A 2/3 Lower Bound for Progression-Free Harmonic Sums in Terms of van der Waerden Numbers

Tim Held

Independent researcher, Switzerland

Version 1.0, August 28, 2026

Abstract

For an integer $k \geq 3$, let $f(k)$ be the supremum of $\sum_{n \in A} 1/n$ over all sets A of positive integers containing no k-term arithmetic progression, and let $W(k) = W(2, k)$ be the two-color van der Waerden number. The classical coloring argument gives $f(k) \geq \frac{1}{2} H_{W(k)-1}$ and hence $\liminf_{k\to\infty} f(k)/\log W(k) \geq \frac{1}{2}$. We prove the stronger explicit estimate

$$f(k) \geq 1 + \frac{2W(k) - 3}{3W(k) - 5}\left(H_{W(k)-1} - \frac{3}{2}\right),$$

which implies

$$\liminf_{k\to\infty} \frac{f(k)}{\log W(k)} \geq \frac{2}{3}.$$

The proof starts with a two-coloring of $[W(k) - 1]$ having no monochromatic k-term progression. Each finite color class is normalized and used as the digit set of a shifted Kempner set in base $2W(k) - 3$. A modular progression argument proves that the resulting infinite set is k-term-progression-free, while a leading-digit decomposition gives a geometric amplification of its harmonic sum. A weighted averaging step then combines the two color classes. We also give an r-color extension and a coloring-sensitive refinement. The result is a strict improvement of the constant 1/2, but it does not prove the stronger conjecture that $f(k)/\log W(k) \to \infty$.

Status and scope. This manuscript is a non-peer-reviewed preprint, and independent verification is welcomed. The principal result is a partial advance on Erdős Problem 169: it proves the weaker assertion that the classical constant 1/2 can be replaced by a larger absolute constant, while leaving the stated limit-to-infinity problem open. No claim is made here that the full problem has been solved.

2020 Mathematics Subject Classification. Primary 11B25; Secondary 05D10, 11B05.
Keywords. Arithmetic progressions, van der Waerden numbers, harmonic sums, Kempner sets, digit-restricted sets, additive combinatorics.

1 Introduction

A k-term arithmetic progression, abbreviated k-AP, is a set of the form

$$\{a, a + d, \ldots, a + (k - 1)d\}, \qquad a, d \in \mathbb{N},\ d > 0.$$

A set of positive integers is called *k-AP-free* if it contains no such progression. Define

$$f(k) := \sup\left\{\sum_{n\in A} \frac{1}{n} : A \subseteq \mathbb{N} \text{ is } k\text{-AP-free}\right\}. \tag{1}$$

The finiteness of these quantities is tied to the Erdős–Turán conjecture on arithmetic progressions [5, 6]. The present paper concerns only unconditional lower bounds, so no finiteness assumption is used.

For integers $r \geq 2$ and $k \geq 2$, let $W(r,k)$ be the least positive integer N such that every coloring of $[N] := \{1, \ldots, N\}$ with r colors contains a monochromatic k-AP. We write

$$W(k) := W(2,k).$$

Van der Waerden's theorem guarantees that these numbers exist [9]. Erdős and Graham asked for estimates on $f(k)$ in terms of $W(k)$ and observed the elementary inequality

$$f(k) \geq \frac{1}{2} H_{W(k)-1}, \tag{2}$$

where $H_N := \sum_{n=1}^{N} 1/n$ is the Nth harmonic number. In particular,

$$\liminf_{k\to\infty} \frac{f(k)}{\log W(k)} \geq \frac{1}{2}.$$

They explicitly asked whether the constant $1/2$ could be increased, and more strongly whether

$$\lim_{k\to\infty} \frac{f(k)}{\log W(k)} = \infty. \tag{3}$$

See [3, 4, 2]. Berlekamp obtained the linear lower bound

$$f(k) \geq \frac{\log 2}{2} k,$$

and Gerver later proved the stronger asymptotic estimate

$$f(k) \geq (1 - o(1)) k \log k \tag{4}$$

[1, 6]. The estimate in this paper addresses a different scale: it specifically improves the coefficient attached to $\log W(k)$.

Our main theorem is the following.

