Fractional Chromatic Number of Distance Graphs Generated by Two-Interval Sets

Daphne Der-Fen Liu *
Department of Mathematics
California State University, Los Angeles
Los Angeles, CA 90032, USA
Email: dliu@calstatela.edu

Xuding Zhu †
Department of Applied Mathematics
National Sun Yat-sen University
Kaohsiung, Taiwan 80424
and
National Center for Theoretical Sciences
Email: zhu@math.nsysu.edu.tw

September 30, 2007

Abstract

Let \( D \) be a set of positive integers. The distance graph generated by \( D \), denoted by \( G(Z, D) \), has the set \( Z \) of all integers as the vertex

---

*Supported in part by the National Science Foundation under grant DMS 0302456.
†Supported in part by the National Science Council, R. O. C., under grant NSC94-2115-M-110-001.
set, and two vertices \( x \) and \( y \) are adjacent whenever \( |x - y| \in D \). For integers \( 1 < a \leq b < m - 1 \), denote \( D_{a,b,m} = \{1, 2, \ldots, a - 1\} \cup \{b + 1, b + 2, \ldots, m - 1\} \). For the special case \( a = b \), the chromatic number for the family of distance graphs \( G(Z, D_{a,a,m}) \) was first studied by Eggleton, Erdős and Skilton [5] and was completely solved by Chang, Liu and Zhu [3]. For the general case \( a \leq b \), the fractional chromatic number for \( G(Z, D_{a,b,m}) \) was studied by Lam and Lin [14] and by Wu and Lin [23], in which partial results for special values of \( a, b, m \) were obtained. In this article, we completely settle this problem for all possible values of \( a, b, m \).

2000 Mathematics Subject Classification: Primary 05C15 and 11B05.

Keywords: distance graphs, fractional chromatic number, density of integral sequences, \( T \)-coloring.

1 Introduction

Let \( D \) be a set of positive integers. The distance graph generated by \( D \), denoted by \( G(Z, D) \), has the set \( Z \) of all integers as the vertex set, and two vertices \( x \) and \( y \) are adjacent whenever \( |x - y| \in D \). Initiated by Eggleton, Erdős and Skilton [5], the study of distance graphs has attracted considerable attention ([2–8, 11-18, 20-25]).

A fractional coloring of a graph \( G \) is a mapping \( f \) which assigns to each independent set \( I \) of \( G \) a non-negative weight \( f(I) \) such that for each vertex \( x \), \( \sum_{x \in I} f(I) \geq 1 \). The fractional chromatic number \( \chi_f(G) \) of \( G \) is the least total weight of a fractional coloring for \( G \).

The problem of determining the fractional chromatic number for distance graphs has been studied in different research areas under different names. Firstly, it is equivalent to a sequence density problem in number theory. For a set \( D \) of positive integers, a sequence \( S \) of non-negative integers is called a \( D \)-sequence if \( a - b \notin D \) for any \( a, b \in S \). Let \( S(n) \) denote \( \{0, 1, \ldots, n-1\} \cap S \). The upper density and the lower density of \( S \) are defined, respectively, by

\[
\overline{\delta}(S) = \lim_{n \to \infty} \frac{S(n)}{n}, \quad \underline{\delta}(S) = \lim_{n \to \infty} \frac{S(n)}{n}.
\]
We say $S$ has density $\delta(S)$ if $\delta(S) = \bar{\delta}(S) = \bar{\delta}(S)$. The parameter of interest is the maximum density of a $D$-sequence, defined by

$$\mu(D) = \sup \{\delta(S) : S \text{ is a } D\text{-sequence}\}.$$ 

The problem of determining or estimating $\mu(D)$ was initially posed by Motzkin in an unpublished problem collection (cf. [1]), and has been studied in [1, 10, 19, 9, 18]. Note that, $S$ is a $D$-sequence if and only if $S$ (as a set of integers) is an independent set of $G(Z, D)$. It was proved by Chang et al. [3] that for any finite set $D$,

$$\mu(D) = \frac{1}{\chi_f(G(Z, D))}.$$ 

Secondly, the fractional chromatic number of a distance graph is equivalent to an asymptotic problem in $T$-coloring. The $T$-coloring problem was motivated by the channel assignment problem introduced by Hale [10], in which an integer broadcast channel is assigned to each of a given set of stations or transmitters so that interference among nearby stations is avoided. Interference is modeled by a set of non-negative integers $T$ containing 0 as the forbidden channel separations. By using a graph $G$ to represent the broadcast network, a valid channel assignment is defined as a $T$-coloring for $G$, which is a mapping $f : V(G) \to Z$ such that $|f(x) - f(y)| \not\in T$ whenever $xy \in E$. The span of a $T$-coloring $f$ is the difference between the largest and the smallest numbers in $f(V)$, i.e., $\max \{|f(u) - f(v)| : u, v \in V\}$. Given $T$ and $G$, the $T$-span of $G$, denoted by $\text{sp}_T(G)$, is the minimum span among all $T$-colorings of $G$. As for any graph $G$, $\text{sp}_T(G) \leq \text{sp}_T(K_{\chi(G)})$, it is useful to estimate $\text{sp}_T(K_n)$. Let $\sigma_n$ denote $\text{sp}_T(K_n)$. Griggs and Liu [9] proved that for any set $T$ the asymptotic $T$-coloring ratio

$$R(T) := \lim_{n \to \infty} \frac{\sigma_n}{n}$$

exists and is a rational number. It was proved in [3] that for any $T$, by letting $D = T - \{0\}$, we have $R(T) = \chi_f(G(Z, D))$.

