i$. Note that $|f_\omega(0, n_i) - P/Q| > 1/2kQ > 1/2Q^2 > |L - P/Q|$, so that

$$|f_\omega(0, n_i) - \frac{P}{Q}| - |L - \frac{P}{Q}| > \frac{1}{2kQ} - \frac{1}{2Q^2} = \epsilon > 0.$$

Hence, for each $i$,

$$|f_\omega(0, n_i) - L| = |(f_\omega(0, n_i) - \frac{P}{Q}) - (L - \frac{P}{Q})| \geq |f_\omega(0, n_i) - \frac{P}{Q}| - |L - \frac{P}{Q}| > \epsilon.$$

This implies that $\{f_\omega(0, i)\}$ has a limit point other than $L$, a contradiction. □

## 5. Remarks on Collinear Planar Lattice Points

Let $\omega = x_1x_2x_3\cdots$ (where each $x_i \in S, S$ a finite set of integers). Define a sequence of plane latttice points $\mathbf{P} = \{P_i\}_{i=0}^{\infty}$ by setting

$$P_0 = (0, 0), \quad P_{i+1} - P_i = (1, x_{i+1}), i \geq 0,$$

so that $P_n = (n, x_1 + \cdots + x_n)$. Now let $m < n < q$. Then $P_m, P_n, P_q$ are collinear exactly when the slope of the line through $P_m, P_n$ equals the slope of the line through $P_n, P_q$, that is, exactly when

$$\frac{1}{n-m}((x_1+\cdots+x_n)-(x_1+\cdots+x_m)) = \frac{1}{q-m}((x_1+\cdots+x_q)-(x_1+\cdots+x_n)),$$

or

$$\frac{1}{n-m}(x_{m+1}+\cdots+x_n) = \frac{1}{q-m}(x_{n+1}+\cdots+x_q),$$

or

$$f_\omega(m,n) = f_\omega(n,q).$$

Thus the sequence $\{P_i\}$ contains $M+1$ collinear points iff the word $\omega$ contains $M$ consecutive equal average blocks, for any $M > 0$.

A more general class of sequences $\{P_i\}$ of planar points is obtained by specifying a set $A$ of planar points and then requiring $P_0 = (0,0), P_{i+1} - P_i \in A, i \geq 0$. Such sequences are considered in [2], [3], [5], [6], [7].

## 6. S={0,1} Suffices

One version of van der Waerden's famous theorem on arithmetic progressions [8] is this: given any infinite word $\omega$ on a finite set of positive integers, and any $k$, there are $k$ consecutive blocks in $\omega$ all having the same *sum*.

To see that this statement is implied by van der Waerden's theorem, let $\omega = x_1x_2x_3\cdots$, where each $x_i \in S$, $S$ a finite subset of $\mathbb{N}$, and let $T = \{t_i\}_{i=1}^{\infty}$, where $t_i = x_1+x_2+\cdots+x_i$, $i \geq 1$. Since $t_{i+1}-t_i \leq \max S$, a finite number of translates of $T$ covers $\mathbb{N}$. Removing elements in overlapping translates, we obtain a finite coloring of $\mathbb{N}$, hence by van der Waerden's theorem, there are arbitrarily large monochromatic arithmetic progressions. Each monochromatic arithmetic progression is a subset of a translate of $T$, hence $T$ itself contains arbitrarily large arithmetic progressions. If, for example, $t_5, t_9, t_{18}$, are in arithmetic progression, then $(x_1+\cdots+x_9)-(x_1+\cdots+x_5) = (x_1+\cdots+x_{18})-(x_1+\cdots+x_9)$, or $x_6+x_7+x_8+x_9 = x_{10}+\cdots+x_{18}$.

The above statement seems similar in spirit to the equal average property, which says: given any infinite word $\omega$ on a finite set of positive integers, and any $k$, there are $k$ consecutive blocks in $\omega$ all having the same *average*.

The usual version of van der Waerden's theorem is that if $\mathbb{N}$ is finitely colored, there are arbitrarily large (finite) monochromatic arithmetic progressions. It is well known that it suffices to show this for just the case of two colors.

We now show (Theorem 9) that to prove that any infinite word on a finite set $S$ of integers has the average property, it suffices to show this just for the case $S = \{0,1\}$.

First we give a somewhat easier result.

**Theorem 8.** *Assume that every infinite word on* $\{0,1\}$ *has the average property. Let* $\{P\}_{i=0}^{\infty}$ *be any sequence of points in the plane such that* $P_i - P_{i-1} \in \{(0,1),(1,0)\}$, $i \geq 1$. *Then* $\{P\}_{i=0}^{\infty}$ *contains* $k$ *collinear points for every* $k$.