Theorem 1.1 (Main theorem). *For every integer $k \geq 3$,*

$$f(k) \geq 1 + \frac{2W(k) - 3}{3W(k) - 5} \left(H_{W(k)-1} - \frac{3}{2} \right). \tag{5}$$

Consequently,

$$\liminf_{k\to\infty} \frac{f(k)}{\log W(k)} \geq \frac{2}{3}. \tag{6}$$

The original derivation of the argument gave the slightly weaker but still sufficient estimate

$$f(k) \geq 1 + \frac{2W(k) - 3}{3W(k) - 5} \left(H_{W(k)-1} - 2 \right). \tag{7}$$

Keeping track of the distinct minima of the two color classes replaces the constant 2 by $3/2$, producing equation (5).

The mechanism is simple. Begin with a two-coloring of $[W(k) - 1]$ having no monochromatic k-AP. Normalize a color class C by subtracting its minimum, and use the resulting finite set as the allowed base-b digits of a Kempner set. If b is greater than twice the diameter of C, ordinary k-AP-freeness implies modular k-AP-freeness. A base-b descent then shows that the infinite digit-restricted set is k-AP-free. Its successive digit layers contribute a geometric series to the harmonic sum, with ratio $|C|/b$. Applying this construction to both color classes and averaging the two resulting bounds yields the constant $2/3$.

Digit-restricted harmonic series originate with Kempner's classical example [7]. The relevant facts about progression-free digit sets are closely related to the framework developed by Walker and Walker [11] and by Walker [10]. Complete proofs are included here so that the present argument is self-contained.

2 Notation and elementary inequalities

Throughout, $k \geq 3$ unless stated otherwise. For a set $A \subseteq \mathbb{N}$, write

$$\mathcal{H}(A) := \sum_{a \in A} \frac{1}{a}$$

with the value $+\infty$ permitted. For a nonempty finite set C of integers, define

$$|C| = \text{its cardinality}, \qquad \operatorname{diam}(C) := \max C - \min C.$$

Translations preserve arithmetic progressions: C is k-AP-free if and only if $C+t$ is k-AP-free for every integer t.

We will repeatedly use a weighted averaging inequality.

Lemma 2.1 (Weighted maximum)**.** *Let $u_1, \dots, u_q \in \mathbb{R}$ and let $d_1, \dots, d_q > 0$. Then*

$$\max_{1 \leq i \leq q} \frac{u_i}{d_i} \geq \frac{\sum_{i=1}^{q} u_i}{\sum_{i=1}^{q} d_i}. \tag{8}$$

Proof. The quotient on the right is the weighted average

$$\frac{\sum_i d_i (u_i/d_i)}{\sum_i d_i}$$

of the numbers u_i/d_i, with positive weights d_i. A weighted average cannot exceed the largest entry. No sign condition on the u_i is needed. $\square$

For later limit arguments, note also that

$$W(r,k) \geq k+1 \qquad (r \geq 2,\ k \geq 3). \tag{9}$$

Indeed, color the interval $[k]$ with at least two colors. Its only k-term arithmetic progression is the whole interval $[k]$, which is not monochromatic. Thus $W(r,k) > k$.

We also record the standard harmonic-number asymptotic

$$H_N = \log N + \gamma + O(N^{-1}), \tag{10}$$

where γ is the Euler–Mascheroni constant.

3 The classical $1/2$ bound

The starting point is the short coloring argument of Erdős and Graham.

Proposition 3.1. *For every $k \geq 3$,*

$$f(k) \geq \frac{1}{2} H_{W(k)-1}.$$

Consequently,

$$\liminf_{k\to\infty} \frac{f(k)}{\log W(k)} \geq \frac{1}{2}.$$

Proof. Put $N = W(k) - 1$. By the minimality of $W(k)$, there is a two-coloring

$$[N] = C_0 \sqcup C_1$$

with no monochromatic k-AP. Thus both C_0 and C_1 are k-AP-free. Moreover,

$$\mathcal{H}(C_0) + \mathcal{H}(C_1) = H_N.$$

At least one of the two sums is at least $H_N/2$, and that color class is admissible in the definition of $f(k)$. This proves the finite inequality. The asymptotic statement follows from equation (10) and $W(k) \to \infty$. □

The rest of the paper amplifies the harmonic sum of each finite color class before the final averaging step.