Partially due to its rich connections to other problems, the fractional chromatic number for various classes of distance graphs has been studied in the literature (cf. [2, 3, 17, 18, 23, 14, 24, 25]). If $D$ is a singleton, trivially
\(\chi_f(G(Z, D)) = 2.\) If \(D = \{a, b\}\) and \(\gcd(a, b) = 1,\) it is known [1] that \(\chi_f(G(Z, D)) = \frac{a+b}{\lfloor (a+b)/2 \rfloor} .\) For \(|D| \geq 3,\) the exact values of \(\chi_f(G(Z, D))\) are known only for some special sets \(D.\) For \(D = \{a, b, a+b\},\) upper and lower bounds for \(\chi_f(G(Z, D))\) were obtained by Rabinowitz and Proulx [19]. Let \(\chi(G)\) and \(\omega(G)\) denote, respectively, the chromatic number and the clique number of \(G.\) It is easy to see that \(\omega(G) \leq \chi_f(G) \leq \chi(G)\) holds for any graph \(G,\) and \(\chi(G(Z, D)) \leq |D| + 1 \) ([4, 20]) if \(D\) is finite. In [18], the sets \(D\) with \(\omega(G(Z, D)) \geq |D|\) were characterized and the value of \(\chi_f(G(Z, D))\) for most of this class of graphs, including \(D = \{a, b, a+b\},\) was determined.

For any two integers \(a \leq b,\) let \([a, b]\) denote the interval of consecutive integers \(\{a, a+1, \ldots, b\}.\) It is known that if \(D = [a, b],\) then \(\chi_f(G(Z, D)) = (a+b)/a\) [9, 2]. For the sets \(D\) of the form \(D = [1, m] - \{k, 2k, \ldots, sk\}\) for integers \(m, k\) and \(s,\) the values of \(\chi_f(G(Z, D))\) were determined in [17].

For \(1 < a \leq b < m - 1,\) let \(D_{a,b,m}\) denote the two-interval set

\[D_{a,b,m} = [1, a - 1] \cup [b + 1, m - 1].\]

Note, if \(a = b,\) then \(D_{a,a,m} = [1, m - 1] - \{a\}.\) The chromatic number for \(G(Z, D_{a,a,m})\) was first studied by Eggleton, Erdős and Skilton [5] and the problem was completely solved in [3]. For the general case \(a \leq b,\) both the fractional chromatic number and the chromatic number for \(G(Z, D_{a,b,m})\) were studied by Wu and Lin [23], and by Lam and Lin [14]. Some partial results were obtained. In this article, we completely determine the fractional chromatic number of \(G(Z, D_{a,b,m})\) for all \(1 < a \leq b < m - 1.\)

2 The main result and some preliminaries

For some special cases, the values of \(\chi_f(G(Z, D_{a,b,m}))\) for the two-interval set \(D_{a,b,m} = [1, a - 1] \cup [b + 1, m - 1]\) were solved in [23] and [14]. If \(b < 2a,\) then \(\chi_f(G(Z, D_{a,b,m}))\) is determined in [23]. Let \(\Delta = m - b.\) If \(\Delta \leq a\) or \(\Delta \geq 2a,\) then \(\chi_f(G(Z, D_{a,b,m}))\) is determined in [14]. Some other special cases (which cannot be easily described) are discussed in [14].

The main result of this article is the following which completely determines the value of \(\chi_f(G(Z, D_{a,b,m}))\) for all \(1 < a \leq b < m - 1.\)
Theorem 1 For integers $1 < a \leq b < m - 1$. Suppose $G = G(Z, D_{a,b,m})$ where $D_{a,b,m} = [1, a - 1] \cup [b + 1, m - 1]$. Let $\Delta = m - b$, $s = \lfloor b/a \rfloor$, and $q = \lfloor m/\Delta \rfloor$.

• If $\Delta \geq 2a$, then $\chi_f(G) = (sa + m)/(s + 1)$.

• If $\Delta \leq a$, then $\chi_f(G) = \max\{a, m/(s + 1)\}$.

• If $a < \Delta < 2a$, then

\[
\chi_f(G) = \begin{cases}
\frac{sa+m}{s+1}, & \text{if } 2qa \leq m < a + q\Delta \text{ or } m \geq (2q+1)a; \\
\frac{m}{q}, & \text{if } m < \min\{q\Delta + a, 2qa\}; \\
\frac{(2q-1)m+a}{2q^2}, & \text{if } q\Delta + a \leq m < (2q+1)a.
\end{cases}
\]

The cases for $\Delta \geq 2a$ and $\Delta \leq a$ were solved in [14]. However, for completeness, we include these cases in the statement and give a short proof for them.

Recall the result in [3] mentioned in Section 1, the fractional chromatic number of $G$ is equal to the reciprocal of $\mu(D_{a,b,m})$, which is the maximum density of a $D_{a,b,m}$-sequence. Let $I = \{x_1, x_2, \cdots\}$ be a $D_{a,b,m}$-sequence where $x_i < x_{i+1}$. Let $\delta_i = x_{i+1} - x_i$. The sequence $\Omega = (\delta_1, \delta_2, \cdots)$ is called the gap sequence of $I$. In the following, we call a sequence $(\delta_1, \delta_2, \cdots)$ a $D$-gap sequence if it is the gap sequence of a $D$-sequence. Observe that a sequence $(\delta_1, \delta_2, \cdots)$ is a $D$-gap sequence if and only if for any $j \leq j'$, $\sum_{i=j}^{j'} \delta_i \notin D$. In particular, the following observation is frequently used, usually implicitly, in our proofs.