*Proof.* Let $P(0,0) = (0,0)$ and, for $i \geq 1, P_i - P_{i-1} \in \{(0,1),(1,0)\}$. For $i \geq 1$, define

$$x_i = \begin{cases} 1, & \text{if } P_i - P_{i-1} = (0,1) \\ 0, & \text{if } P_i - P_{i-1} = (1,0) \end{cases}$$

and let $\omega = x_1 x_2 x_3 \cdots$. Let $i < j < k$, where

$$x_{i+1}x_{i+2}\cdots x_j, \quad x_{j+1}x_{j+2}\cdots x_k$$

have the same average. In the first block, $x_{i+1} + x_{i+2} + \cdots + x_j$ is the number of ones in the block, that is, the number of vertical (unit) steps made from $P_i$ to $P_j$. Also, $(j-i) - (x_{i+1} + x_{i+2} + \cdots + x_j)$ is the number of zeros in $x_{i+1}x_{i+2}\cdots x_j$, that is, the number of horizontal (unit) steps made from $P_i$ to $P_j$. Hence

$$\frac{x_{i+1} + x_{i+2} + \cdots + x_j}{(j-i) - (x_{i+1} + x_{i+2} + \cdots + x_j)}$$

is the slope of the line connecting $P_i$ to $P_j$. Replacing $i, j$ by $j, k$, the same expression gives the slope of the line connecting $P_j$ to $P_k$.

But
$$\frac{x_{i+1} + x_{i+2} + \cdots + x_j}{j-i} = \frac{x_{j+1} + x_{j+2} + \cdots + x_k}{k-j}$$

(the two blocks have the same average), hence, since

$$\frac{a}{b} = \frac{c}{d} \quad \text{if and only if} \quad \frac{a}{b-a} = \frac{c}{d-c},$$

we have finally

$$\frac{x_{i+1} + x_{i+2} + \cdots + x_j}{(j-i) - (x_{i+1} + x_{i+2} + \cdots + x_j)} = \frac{x_{j+1} + x_{j+2} + \cdots + x_k}{(k-j) - (x_{j+1} + x_{j+2} + \cdots + x_k)}.$$

□

The above argument is reversible, so the converse of Theorem 8 is also true.

**Theorem 9.** *Assume that every infinite word on* $\{0,1\}$ *has the average property. Then every infinite word on any finite set* $S$ *of integers has the average property.*

*Proof.* Let

$$\omega = y_1 y_2 y_3 \cdots$$

be an infinite word on a finite set $S$ of non-negative integers. (The more general case, allowing $S$ to be a finite set of integers, follows by subtracting a suitable positive integer from each $y_i$, which does not disturb the average property.)

For each $i$, let $u_i = 1^{y_i}0$. For example, if $y_i = 3$, then $u_i = 1110$. If $y_i = 0$, we set $u_i = 0$.

Let

$$\beta = u_1u_2u_3\cdots \;=\; x_1x_2x_3\cdots\,,\; x_i \in \{0,1\},\; i \geq 1.$$

By assumption, for any $M > 0$, $\beta$ has $M$ consecutive blocks

$$x_{t_1+1}x_{t_1+2}\cdots x_{t_2}, x_{t_2+1}x_{t_2+2}\cdots x_{t_3}, \quad \cdots, \quad x_{t_M+1}x_{t_M+2}\cdots x_{t_{M+1}}$$

each with the same average, and we wish to show that $\omega$ has $K$ consecutive blocks with equal averages, for any $K$.

Let $K$ be given. Since each $x_{t_i}$ occurs in some $u_j$, and there are only finitely many distinct $u_j$, if we choose $M$ large enough, we can find $K+1$ indices amongst the $t_i$, say

$$j_1 < j_2 < j_3 < \cdots < j_{K+1},$$

such that each $x_{j_i}$ occurs in the *same position* of the *same block* $u$ ($u \in \{1^a0|\ a \in S\}$). By Lemma 1, the $K$ blocks

$$x_{j_1+1}x_{j_1+2}\cdots x_{j_2}, x_{j_2+1}x_{j_2+2}\cdots x_{j_3}, \quad \cdots, \quad x_{j_K+1}x_{j_K+2}\cdots x_{j_{K+1}},$$

also have the same average.