4 Modular arithmetic progressions and Kempner sets

4.1 Progressions modulo a base

Definition 4.1. Let $b \geq 2$ and $S \subseteq \{0, 1, \dots, b-1\}$. We say that S contains a *k-term arithmetic progression modulo b* if there are integers x and Δ with $b \nmid \Delta$ such that

$$x + j\Delta \pmod{b} \in S \qquad (j = 0, 1, \dots, k-1).$$

If no such pair exists, S is called *k-AP-free modulo b*.

The difference Δ need not be coprime to b, and the residues need not be distinct. This is the notion needed for the digit descent below.

Proposition 4.2 (Large-base modular freeness)**.** *Let $S \subseteq [0, M]$ be an ordinary k-AP-free set. If $b > 2M$, then S is k-AP-free modulo b.*

Proof. Assume for contradiction that there are x and Δ as in theorem 4.1. Replacing Δ by its least positive residue, we may suppose

$$1 \leq \delta \leq b-1, \qquad \delta \equiv \Delta \pmod{b}.$$

Let

$$r_j \in S \subseteq [0, M]$$

be the least nonnegative residue of $x + j\delta$ modulo b. Since $0 < \delta < b$, each successive difference is either

$$r_{j+1} - r_j = \delta \quad \text{or} \quad r_{j+1} - r_j = \delta - b.$$

If a step of the first type occurs, then $\delta \le M$, because both residues lie in $[0, M]$. If a step of the second type occurs, then $b - \delta \le M$.

If both step types occur, then

$$b = \delta + (b - \delta) \le 2M,$$

contrary to $b > 2M$. Hence every step has the same type. If all steps are $+\delta$, then $r_0, r_1, \dots, r_{k-1}$ is an ordinary increasing k-AP in S. If all steps are $\delta - b$, then the same residues, read in reverse order, form an ordinary increasing k-AP in S. Both alternatives contradict the assumed k-AP-freeness of S. $\square$

4.2 Digit-restricted sets

Definition 4.3 (Kempner set)**.** Let $b \ge 2$ and let $S \subseteq \{0, 1, \dots, b-1\}$ with $0 \in S$. The associated Kempner set is

$$\mathcal{K}(S, b) := \left\{ \sum_{j=0}^{L} a_j b^j : L \ge 0,\ a_j \in S \right\}.$$

Equivalently, $\mathcal{K}(S, b)$ is the set of nonnegative integers all of whose canonical base-b digits belong to S.

Theorem 4.4 (Digit descent)**.** *If S is k-AP-free modulo b and $0 \in S$, then $\mathcal{K}(S, b)$ is ordinary k-AP-free. Consequently, $\mathcal{K}(S, b) + 1$ is a k-AP-free set of positive integers.*

Proof. Suppose, toward a contradiction, that

$$x, x + d, \dots, x + (k-1)d$$

is contained in $\mathcal{K}(S, b)$, where $d > 0$. Let $t \ge 0$ be the largest integer for which $b^t \mid d$. We remove t common terminal digits.

If $b \mid d$, all progression terms have the same units digit, say $a_0 \in S$. After subtracting a_0 from every term and dividing by b, we obtain another k-AP contained in $\mathcal{K}(S, b)$, now with common difference d/b. Repeating this operation exactly t times produces a k-AP in $\mathcal{K}(S, b)$ whose common difference $d' = d/b^t$ is not divisible by b.

Reducing the final progression modulo b gives

$$x' + jd' \pmod{b} \in S \qquad (j = 0, \dots, k-1),$$

with $b \nmid d'$. This is a k-term arithmetic progression modulo b in S, contradicting the hypothesis. Therefore $\mathcal{K}(S, b)$ is k-AP-free. Translation by 1 preserves arithmetic progressions, so $\mathcal{K}(S, b) + 1$ is also k-AP-free. $\square$

Combining theorems 4.2 and 4.4 gives the form used later.

Corollary 4.5 (Finite seed to infinite set)**.** *Let C be a nonempty finite k-AP-free set of integers. Put*

$$m = \min C, \qquad S = C - m, \qquad M = \operatorname{diam}(C).$$

For every integer $b > 2M$, the shifted Kempner set

$$\mathcal{K}(S, b) + 1$$

is a k-AP-free set of positive integers.

Proof. The translated set S is ordinary k-AP-free, lies in $[0, M]$, and contains 0. Apply theorems 4.2 and 4.4. $\square$

5 Harmonic-sum amplification

The next estimate is the quantitative component of the argument.