• A sequence $(\delta_1, \delta_2, \cdots)$ is a $D_{a,b,m}$-gap sequence if and only if

1. $\delta_i \geq a$ for each $i$; and
2. for any $j \leq j'$, either $\sum_{i=j}^{j'} \delta_i \leq b$ or $\sum_{i=j}^{j'} \delta_i \geq m$.

By definition,

$$
\mu(D_{a,b,m}) = \max \lim_{n \to \infty} \frac{|I \cap [0, n - 1]|}{n},
$$

5
where the maximum is taken over all \(D_{a,b,m}\)-sequences \(I\). Hence

\[
\chi_f(G) = \frac{1}{\mu(D_{a,b,m})} = \min \lim_{n \to \infty} \frac{n}{|I \cap [0, n-1]|} = \min \lim_{k \to \infty} \sum_{i=1}^{k} \frac{\delta_i}{k}.
\]

Again, the minimum is taken over all \(D_{a,b,m}\)-sequences \(I\) with gap sequence \((\delta_1, \delta_2, \cdots)\).

For an interval of integers \([a, b]\), we call its cardinality \(|[a, b]|\) the length of \([a, b]\). Given a \(D_{a,b,m}\)-gap sequence \(Y = (\delta_1, \delta_2, \delta_3, \cdots)\), the average gap length of \(Y\) is

\[
\langle \delta \rangle = \lim_{k \to \infty} \frac{\sum_{i=1}^{k} \delta_i}{k} \quad \text{if exists}.
\]

Thus to determine the fractional chromatic number of \(G(Z, D_{a,b,m})\), it amounts to determine the minimum average gap length of a \(D_{a,b,m}\)-gap sequence. Usually, the gap sequences we concern are periodic. For a periodic gap sequence, it suffices to present one period of the sequence. We shall denote by \(\langle y_1, y_2, \cdots, y_k \rangle\) the infinite periodic sequence with period \(k\). That is, \(\langle y_1, y_2, \cdots, y_k \rangle = \langle y_1, y_2, \cdots, y_j, \cdots \rangle\) where for \(j > k\), \(y_j = y_{j-k}\). For convenience, we denote by \(p \otimes t\), for any integers \(p\) and \(t\), the \(p\) repetitions of \(t\). For example, \(\langle 3 \otimes 5, 2 \otimes 7 \rangle\) is the periodic sequence \(\langle 5, 5, 5, 7, 7, 5, 5, 5, 7, 7, \cdots \rangle\).

We now give a short proof for the cases \(\Delta \leq a\) and \(\Delta \geq 2a\). As each gap of a \(D_{a,b,m}\)-gap sequence is at least \(a\), we have \(\chi_f(G) \geq a\). If \(m \leq (s+1)a\), then \(\langle a \rangle\) is a \(D_{a,b,m}\)-gap sequence with average gap length \(a\). Hence \(\chi_f(G) = a\).

Assume \(m > (s+1)a\) and \(\Delta \leq a\). Then the sequence \(\langle s \otimes a, m - sa \rangle\) is a \(D_{a,b,m}\)-gap sequence of average gap length \(m/(s+1)\). So \(\chi_f(G) \leq m/(s+1)\).

On the other hand, for any \(D_{a,b,m}\)-gap sequence \((\delta_1, \delta_2, \cdots)\), since \(\sum_{i=1}^{s+1} \delta_i \geq (s+1)a \geq b+1\), we must have \(\sum_{i=1}^{s+1} \delta_i \geq m\). Hence the average gap length is at least \(m/(s+1)\). So \(\chi_f(G) = m/(s+1)\).

Assume \(\Delta \geq 2a\). It is easy to verify that the sequence \(\langle s \otimes a, m \rangle\) is a \(D_{a,b,m}\)-gap sequence with average gap length \((m+sa)/(s+1)\). Hence \(\chi_f(G) \leq (m+sa)/(s+1)\)

On the other hand, if \(\chi_f(G) = 1/\mu(D_{a,b,m}) < (m+sa)/(s+1)\), then there is a \(D_{a,b,m}\)-sequence \(I\) with \(|[0, sa+m-1] \cap I| \geq s+2\). Without loss of generality, we may assume \(0 \in I\). Let \(I' = \{i : i \in I, i \leq b\} \cup \{i - m + a : i \in I, i \geq m - a\}\). It is easy to verify that \(|I| = |I'|, I' \subseteq [0, (s+1)a - 1]\) and for any \(x, y \in I', |x - y| \geq a\). This is in contrary to the assumption that \(|I| \geq s + 2\). Therefore we have \(\chi_f(G) = (m+sa)/(s+1)\).
3 Proof of the upper bound

In the rest of the paper, we assume that $a < \Delta < 2a$, and let

$$\tau(D_{a,b,m}) = \begin{cases} \frac{sa+m}{s+1}, & \text{if } 2qa \leq m < a + q\Delta \text{ or } m \geq (2q + 1)a; \\ \frac{m}{q}, & \text{if } m < \min\{q\Delta + a, 2qa\}; \\ \frac{(2q-1)m+a}{2q^2}, & \text{if } q\Delta + a \leq m < (2q + 1)a. \end{cases}$$

In this section, we prove that $\chi_f(G) \leq \tau(D_{a,b,m})$. This amounts to present a $D_{a,b,m}$-gap sequence whose average gap length is at most $\tau(D_{a,b,m})$.

Lemma 2 Suppose $G = G(Z, D_{a,b,m})$. Then $\chi_f(G) \leq \tau(D_{a,b,m})$.

