International Journal of Mathematics and
Computer Science, **15**(2020), no. 2, 525–532

An Enumeration Problem Arising From a Result About Arithmetic Progressions

Julian Benali[1], **Harris Cobb**[2], **Peter Johnson**[3]

[1]Department of Mathematics
George Mason University
Fairfax, VA 22030, USA

[2]Department of Mathematics
Texas A & M University
College Station, TX 77843, USA

[3]Department of Mathematics and Statistics
Auburn University
Auburn, AL 36849, USA

email: johnspd@auburn.edu

(Received November 19, 2019, Accepted January 8, 2020)

Abstract

For positive integers m, n, q, $g(m, n, q)$ is the number of words of length n over an alphabet $A = \{a_1 \ldots, a_q\}$ such that there is no block of m consecutive a_q's in the word. A recent result shows that when q is a prime and $m \leq n$ then $g(m, n, q)$ is a lower estimate of the cardinality of the largest set of integers in $\{0, \ldots, q^n - 1\}$ which contains no q^m-term arithmetic progression. We give formulas for and estimates of $g(m, n, q)$ in special cases, and also a linear difference equation satisfied by $g(m, n, q)$ as a function of n.

1 Introduction

Suppose that $N \geq k \geq 3$ are integers, and let $[N] = \{0, \ldots, N - 1\}$. A *k-term arithmetic progression* is a set of integers expressible as

Key words and phrases: arithmetic progression, cyclic arithmetic progression, Van der Waerden hypergraph, chromatic number, vertex independence number, linear homogeneous difference equation, Fibonacci sequence.
This research was supported by NSF grant number 1560257.
AMS (MOS) Subject Classifications: 05A15, 05A16, 05C65, 05C69.
ISSN 1814-0432, 2020, http://ijmcs.future-in-tech.net

$\{a + jd \mid j = 0, \ldots, k-1\}$ for some integers $d > 0$ and a. The *Van der Waerden hypergraph* $W(k, N)$ has vertex set $[N]$ and for *hyperedges* (or just *edges*) all the k-term arithmetic progressions contained in $[N]$. The *cyclic Van der Waerden* hypergraph $W_c(k, N)$ is defined similarly, except that its edges are the *cyclic k-term arithmetic progressions mod N*; these are the k-element subsets of $[N]$ obtained by reducing mod N the integers in a k-term arithmetic progression. For example, $\{1, 6, 8\}$ is a cyclic 3-term arithmetic progression mod 9. Note that, for instance, the reductions mod 4 of the elements of the 3-term arithmetic progression $\{0, 2, 4\}$ form the set $\{0, 2\}$, which is not a cyclic 3-term arithmetic progression mod 4. Also note that every ordinary k-term arithmetic progression contained in $[N]$ is also a cyclic k-term arithmetic progression mod N. Therefore, $W(k, N)$ is a spanning subhypergraph of $W_c(k, N)$.

These hypergraphs are of interest not least because of a result that B. L. van der Waerden published in 1927 [5], well before hypergraphs became common objects of mathematical study. The result may be stated thus: For each $k \geq 3$, the *chromatic number* of $W(k, n)$, which is the minimum number of colors needed to color $[N]$ so that no k-term arithmetic progression in $[N]$ is monochromatic, tends to infinity as $N \to \infty$. Since $W(k, N)$ is a subhypergraph of $W_c(k, N)$, the same conclusion holds for $W_c(k, N)$.

Another hypergraph parameter closely related to the chromatic number is the (vertex) *independence number*. The independence number of $W(k, N)$ is the size of a largest subset of $[N]$ that does not contain any k-term arithmetic progression. Szemerédi's Lemma [4] implies that, if we let $\alpha(k, N)$ denote the independence number of $W(k, N)$, then $\alpha(k, N)/N \to 0$ as $N \to \infty$. Letting $\alpha_c(k, N)$ denote the independence number of $W_c(k, N)$, we have $\alpha_c(k, N) \leq \alpha(k, N)$, and so $\alpha_c(k, N)/N \to 0$ as $N \to \infty$.

Szemerédi's Lemma is stronger than van der Waerden's theorem: the later result implies the earlier, almost instantly. However, it appears that the actual behavior of $\alpha(k, N)$ as a function of N, for fixed $k \geq 3$, has not attracted anywhere near the interest focused on the chromatic number of $W(k, N)$. (The values of N at which the chromatic number increases are called *van der Waerden numbers*, and finding them is an obsession in some precincts.)