Proposition 5.1 (Leading-digit lower bound)**.** *Let $b \geq 2$, let $S \subsetneq \{0, 1, \dots, b-1\}$ contain 0, and write $s = |S|$. Then*

$$\mathcal{H}(\mathcal{K}(S,b) + 1) \geq 1 + \frac{\displaystyle\sum_{a \in S \setminus \{0\}} \frac{1}{a+1}}{1 - \dfrac{s}{b}}. \tag{11}$$

In particular, the harmonic series of $\mathcal{K}(S,b) + 1$ converges whenever $s < b$.

Proof. The element $0 \in \mathcal{K}(S,b)$ contributes 1 after the shift by 1. Now fix an integer $\ell \geq 1$ and consider the positive elements of $\mathcal{K}(S,b)$ whose canonical base-b expansion has exactly ℓ digits. Each such integer has a unique representation

$$n = ab^{\ell-1} + q,$$

where

$$a \in S \setminus \{0\}$$

is the leading digit and q is an integer represented by $\ell - 1$ digits from S, allowing leading zeroes in this lower block. There are exactly $s^{\ell-1}$ choices for q.

Since $0 \leq q \leq b^{\ell-1} - 1$, we have

$$n + 1 = ab^{\ell-1} + q + 1 \leq (a+1)b^{\ell-1}.$$

Therefore

$$\frac{1}{n+1} \geq \frac{1}{(a+1)b^{\ell-1}}.$$

Summing first over all q, then over all nonzero leading digits a, the contribution of the ℓ-digit layer is at least

$$\left(\frac{s}{b}\right)^{\ell-1} \sum_{a \in S \setminus \{0\}} \frac{1}{a+1}.$$

Summing these lower bounds over $\ell \geq 1$ and adding the contribution of 0 gives

$$\mathcal{H}(\mathcal{K}(S,b) + 1) \geq 1 + \left(\sum_{a \in S \setminus \{0\}} \frac{1}{a+1} \right) \sum_{j=0}^{\infty} \left(\frac{s}{b}\right)^{j}.$$

Because S is a proper subset of the b possible digits, $s < b$, so the geometric series equals $(1 - s/b)^{-1}$. This proves equation (11). For completeness, convergence follows from a separate upper estimate. Every positive element in the ℓ-digit layer is at least $b^{\ell-1}$, while that layer contains at most s^{ℓ} elements. Consequently its harmonic contribution after shifting is at most $b(s/b)^{\ell}$. The sum of these upper bounds is geometric because $s < b$. $\square$

For a finite seed C, it is convenient to introduce its normalized harmonic weight

$$\Phi(C) := \sum_{c \in C} \frac{1}{c - \min C + 1}. \tag{12}$$

Theorem 5.2 (Finite-seed amplification)**.** *Let C be a nonempty finite k-AP-free set of positive integers, and let b be an integer satisfying*

$$b > 2\,\mathrm{diam}(C).$$

Then

$$f(k) \geq 1 + \frac{\Phi(C) - 1}{1 - |C|/b}. \tag{13}$$

Proof. Let $m = \min C$ and $S = C - m$. By theorem 4.5, the set $\mathcal{K}(S, b) + 1$ is k-AP-free. Moreover,

$$\sum_{a \in S \setminus \{0\}} \frac{1}{a+1} = \sum_{c \in C \setminus \{m\}} \frac{1}{c - m + 1} = \Phi(C) - 1.$$

Applying theorem 5.1 and then the definition of $f(k)$ gives equation (13). □

Remark 5.3. For every nonempty finite $C \subseteq \mathbb{N}$ with $m = \min C$,

$$\Phi(C) \geq \mathcal{H}(C) + 1 - \frac{1}{m}. \tag{14}$$

Indeed, the term $c = m$ changes from $1/m$ to 1, and every other denominator decreases from c to $c - m + 1 \leq c$. The correction $1 - 1/m$ is the source of the improvement from $H_N - 2$ to $H_N - 3/2$ in the two-color theorem.

6 Amplifying a partition of an interval

We first state a coloring-sensitive form that retains the individual class sizes, spans, and normalized harmonic weights.