Proof. First note that the following are two $D_{a,b,m}$-gap sequences:

$$\langle s \otimes a, m \rangle \text{ and } \langle (q-1) \otimes \Delta, m - ((q-1)\Delta) \rangle,$$

where the average gap lengths, respectively, are $(sa+m)/(s+1)$ and $m/q$. This proves the result for all the cases, except the very last one.

For the last case, $q\Delta + a \leq m < (2q + 1)a$, the gap sequence is more complicated. We shall define some special sequences, then combine them to form the required periodic sequence.

For $i = 1, 2, \ldots, q-1$, let $Y_i$ and $Y_i'$ and $Z$ be finite sequences of integers defined as follows:

$$Y_i = (i \otimes \Delta, a, (q-1-i) \otimes \Delta, m - (a + (q-1)\Delta))$$

$$Y_i' = ((i-1) \otimes \Delta, \Delta + a, (q-1-i) \otimes \Delta, m - (a + (q-1)\Delta))$$

$$Z = (a)$$

Let $$Y_q' = ((q-1) \otimes \Delta, m - (q-1)\Delta).$$

For finite sequences $A = (a_1, a_2, \ldots, a_s)$ and $B = (b_1, b_2, \ldots, b_t)$, the concatenation of $A$ and $B$, denoted by $AB$, is the sequence

$$AB = (a_1, a_2, \ldots, a_s, b_1, b_2, \ldots, b_t).$$
The concatenation of sequences is associative. Thus for finite sequences $A_1, A_2, \cdots, A_t$, the sequence $A_1A_2\cdots A_t$ is well-defined. Define the periodic gap sequence as

$$\langle Y'_qY'_{q-1}Y'_{q-2}Y'_{q-3}\cdots Y'_{1}Y'_1Z \rangle.$$ 

Now we show that this sequence is indeed a $D_{a,b,m}$-gap sequence. Since

$$m - (a + (q - 1)\Delta) = m - q\Delta - a + \Delta \geq \Delta > a,$$

each entry of the sequence is at least $a$. It remains to show that the sum of any number of consecutive entries of the sequence is either at most $b$ or at least $m$. Observe that the sum of the entries in each $Y_i$ or $Y'_i$ is equal to $m$. Consider the sum of any $t$ consecutive entries in the sequence. Straightforward calculation shows that if $t \geq q+1$, then the sum is at least $m$; if $t \leq q-1$, then the sum is at most $b$; if $t = q$, then the sum is either equal to $m$ or at most $b$. (Here we use the condition that $(q-1)\Delta + a \leq (q-1)\Delta + m - q\Delta = b$.) Thus the sequence defined above is a $D_{a,b,m}$-gap sequence.

Straightforward calculation shows that this gap sequence has average gap length $\frac{(2q-1)m+a}{2q^2}$.

4 Proof of the lower bound

To complete the proof of Theorem 1, it remains to show that $\chi_f(G) \geq \tau(D_{a,b,m})$. To this end, we need some more definitions.

In the following, we assume that $I = \{x_1, x_2, \cdots \}$ is a $D_{a,b,m}$-sequence, i.e., an independent set in $G = G(Z, D_{a,b,m})$. We shall prove that the gap sequence of $I$ has average gap length at least $\tau(D_{a,b,m})$.

Let

$$L = \{i : x_{i+1} - x_i \geq \Delta\}.$$ 

For each $x_i \in I$, we associate it with a set $X_i$ of integers as follows.

$$X_i = \begin{cases} 
[x_i, x_i + \Delta - 1], & \text{if } i \in L; \\
[x_i, x_i + a - 1] \cup [x_i + m, x_i + m + a - 1], & \text{if } i \not\in L.
\end{cases}$$

Lemma 3 If $i \neq j$, then $X_i \cap X_j = \emptyset$. 

8
Proof. Assume \( i < j \). If \( i \in L \), then \( X_i = [x_i, x_i + \Delta - 1] \) and by definition, \( x_j \geq x_i + \Delta \). As \( t \in X_j \) implies that \( t \geq x_j \), we have \( X_i \cap X_j = \emptyset \). Assume \( i \notin L \). Then \( X_i = [x_i, x_i + a - 1] \cup [x_i + m, x_i + m + a - 1] \). As \( x_j \geq x_i + a \), we know that \( X_j \cap [x_i, x_i + a - 1] = \emptyset \). Assume \( X_j \cap [x_i + m, x_i + m + a - 1] \neq \emptyset \).

Then by the definition of \( X_j \), we have either \( x_j \in [x_i + m - \Delta + 1, x_i + m - 1] \) or \( x_j \in [x_i + m, x_i + m + a - 1] \). The former case implies \( b + 1 \leq x_j - x_i \leq m - 1 \); and the latter case implies \( b + 1 \leq x_j - x_i + 1 \leq m - 1 \) (since \( i \notin L \), we have \( a \leq x_{i+1} - x_i < \Delta \)). For both cases, it contradicts the assumption that \( I \) is a \( D_{a,b,m} \)-sequence.

We call intervals of the form \([x_i + m, x_i + m + a - 1]\) for \( i \notin L \) Type-B \( I \)-intervals. Intervals of the form \([x_i, x_i + \Delta - 1]\) for \( i \in L \), and intervals of the form \([x_i, x_i + a - 1]\) for \( i \notin L \) are called Type-A \( I \)-intervals. Both Type-A and Type-B \( I \)-intervals are referred as \( I \)-intervals. The length of an \( I \)-interval is either \( \Delta \) or \( a \), and they are called, respectively, long or short \( I \)-intervals.