From these $K$ blocks, let us consider two that are adjacent. To ease the notation, let $f = j_i, g = j_{i+1}$, and $h = j_{i+2}$. We must have, for the two consecutive blocks in $\beta$ determined by $f, g$, and $h$, two corresponding consecutive blocks in $\omega$. Precisely, we have

$$x_{f+1}x_{f+2}\cdots x_g = ru_{m+1}u_{m+2}\cdots u_{n-1}s \text{ (corresponding to } y_{m+1}y_{m+2}\cdots y_n)$$

and

$$x_{g+1}x_{g+2}\cdots x_h = ru_{n+1}u_{n+2}\cdots u_{p-1}s \text{ (corresponding to } y_{n+1}y_{n+2}\cdots y_p)$$

for appropriate $m, n$, and $p$, and words $r$ and $s$, where $sr = u = u_m = u_n = u_p$ ($r$ will be the empty word, in the case $x_f = x_g = x_h = 0$.)

Summing these blocks, we get (using $\sum r + \sum s = \sum u_n$)

$$x_{f+1} + x_{f+2} + \cdots + x_g = \sum r + \sum u_{m+1} + \sum u_{m+2} + \cdots + \sum u_{n-1} + \sum s$$

$$= \sum u_{m+1} + \sum u_{m+2} + \cdots + \sum u_{n-1} + \sum u_n = y_{m+1} + y_{m+2} + \cdots + y_n,$$

and, similarly,

$$x_{g+1} + x_{g+2} + \cdots + x_h = \sum r + \sum u_{n+1} + \sum u_{n+2} + \cdots + \sum u_{p-1} + \sum s$$

$$= \sum u_{n+1} + \sum u_{n+2} + \cdots + \sum u_{p-1} + \sum u_p = y_{n+1} + y_{n+2} + \cdots + y_p.$$

Now, in the blocks $x_{f+1}x_{f+2}\cdots x_g$ and $x_{g+1}x_{g+2}\cdots x_h$, the numbers of 1s are

$$x_{f+1}+x_{f+2}+\cdots+x_g \text{ and } x_{g+1}+x_{g+2}+\cdots+x_h$$

respectively and the numbers of 0s are $n-m$ and $p-n$, respectively, since each $u_i$, as well as $sr$, contains exactly one 0. Hence,

$$g-f=(n-m)+x_{f+1}+x_{f+2}+\cdots+x_g$$

and

$$h-g=(p-n)+x_{g+1}+x_{g+2}+\cdots+x_h.$$

The blocks $x_{f+1}x_{f+2}\cdots x_g$ and $x_{g+1}x_{g+2}\cdots x_h$ have the same average, so

$$\frac{x_{f+1}+x_{f+2}+x_{f+3}+\cdots+x_g}{(n-m)+x_{f+1}+x_{f+2}+x_{f+3}+\cdots+x_g}=\frac{x_{g+1}+x_{g+2}+x_{g+3}+\cdots+x_h}{(p-n)+x_{g+1}+x_{g+2}+x_{g+3}+\cdots+x_h}.$$

Since $a/b=c/d$ if and only if $a/(b+a)=c/(d+c)$, we get

$$\frac{x_{f+1}+x_{f+2}+x_{f+3}+\cdots+x_g}{(n-m)}=\frac{x_{g+1}+x_{g+2}+x_{g+3}+\cdots+x_h}{(p-n)};$$

that is,

$$\frac{y_{m+1}+y_{m+2}+y_{m+3}+\cdots+y_n}{(n-m)}=\frac{y_{n+1}+y_{n+2}+y_{n+3}+\cdots+y_p}{(p-n)}.$$

$\square$

## References

[1] T. C. Brown, Approximations of additive squares in infinite words, *Integers, Electron. J. Combin. Number Theory* **12** (2012), A22.

[2] J. L. Gerver, Long walks in the plane with few collinear points, *Pacific J. Math.* **83** (1979), 349-355.

[3] J. L. Gerver and L. T. Ramsey, On certain sequences of lattice points, *Pacific J. Math.* **83** (1979), 357-363.

[4] R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, 2nd ed., John Wiley & Sons, New York, 1980.

[5] P. L. Montgomery, Solution of problem 5811 (proposed by T. C. Brown), *Amer. Math. Monthly* **79** (1972), 1143-1144.

[6] C. Pomerance, Collinear subsets of lattice point sequences - an analog of Szemerédi's theorem, *J. Combin. Theory Ser. A* **28** (1980), 140-149.

[7] L. T. Ramsey, Fourier-Stieltjes transforms of measures with a certain continuity property, *J. Funct. Anal.* **25** (1977), no. 3, 306-316.

[8] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Archief voor Wiskunde* **15** (1927), 212-216.