Proposition 6.1 (Coloring-sensitive partition bound)**.** *Suppose*

$$[N] = C_1 \sqcup \cdots \sqcup C_q$$

is a partition into nonempty k-AP-free sets. For each i, choose an integer $b_i > 2\,\mathrm{diam}(C_i)$ and put

$$s_i = |C_i|, \qquad d_i = 1 - \frac{s_i}{b_i}.$$

Then

$$f(k) \geq 1 + \frac{\sum_{i=1}^{q} (\Phi(C_i) - 1)}{\sum_{i=1}^{q} \left(1 - \frac{s_i}{b_i}\right)}. \tag{15}$$

Proof. By theorem 5.2, for every i,

$$f(k) \geq 1 + \frac{\Phi(C_i) - 1}{d_i}.$$

The denominators d_i are positive because $s_i \leq \mathrm{diam}(C_i) + 1 < b_i$. Taking the maximum over i and applying theorem 2.1 with $u_i = \Phi(C_i) - 1$ proves equation (15). □

For a universal estimate depending only on N and the number of classes, we use a common safe base.

Theorem 6.2 (Universal partition amplification)**.** *Suppose*

$$[N] = C_1 \sqcup \cdots \sqcup C_q$$

is a partition into $q \geq 1$ nonempty k-AP-free sets. Then

$$f(k) \geq 1 + \frac{H_N - H_q}{q - \dfrac{N}{2N-1}} = 1 + \frac{2N-1}{(2q-1)N - q}(H_N - H_q). \tag{16}$$

Proof. Choose the common base

$$b = 2N - 1.$$

For every class C_i,

$$\operatorname{diam}(C_i) \leq N - 1,$$

so $b > 2\operatorname{diam}(C_i)$. Apply theorem 6.1 with $b_i = b$ for all i. Since $\sum_i s_i = N$, its denominator is

$$\sum_{i=1}^{q}\left(1 - \frac{s_i}{b}\right) = q - \frac{N}{b} = q - \frac{N}{2N-1}. \tag{17}$$

It remains to lower-bound the numerator. Let

$$m_i = \min C_i.$$

By equation (14),

$$\Phi(C_i) - 1 \geq \mathcal{H}(C_i) - \frac{1}{m_i}.$$

Summing over the partition gives

$$\sum_{i=1}^{q}(\Phi(C_i) - 1) \geq H_N - \sum_{i=1}^{q}\frac{1}{m_i}. \tag{18}$$

The minima $m_1, \ldots, m_q$ are distinct positive integers. The sum of their reciprocals is therefore at most the sum of the reciprocals of the q smallest positive integers:

$$\sum_{i=1}^{q}\frac{1}{m_i} \leq H_q. \tag{19}$$

Combining equations (17) to (19) with equation (15) proves the first expression in equation (16). Multiplying numerator and denominator by $2N - 1$ gives the second expression. □

Remark 6.3 (The initially advertised estimate). If one uses only the cruder inequality $\Phi(C_i) \geq \mathcal{H}(C_i)$, then the numerator in equation (16) is replaced by $H_N - q$. For $q = 2$, this gives exactly

$$f(k) \geq 1 + \frac{2N-1}{3N-2}(H_N - 2),$$

which is the bound in the original proof sketch. The minima argument costs nothing and improves $H_N - 2$ to $H_N - H_2 = H_N - 3/2$.

7 Proof of the main theorem

Proof of theorem 1.1. Set

$$N = W(k) - 1.$$

By the definition of $W(k)$, there exists a two-coloring

$$[N] = C_0 \sqcup C_1$$

with neither color class containing a k-AP. The two classes are nonempty: if one were empty, the other would equal $[N]$, which contains k consecutive integers and therefore a k-AP, since $N \geq k$.

Apply theorem 6.2 with $q = 2$. Since $H_2 = 3/2$,

$$f(k) \geq 1 + \frac{2N-1}{3N-2}\left(H_N - \frac{3}{2}\right).$$

Substituting $N = W(k) - 1$ yields

$$2N - 1 = 2W(k) - 3, \qquad 3N - 2 = 3W(k) - 5,$$

and therefore

$$f(k) \geq 1 + \frac{2W(k)-3}{3W(k)-5}\left(H_{W(k)-1} - \frac{3}{2}\right).$$

This is equation (5).