**Lemma 4** If \( T = [x_i, x_i + a - 1] \) is a short Type-A \( I \)-interval, then the first \( I \)-interval \( T' = [u,v] \) with \( u \geq x_i + a \) is Type-A.

**Proof.** Assume to the contrary that \( T' = [u,v] = [x_j + m, x_j + m + a - 1] \) for some \( j \). As \( x_j + m \geq x_i + a \), which implies \( x_i - x_j \leq m - a \), we have \( x_i - x_j \leq b \). So \( x_j + m \geq x_i + \Delta \). In addition, since \( T \) is a short Type-A \( I \)-interval, \( x_{i+1} < x_i + \Delta \). Hence, \( x_{i+1} < x_j + m \), contradicting the choice of \( T' \).

**Lemma 5** There are at most \( s \) short consecutive \( I \)-intervals that are of the same type.

**Proof.** First we show that there are at most \( s \) short consecutive Type-A \( I \)-intervals. Assume \( T_1 = [u_1, v_1], T_2 = [u_2, v_2], \ldots, T_{j} = [u_{j}, v_{j}] \) are consecutive \( I \)-intervals and \( T_1, T_2, \ldots, T_{j-1} \) are short and Type-A. By Lemma 4, \( T_{j} \) is also Type-A. So \( u_1, u_2, \ldots, u_{j} \in I \). We prove by induction on \( i \) that \( u_i \leq u_1 + b \) for \( i = 1, 2, \ldots, j \). It is trivial for \( i = 1 \). Assume \( i < j \) and \( u_i \leq u_1 + b \). By definition of \( I \)-intervals, \( u_{i+1} - u_i < \Delta \). Hence \( u_{i+1} < u_i + \Delta \leq u_1 + m \). As \( u_1, u_{i+1} \in I \), it follows that \( u_{i+1} \leq u_1 + b \).
Because $s = \lfloor b/a \rfloor$ and $|T_i| \geq a$, we conclude that there are at most $s$ consecutive short Type-A $I$-intervals. By definition, consecutive Type-B $I$-intervals correspond to consecutive short Type-A $I$-intervals. So the result follows.

Suppose $T$ is an $I$-interval. Define the weight of $T$ by

$$w(T) = \begin{cases} 1, & \text{if } T \text{ is long;} \\ 1/2, & \text{if } T \text{ is short.} \end{cases}$$

For any interval of integers $[u, v]$, let

$$w([u, v]) = \sum_{T \text{ is an } I\text{-interval and } T \subseteq [u, v]} w(T).$$

By definition, every integer in $I$ creates either a long interval of weight 1 or two short intervals of weight $1/2$ each. By Lemma 3, all these intervals are disjoint, and by definition the two short intervals induced by an integer in $I$ are of distance $m - a$ apart. Hence, by Lemma 5, for any $n$,

$$w([0, n - 1]) - s/2 \leq |I \cap [0, n - 1]| \leq w([0, n - 1]) + s/2.$$

Thus to prove that $\lim_{n \to \infty} \frac{|I \cap [0, n - 1]|}{w([0, n - 1])} \geq \tau(D_{a,b,m})$, it suffices to show that $\lim_{n \to \infty} \frac{n}{w([0, n - 1])} \geq \tau(D_{a,b,m})$.

An interval $W = [x, y]$ of integers is called **neat** if every $I$-interval is either contained in $W$ or disjoint from $W$. Suppose $W$ is a neat interval. We define the $X$-ratio of $W$ to be

$$r(W) = \frac{|W|}{w(W)}.$$

To prove that $\lim_{n \to \infty} \frac{n}{|I \cap [0, n - 1]|} \geq \tau(D_{a,b,m})$, it suffices to find integers $a_1 < a_2 < \cdots$ such that for any $i$, $R_i = [a_i, a_{i+1} - 1]$ is a neat interval and $r(R_i) \geq \tau(D_{a,b,m})$.

We say an integer $p$ has property (*) if

(*) for the first Type-B $I$-interval $[u, u + a - 1]$ with $u \geq p$, we have $u \geq p + \Delta$. 

10
Lemma 6 Each $x_i \in I$ has property (*). Moreover, if $i \in L$, then $x_i + m$ also has property (*) and $[x_i, x_i + m - 1]$ is neat.

Proof. If $i \not\in L$, by Lemma 4, $x_i$ has property (*). Assume $i \in L$. By definition, $x_i$ has property (*). Suppose $x_i + m$ does not have property (*). Then, there exists some $u$ with $x_i + m \leq u < x_i + m + \Delta$ such that $[u, u + a - 1]$ is a Type-B $I$-interval. By definition, $u - m \in I$ and $[u - m, u - m + a - 1]$ is Type-A. This is impossible as $x_i \leq u - m < x_i + \Delta \leq x_{i+1}$ but $i \in L$. Hence, $x_i + m$ has property (*).

Now, assume to the contrary that $[x_i, x_i + m - 1]$ is not neat. Let $T = [u, v]$ be an $I$-interval that $T \cap [x_i, x_i + m - 1] \neq \emptyset$ and $T \not\subseteq [x_i, x_i + m - 1]$. By definition and as $i \in L$, $T$ must be Type-A. Hence, $u \in I$. Let $u = x_t$ for some $t$. Then $x_i + m - \Delta + 1 \leq t \leq x_i + m - 1$. This implies $b + 1 \leq t - x_i \leq m - 1$, a contradiction.

To complete the proof of Theorem 1, it suffices to find an infinite sequence of integers $a_1 < a_2 < \cdots$ such that the following hold for all $i$:

1. $a_i$ has property (*),
2. $R_i = [a_i, a_{i+1} - 1]$ is neat, and
3. $r(R_i) \geq \tau(D_{a,b,m})$.

We shall construct such a sequence of integers $a_1 < a_2 < \cdots$ inductively. Initially, set $a_1 = x_1$. By Lemma 6, $a_1$ has property (*). Assume we have determined $a_1, a_2, \ldots, a_i$, where (1 - 3) in the above are satisfied. We shall determine $a_{i+1}$ so that (1 - 3) still hold.

Let $[u, v]$ be the first $I$-interval with $u \geq a_i$. If $[u, v]$ is Type-B, then as $a_i$ has property (*), $u \geq a_i + \Delta$. Let $a_{i+1} = x_t$, where $x_t$ is the smallest element of $I$ for which $x_t > a_i$. Then all the $I$-intervals contained in $R_i = [a_i, a_{i+1} - 1]$ are Type-B, and $R_i$ is neat. Assume $R_i$ contains $j$ Type-B $I$-intervals. By Lemma 5, $j \leq s$. Since $w(R_i) = j/2$ and $|R_i| \geq \Delta + ja_i$, it follows that

$$r(R_i) \geq \frac{2(\Delta + ja_i)}{j} \geq 2a + \frac{2\Delta}{s} \geq \tau(D_{a,b,m}).$$
Observe that \( \frac{sa+m}{s+1} < a + \frac{b}{s+1} + \frac{\Delta}{s+1} < 2a + \frac{\Delta}{s+1} \). If \( m < 2qa \), then \( \frac{m}{q} < 2a \). If \( m < (2q+1)a \), then \( \frac{(2q+1)a+m}{2q^2} < 2a \). Moreover, by Lemma 6, \( a_{i+1} = x_t \) has property (*). Thus (1 - 3) in the above are satisfied.

In the following, assume \([u, v]\) is Type-A. Then \( u \in I \). Let \( x_h \) be the first element of \( I \) such that \( x_h \geq u \) and \( h \in L \). Let \( a_{i+1} = x_h + m \). By Lemma 6, \( R_i = [a_i, a_{i+1} - 1] \) is neat and \( a_{i+1} \) has property (*).

It remains to show (3). Assume the interval \([u, x_h - 1]\) contains \( j \) I-intervals for some \( j \geq 0 \). By Lemma 4, all the I-intervals contained in \([u, x_h - 1]\) are Type-A and short.

Since an I-interval of weight 1 has length \( \Delta \) and an I-interval of weight \( 1/2 \) has length \( a > \Delta/2 \), so for any interval \( T \) of length \( m \), we have

\[
w(T) \leq \begin{cases} 
q, & \text{if } m < q\Delta + a; \\
q + \frac{1}{2}, & \text{if } m \geq q\Delta + a.
\end{cases}
\]

Because \( R_i = [a_i, x_h - 1] \cup [x_h, x_h + m - 1] \), it follows that

\[
w(R_i) \leq \begin{cases} 
q + \frac{j}{2}, & \text{if } m < q\Delta + a; \\
q + \frac{j + 1}{2}, & \text{if } m \geq q\Delta + a.
\end{cases}
\]

Now we consider three cases.

**Case 1** \( m < q\Delta + a \). As \( |R_i| \geq ja + m \), by the above discussion, \( r(R_i) \geq \frac{ja+m}{q+j/2} \). Observe that \( \frac{ja+m}{q+j/2} \) is a function of \( j \) which is increasing if \( m \leq 2qa \) and decreasing if \( m \geq 2qa \). Hence, as \( j \leq s \), we have

- if \( m \geq 2qa \), then \( r(R_i) \geq \frac{sa+m}{q+1} \geq \frac{sa+m}{s+1} \);
- if \( m < 2qa \), then \( r(R_i) \geq \frac{0a+m}{q+0} \geq \frac{m}{q} \).

Hence, (3) holds.

**Case 2** \( m \geq (2q+1)a \). Similar to Case 1, we have \( r(R_i) \geq \frac{ja+m}{q+(j+1)/2} \). Because \( m \geq (2q+1)a \), which implies that \( \frac{ja+m}{q+(j+1)/2} \) is a decreasing function of \( j \), we conclude that \( r(R_i) \geq \frac{sa+m}{q+(s+1)/2} \). As \( \frac{b}{a} = \frac{m}{a} - \frac{\Delta}{a} \geq 2q + 1 - 2 \), we
have \( s = [b/a] \geq 2q - 1 \), i.e., \( q \leq (s + 1)/2 \). Hence \( r(R_i) \geq (sa + m)/(s + 1) \), so (3) holds.

**Case 3** \( a + q\Delta \leq m < (2q + 1)a \). Then \( r(R_i) \geq \frac{j a + m}{q + (j + 1)/2} \). Because \( m < (2q + 1)a \), \( \frac{j a + m}{q + (j + 1)/2} \) is an increasing function of \( j \). If \( j \geq 1 \), then \( r(R_i) \geq \frac{a + m}{q + 1} > \frac{(2q - 1)m + a}{2q^2} \). If \( j = 0 \) and \( w(R_i) \leq q \), then \( r(R_i) \geq \frac{m}{q} > \frac{(2q - 1)m + a}{2q^2} \), and we are done.

Assume \( j = 0 \) and \( w(R_i) = q + 1/2 \). Then \( u = h \) and \( r(R_i) \geq m/(q + 1/2) \). As \( \frac{m}{q + 1/2} < \frac{(2q - 1)m + a}{2q^2} = \tau(D_{a,b,m}) \), this “\( a_{i+1} \)” does not satisfy our requirement. We need to find a different \( a_{i+1} \) so that (1 - 3) are satisfied. In the following, we re-name the interval \([u, u + m - 1]\) just obtained by \( R_i^1 \). (The correct \( R_i \) is not found yet.)