For the asymptotic consequence, note that $W(k) \geq k$, hence $W(k) \to \infty$. With $N = W(k) - 1$, we have

$$\frac{2N-1}{3N-2} \longrightarrow \frac{2}{3}, \qquad \frac{H_N - 3/2}{\log(N+1)} \longrightarrow 1$$

by equation (10). Dividing the finite inequality by $\log W(k) = \log(N+1)$ and taking the lower limit gives

$$\liminf_{k\to\infty} \frac{f(k)}{\log W(k)} \geq \frac{2}{3}.$$

□

The new finite bound is not merely asymptotically stronger than the classical coloring bound. It is strictly stronger for every relevant N.

Corollary 7.1 (Exact gain over the classical bound)**.** *Let $N = W(k) - 1$. Then the right-hand side of equation* (5) *exceeds $H_N/2$ by*

$$\frac{NH_N - 1}{2(3N-2)}. \tag{20}$$

In particular, the improvement is strict for every $k \geq 3$.

Proof. Write $a_N = (2N-1)/(3N-2)$. Direct algebra gives

$$\begin{aligned}\left[1 + a_N\left(H_N - \frac{3}{2}\right)\right] - \frac{H_N}{2} &= 1 - \frac{3}{2}a_N + \left(a_N - \frac{1}{2}\right)H_N \\ &= \frac{NH_N - 1}{2(3N-2)}.\end{aligned}$$

For $k \geq 3$, one has $N \geq 2$, so $NH_N > 1$. □

Since $H_N - 3/2 \geq H_N - 2$, equation (5) also immediately implies the originally claimed bound equation (7).

8 An r-color extension

The same argument applies to any number of colors.

Theorem 8.1 (Van der Waerden extension)**.** *Let $r \geq 2$ and $k \geq 3$, and put*

$$N = W(r, k) - 1.$$

Then

$$f(k) \geq 1 + \frac{2N - 1}{(2r - 1)N - r} (H_N - H_r). \tag{21}$$

Consequently, for every fixed $r \geq 2$,

$$\liminf_{k \to \infty} \frac{f(k)}{\log W(r, k)} \geq \frac{2}{2r - 1}. \tag{22}$$

Proof. There is an r-coloring of $[N]$ without a monochromatic k-AP. Discard any empty color classes and let $q \leq r$ be the number of nonempty classes. By theorem 6.2,

$$f(k) \geq 1 + \frac{H_N - H_q}{q - N/(2N - 1)}.$$

Because $q \leq r$, we have $H_N - H_q \geq H_N - H_r$, while $q - N/(2N - 1) \leq r - N/(2N - 1)$. Whenever $H_N - H_r \geq 0$, replacing the numerator by the smaller quantity and the denominator by the larger one preserves a valid lower bound. If $H_N - H_r < 0$, the displayed target bound is weaker than the trivial inequality $f(k) \geq 1$ and is automatic. Thus

$$f(k) \geq 1 + \frac{H_N - H_r}{r - N/(2N - 1)},$$

which is equivalent to equation (21).

For fixed r, $W(r, k) \to \infty$ as $k \to \infty$. Hence $H_N - H_r \sim \log N$ and

$$\frac{2N - 1}{(2r - 1)N - r} \longrightarrow \frac{2}{2r - 1}.$$

Dividing by $\log W(r, k) = \log(N + 1)$ proves equation (22). ☐

9 Numerical illustrations

The exact two-color van der Waerden numbers currently known are

$$W(3) = 9, \qquad W(4) = 35, \qquad W(5) = 178, \qquad W(6) = 1132.$$

These values are recorded in OEIS A005346 [8]. The following table compares three bounds obtained from them. The second column of bounds is the classical coloring estimate $H_N/2$. The third is the initially advertised estimate with $H_N - 2$. The fourth is the strengthened estimate of equation (5), with $N = W(k) - 1$.

These values are included only to illustrate the finite formula. For small k, specialized constructions, including those of Gerver and later Kempner-set searches, give stronger numerical lower bounds. The point of theorem 1.1 is the uniform coefficient of $\log W(k)$.