Since \( w(R_i^1) = q + 1/2 \), \( R_i^1 \) contains a short \( I \)-interval. Let \( p_1 \leq q \) be the total weight of \( I \)-intervals preceding the last short \( I \)-interval in \( R_i^1 \). As \( w(R_i^1) = q + 1/2 \) and the first \( I \)-interval of \( R_i^1 \) is long, we know that \( p_1 \geq 1 \) is an integer.

Before reaching the correct interval \( R_i \), we may need a (finite) sequence of intervals \( R_i^j \), where \( R_i^1 \) is just the first one of them. In the following, we describe the inductive step of finding \( R_i^j \).

Suppose \( z \) is an integer, \( 1 \leq z \leq 2q - 1 \), and for \( j = 1, 2, \cdots, z \), we have obtained \( R_i^j = [x_{i_j}, x_{i_j} + m - 1] \) with the following properties:

- \( x_{i_j} \in I \) and \( i_j \in L \), and for \( j \geq 2 \), \( x_{i_{j-1}} + m \leq x_{i_j} < x_{i_{j-1}} + m + a \).
- \( w(R_i^j) = q + 1/2 \).

Observe that if \( w(R_i^j) = q + 1/2 \), the \( I \)-intervals in \( R_i^j \) must be “tightly packed”. Namely, if a neat sub-interval \( H \) of \( R_i^j \) has length \( \geq \alpha\Delta + \beta a \), where \( \alpha, \beta \) are non-negative integers, then \( w(H) \geq \alpha + \beta/2 \). For otherwise, \( w(R_i^j) \) will be less than \( q + 1/2 \).

Let \( p_j \) be the total weight of \( I \)-intervals preceding the last short \( I \)-interval in \( R_i^j \). Since \( w(R_i^j) = q + 1/2 \), \( R_i^j \) does contain a short \( I \)-interval. Since the first interval of \( R_i^j \) is a long interval, we have \( p_j \geq 1 \).
Let \([x_{i'}, x_{i'} + \Delta - 1]\) be the first long \(I\)-interval with \(x_{i'} \geq x_i + m\). If \(x_{i'} \geq x_i + m + a\), let \(a_{i+1} = x_{i'}\). Then \(R_i = [a_i, a_{i+1} - 1]\) is neat, \(|R_i| \geq zm + ja\) for some \(j \geq 1\), and \(w(R_i) \leq z(q + 1/2) + j/2\). Hence

\[
r(R_i) \geq \frac{zm + ja}{z(q + 1/2)} + \frac{j}{2} \geq \frac{zm + a}{z(q + 1/2)} + \frac{1}{2} \geq \frac{(2q - 1)m + a}{2q^2}.
\]

Note, all the \(I\)-intervals contained in \([x_i + m, x_{i'} - 1]\), if any, are short. The last inequality in the above holds since \(z \leq 2q - 1\) and \(\frac{zm + a}{z(q + 1/2)}\) is a decreasing function of \(z\). As \(a_{i+1} \in I\) has property (*), we are done for this case.

Assume \(x_{i'} \leq x_i + m + a - 1\). Let \(R_{i+1}^{z+1} = [x_{i'}, x_{i'} + m - 1]\). If \(w(R_{i+1}^{z+1}) \leq q\), then let \(a_{i+1} = x_{i'} + m\). By Lemma 6, \(R_i = [a_i, a_{i+1} - 1]\) is neat and \(a_{i+1}\) has property (*). To verify (3), we note that \(w(R_i) \leq (z+1)q + z/2\) and

\[
r(R_i) \geq \frac{(z+1)m}{(z+1)q + z/2} \geq \frac{2qm}{2q^2 + q - 1/2} \geq \frac{(2q - 1)m + a}{2q^2}.
\]

The second inequality in the above holds because \(\frac{(z+1)m}{(z+1)q + z/2}\) is a decreasing function of \(z\). The third inequality holds because \(m \geq a(q + 1)\). Thus we assume \(w(R_{i+1}^{z+1}) = q + 1/2\).

**Claim:** \(p_{z+1} \leq p_z\). Moreover, if \(p_{z+1} = p_z\), then the last short \(I\)-interval contained in \(R_{i}^z\) is Type-A, and the last short \(I\)-interval in \(R_{i}^{z+1}\) is Type-B.

**Proof of Claim.** Let \(T = [u, u + a - 1]\) and \(T' = [u', u' + a - 1]\) be the last short \(I\)-interval in \(R_{i}^z\) and \(R_{i}^{z+1}\), respectively. If \(T'\) is Type-B, then \(T'' = [u' - m, u' - m + a - 1]\) is a short Type-A \(I\)-interval contained in \(R_{i}^z\). Note, as \([u' - m, x_{i'} - 1] = [u', x_{i'} + m - 1]\) and \(x_{i'} \geq x_i + m\), we have \([x_{i'}/u' - m - 1] = [x_{i'}/u' - m - 1]\). Hence, \([x_{i'}/u' - m - 1]\) is capable of containing \(I\)-intervals of total weight at least \(p_{z+1}\). As the \(I\)-intervals in \(R_{i}^{z+1}\) are “tightly packed,” the \(I\)-intervals contained in \([x_i, u' - m - 1]\) has total weight at least \(p_{z+1}\). Therefore \(p_z \geq p_{z+1}\), and if the equality holds then the last short \(I\)-interval in \(R_{i}^{z+1}\) is of Type A.