Table 1: Illustrative finite lower bounds derived from exact values of $W(k)$.

k	$W(k)$	$\frac{1}{2}H_N$	$1+\frac{2N-1}{3N-2}(H_N-2)$	$1+\frac{2N-1}{3N-2}(H_N-\frac{3}{2})$
3	9	1.358929	1.489448	1.830357
4	35	2.059105	2.419201	2.754201
5	178	2.878094	3.506492	3.840140
6	1132	3.804258	4.739561	5.072944

9.1 A concrete three-term-progression-free example

The case $k=3$ gives a direct check of every part of the construction. Since $W(3)=9$, the interval $[8]$ admits the progression-free two-coloring

$$C_0=\{1,2,5,6\}, \qquad C_1=\{3,4,7,8\}.$$

Both normalized color classes are equal to

$$S=\{0,1,4,5\}.$$

The universal base is $b=2\cdot 8-1=15$, and $15>2\max S=10$. Consequently

$$A=\mathcal{K}(\{0,1,4,5\},15)+1=\{1,2,5,6,16,17,20,21,61,62,65,66,\ldots\}$$

is 3-AP-free. Retaining the exact normalized weight of either color class in theorem 5.2 gives

$$\begin{aligned}\mathcal{H}(A) &\geq 1+\frac{\frac{1}{2}+\frac{1}{5}+\frac{1}{6}}{1-\frac{4}{15}}\\ &=1+\frac{13/15}{11/15}=\frac{24}{11}\approx 2.181818.\end{aligned}$$

This exceeds the universal value 1.830357 in table 1 because the universal theorem deliberately discards coloring-specific information in order to obtain a formula depending only on $W(k)$.

10 Discussion, limitations, and possible refinements

10.1 What has and has not been proved

The argument is unconditional and constructs explicit infinite k-AP-free sets from any extremal van der Waerden coloring. It proves a strict improvement of the coefficient $1/2$ in the lower bound involving $\log W(k)$. It does not establish equation (3); the ratio is shown only to have lower limit at least $2/3$.

No assumption that $f(k)$ is finite is made. If $f(k)=+\infty$ for some k, all stated lower bounds remain valid. The constructed Kempner sets have convergent harmonic sums because their allowed-digit ratio satisfies $|S|/b<1$.

10.2 Why the constant $2/3$ appears

In a balanced two-coloring of an interval of length N, a typical color class has about $N/2$ elements. The universally safe base in the argument is approximately $2N$, because the elementary modular-freeness lemma requires a base greater than twice the class diameter. Thus the geometric ratio in the harmonic amplification is approximately

$$\frac{|C|}{b} \approx \frac{1}{4},$$

so a color-class harmonic sum is multiplied by approximately

$$\frac{1}{1 - 1/4} = \frac{4}{3}.$$

The original coloring supplies approximately one half of the interval's harmonic mass. Multiplying $1/2$ by $4/3$ gives $2/3$.

This heuristic explains the constant produced by the universal argument, but it is not an optimality theorem. Particular colorings may have smaller class diameters, more favorable normalized harmonic weights, or digit sets that are already progression-free modulo bases substantially smaller than twice their diameter. Theorem 6.1 retains exactly this information.

10.3 Possible routes beyond the universal bound

Any improvement beyond $2/3$ by this circle of ideas would need additional input not used in the universal estimate. Potential directions include:

(i) proving that some color class in every extremal two-coloring is progression-free modulo a base smaller than $2\,\mathrm{diam}(C) + 1$;

(ii) exploiting carries rather than excluding them through the condition $b > 2\,\mathrm{diam}(C)$;

(iii) combining information from both color classes inside a single digit system instead of constructing two separate Kempner sets;

(iv) using the class-specific bases and normalized weights in theorem 6.1, together with structural information on extremal van der Waerden colorings.

These are suggestions for further investigation, not claims that any of the approaches must succeed.

10.4 Relation to Gerver's bound

Combining equation (4) and theorem 1.1 gives the unconditional summary

$$f(k) \geq \max\left\{(1 - o(1))k \log k, \left(\frac{2}{3} - o(1)\right) \log W(k)\right\}.$$

The two terms measure different aspects of the problem. Gerver's estimate is direct in k, whereas the present estimate is designed to answer the specific van der Waerden-number question posed by Erdős and Graham.

A Proof audit and edge cases

This appendix isolates the logical dependencies and common points at which a digit argument can fail.

A.1 Canonical expansions and leading zeroes

Every nonnegative integer has a unique canonical base-b expansion, except that arbitrary leading zeroes may be appended. In theorem 5.1, the leading digit a is required to be nonzero, so each positive integer is counted exactly once. The lower block q is allowed to have leading zeroes, which gives exactly $s^{\ell-1}$ choices.

A.2 Why modular freeness is necessary

Ordinary k-AP-freeness of the digit set alone does not generally prevent an arithmetic progression whose residues wrap around modulo b. The condition $b > 2M$ in theorem 4.2 prevents a mixture of wrapped and unwrapped steps. This is precisely where the factor of approximately two in the chosen base enters.

A.3 Termination of the descent

The descent in theorem 4.4 terminates because a positive integer d is divisible by only finitely many powers of b. Primality of b is irrelevant. The relevant exponent is simply the largest t for which $b^t \mid d$.

A.4 Positivity of denominators

In theorem 5.2, C lies in an interval of length $\operatorname{diam}(C) + 1$, so

$$|C| \leq \operatorname{diam}(C) + 1 < b$$

when $b > 2 \operatorname{diam}(C)$ and $\operatorname{diam}(C) \geq 1$. If $\operatorname{diam}(C) = 0$, then $|C| = 1$ and one may take any $b \geq 2$. Thus $1 - |C|/b > 0$ in all cases.

A.5 The weighted average with possibly negative numerators

For a very small class, $\Phi(C_i) - 1$ can be zero, and cruder substitutes such as $\mathcal{H}(C_i) - 1$ can be negative. Theorem 2.1 remains valid without a positivity assumption on the numerators. Only the denominators must be positive.

A.6 No circular use of van der Waerden's theorem

Van der Waerden's theorem is used only to know that $W(k)$ exists and that, by minimality, an avoiding coloring of $[W(k) - 1]$ exists. The construction of the infinite k-AP-free set is then verified directly by modular arithmetic and digit descent.

A.7 Algebraic check of the final coefficient

With $N = W(k) - 1$ and $b = 2N - 1$, the sum of the two geometric denominators is

$$2 - \frac{N}{b} = 2 - \frac{N}{2N-1} = \frac{3N-2}{2N-1}.$$

Taking the reciprocal gives

$$\frac{2N-1}{3N-2} = \frac{2W(k)-3}{3W(k)-5},$$

which tends to $2/3$.

Declarations

Author responsibility and AI-assisted preparation. The manuscript was developed and typeset with substantial assistance from GPT-5.6 Pro, used in the drafting session under the "Sol Pro" reasoning configuration. The system assisted with proof exploration, consistency checks, literature organization, and LaTeX drafting. Tim Held is the sole human author, accepts responsibility for the contents, and is responsible for independently verifying every claim before formal submission or publication. The AI system is not an author.

Funding. No external funding was received for this work.

Competing interests. The author declares no competing interests.

Data and code availability. No external dataset, computer-assisted proof, or custom verification code is required for the argument. The numerical values in table 1 are direct evaluations of the displayed formulas using the exact van der Waerden numbers cited there. The complete LaTeX source accompanies this preprint.

References

[1] E. R. Berlekamp, A construction for partitions which avoid long arithmetic progressions, *Canadian Mathematical Bulletin* **11** (1968), no. 3, 409–414, doi:10.4153/CMB-1968-047-7.

[2] Erdős Problems, Problem 169, https://www.erdosproblems.com/169, accessed August 28, 2026.

[3] P. Erdős and R. L. Graham, Old and new problems and results in combinatorial number theory: van der Waerden's theorem and related topics, *L'Enseignement Mathématique* (2) **25** (1979), no. 3–4, 325–344.

[4] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monographies de L'Enseignement Mathématique, no. 28, Université de Genève, 1980.

[5] P. Erdős and P. Turán, On some sequences of integers, *Journal of the London Mathematical Society* **11** (1936), 261–264.

[6] J. L. Gerver, The sum of the reciprocals of a set of integers with no arithmetic progression of k terms, *Proceedings of the American Mathematical Society* **62** (1977), no. 2, 211–214.

[7] A. J. Kempner, A curious convergent series, *The American Mathematical Monthly* **21** (1914), no. 2, 48–50.

[8] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, A005346: van der Waerden numbers $W(2, n)$, https://oeis.org/A005346, accessed August 28, 2026.

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

[10] A. Walker, Integer sets of large harmonic sum which avoid long arithmetic progressions, arXiv:2203.06045v2 [math.NT], 2025, https://arxiv.org/abs/2203.06045.

[11] Aled Walker and Alexander Walker, Arithmetic progressions with restricted digits, *The American Mathematical Monthly* **127** (2020), no. 2, 140–150, doi:10.1080/00029890.2020.1682888.