Assume \(T'\) is Type-A. Thus \(u' = x_{i'} \in I\) for some \(i'\). Since \(T'\) is short, \(x_{i'+1} \leq x_{i'} + \Delta - 1\). Note, as \(s' \in L\) and \(x_{i'}, x_{i'+1} \in I\), we have \(x_{i'+1} \leq x_{i'} + b\) and \([x_{i'} - m + 1, x_{i'+1} - b - 1] \cap I = \emptyset\).

Consider the interval \([x_{i'}, x_{i'+1} - b - 1]\). If this is a sub-interval of \(R_{i}^z\),
then since
\[ |[x_{i_*}, x_{i_*+1} - 1]| \geq |x_{i_*,} x_{i_*+1} - 1| + \Delta, \]
and the interval \([x_{i_*}, x_{i_*+1} - 1]\) is tightly packed, we conclude that the total weight of the I-intervals that intersect with \([x_{i_*}, x_{i_*+1} - 1]\) is at least \(p_{z+1} + 1 + 1/2\). Moreover, since \(|[x_{i_*}, - m + 1, x_{i_*+1} - b - 1]| \geq \Delta + a - 1\) and \([x_{i_*} - m + 1, x_{i_*+1} - b - 1] \cap I = \emptyset\), we conclude that the last I-interval intersecting with \([x_{i_*}, x_{i_*+1} - 1]\) is Type-B. The total weight of the I-intervals of \(R_i^z\) preceding this Type-B I-interval is at least \(p_{z+1} + 1\). Therefore, \(p_{z+1} < p_z\).

Assume \([x_{i_*}, x_{i_*+1} - 1]\) is not a sub-interval of \(R_i^z\). Then \(x_{i_*+1} - b - 1 \geq x_{i_*} + m\). Since \(x_{i_*} \geq x_{i_*'} - m - a + 1\), we have \(x_{i_*+1} \geq x_{i_*'} + b - a + 3\). This implies that \([x_{i_*+1} + 1, x_{i_*'} + m] \cap I = \emptyset\), and \([x_{i_*+1}, x_{i_*+1} + \Delta - 1] \) is the last I-interval contained in \(R_i^{z+1}\). Hence \(p_{z+1} = q - 1\). If \(p_z \leq q - 2\), then the conclusion follows.

Assume \(p_z = q - 1\). Then the last I-interval in \(R_i^z\) is a long interval. Since \([x_{i_*} - m + 1, x_{i_*} + m - 1] \cap I = \emptyset\), the last integer of \(I\) in \(R_i^z\) is not greater than \(x_{i_*} - m\), implying the last short I-interval in \(R_i^z \) is contained in \([x_{i_*}, x_{i_*} - m - 1]\). Therefore, the interval \([x_{i_*}, x_{i_*} - m - 1]\) has length at least \((q - 1)\Delta + a\). Moreover, since
\[ |[x_{i_*} - m, x_{i_*'} - 1]| \geq |[x_{i_*} - m, x_{i_*+1} - 1]| \geq \Delta + a, \]
we conclude, \([x_{i_*}, x_{i_*'} - 1]\) has length at least \(q\Delta + 2a\). Let \(a_{i+1} = x_{i_*'}\). Then
\[ r(R_i) \geq \frac{(z - 1)m + q\Delta + 2a}{z(q + \frac{1}{2})} \geq \frac{(2q - 2)m + q\Delta + 2a}{(2q - 1)(q + \frac{1}{2})}. \]
The second inequality holds because the formula is a decreasing function on \(z\) and \(z \geq 2q - 1\). To complete the proof of the Claim, it suffices to show
\[ \frac{(2q - 2)m + q\Delta + 2a}{(2q - 1)(q + \frac{1}{2})} \geq \frac{(2q - 1)m + a}{2q^2}. \]
Write \(m = q\Delta + 2a - \lambda\), where \(0 < \lambda \leq a\). The above inequality is equivalent to
\[ 2q^2\lambda - (2q^2 - 1)/2a - m(1/2 - q) \geq 0. \]
By definition, we have:
(1) \( \lambda \geq 2a - \Delta + 1 \) (since \( q = \lfloor m/\Delta \rfloor \))

(2) \( \Delta \leq 2a - 1 \) (since \( 2a > \Delta \))

Therefore,

\[
2q^2\lambda - (2q^2 - 1/2)a - m(1/2 - q) \\
= (2q^2 - q + 1/2)\lambda - a(2q^2 - 2q + 1/2) - \Delta(q/2 - q^2) \\
\geq a(2q^2 + 1/2) - \Delta(q^2 - q/2 + 1/2) + (2q^2 - q + 1/2) \quad \text{(by (1))} \\
\geq a(q - 1/2) + 3q^2 - (3q)/2 + 1 \quad \text{(by (2))} \\
\geq 0 \quad \text{(since } q \geq 1) 
\]

This completes the proof of the Claim. \( \Box \)

Since \( p_i \geq 1 \), so \( p_{2q} \) does not exist. Thus the procedure above terminates at the \( k \)-th step for some \( k \leq 2q \), when the valid \( a_{i+1} \) is obtained. This completes the proof of Theorem 1.

Acknowledgment

The authors wish to thank the anonymous referees for their comments which resulted in a better presentation of the article. Daphne also wishes to thank Professor Ko-Wei Lih and the Institute of Mathematics, Academia Sinica, Taiwan, for their hospitality during her visit in 2005, when part of the work was done.

References