Suppose that $1 \leq m \leq n$ and p are integers, with p being a prime. When $p = 2$ we require that $m \geq 2$. It has recently been discovered [2] that every cyclic p^m-term arithmetic progression mod p^n must contain a term $t = \sum_{j=0}^{n-1} t_j p^j$, in which $t_j \in [p] = \{0, \ldots, p-1\}$, $j = 0, \ldots, n-1$, such that for some $0 \leq j \leq n - m$, $t_{j+i} = p - 1$, $i = 0, \ldots, m-1$. That is, the cyclic *a.p.* must contain an integer whose p-ary (base p) representation

$(t_{n-1}t_{n-2}\cdots t_0)_p$ contains a block of m consecutive $(p-1)$'s. It follows that $I(p^m,p^n) = \{t \in [p^n] |$ the p-ary representation of t does *not* contain a block of m consecutive $(p-1)$'s$\}$ is an independent set of vertices in both $W_c(p^m,p^n)$ and in $W(p^m,p^n)$. (It is also shown in [2] that $I(p^m,p^n)$ is a maximal independent set in both hypergraphs, but not necessarily maximum in $W(p^m,p^n)$.) Therefore,

$$|I(p^m,p^n)| \le \alpha_c(p^m,p^n) \le \alpha(p^m,p^n).$$

Consequently, evaluating $|I(p^m,p^n)|$ will give us some idea about the growth of $\alpha_c(p^m,p^n)$ and $\alpha(p^m,p^n)$ as $n \to \infty$ with m fixed. Indeed, without taking much trouble at all, we already have, for p an odd prime, $(p-1)^n = |I(p,p^n)| \le |I(p^m,p^n)| \le \alpha_c(p^m,p^n) \le \alpha(p^m,p^n) = o(p^n)$ as $n \to \infty$. The first equality is noted in [2] and the last is a consequence of Szemerédi's Lemma.

These bounds on the growth of $\alpha_c(p^m,p^n)$ and $\alpha(p^m,p^n)$, for p odd, are easy to remember, but we wonder if finer estimates are available.

2 A more general enumeration problem

For positive integers $q \ge 2$, m, and n let $A = \{a_1,\ldots,a_q\}$ be an alphabet with q elements, $G(m,n,q) = \{w = a_{i_1}\cdots a_{i_n} \in A^n |$ there is no block of m consecutive a_q's in the word $w\}$, and

$$g(m,n,q) = |G(m,n,q)|.$$

Observe that when p is a prime, $I(p^m,p^n) \simeq G(m,n,p)$ by the map that sends $z \in I(p^m,p^n)$ to its n-tuple of p-ary coefficients, $(c_0,\ldots,c_{n-1}) \in \{0,\ldots,p-1\}^n$, where $z = \sum_{i=0}^{n-1} c_i p^i$.

We aim to evaluate, or at least estimate, the numbers $g(m,n,q)$. From one point of view this is a foolish task to undertake, as the calculation of the $g(m,n,q)$ is but a single instance of a large class of enumeration problems completely and powerfully solved almost 40 years ago in [1]. However, the existence of a more general solution does not mean that there is nothing to be gained from focusing on a particular problem with elementary approaches. For instance, in [3], which is about the special cases when $q = 2$ of our problem, there are connections made with other parts of enumerative combinatorics that we will not touch on here. (However, see our Corollary 3.2, below.)

For another instance, we have in this section Corollary 2.6, which says that $g(m,n,q)/q^n \to 0$ as $n \to \infty$ (for $m \ge 1$, $q \ge 2$), which is a general-

ization of $|I(p^m, p^n)|/p^n \to 0$ as $n \to \infty$, derived at the end of the Introduction. That derivation depended on Szemerédi's Lemma, whereas the proof of Corollary 2.6 is quite elementary. Could we have proven Corollary 2.6 from the general theorem in [1]? No doubt, but there is value in elementary proofs.

We begin with the easy "boundary" values of $g(m, n, q)$. Proofs of the claims in Proposition 2.1 are left to the reader.

Proposition 2.1. *For all positive integers $q \geq 2$, m, and n:*

1. *If $m > n$ then $g(m, n, q) = q^n$.*

2. *$g(n, n, q) = q^n - 1$.*

3. *$g(1, n, q) = (q-1)^n$.*

Proposition 2.2. *For n, $q \geq 2$, $g(n-1, n, q) = q^n - 2q + 1$. If $q \geq 2$ and $m + 2 \leq n \leq 2m$ then $g(m, n, q) = q^n - (n-m+1)q^{n-m} + (n-m)q^{n-m-1}$.*

Remark 2.3. *We are aware that the result for $n = m + 1$ is given in the result for $m + 2 \leq n \leq 2m$; we decided to single out $g(m, m+1, q)$ anyway.*

Proof of Proposition 2.2 Let $H(m, n, q) = A^n \backslash G(m, n, q)$ and $h(m, n, q) = |H(m, n, q)| = q^n - g(m, n, q)$. Observe that $H(m, n, q)$ is the set of all words in A^n that do contain a block of m consecutive a_q's; that is; $H(m, n, q)$ consists of those $w \in A^n$ such that the maximum length of a subword of w consisting of consecutive a_q's is at least m.
If $m < n \leq 2m$ then there can be at most one maximal subword of $w \in A^n$ of length $\geq m$ consisting of consecutive a_q's. For each $b \in \{m, \ldots, n\}$, when $m < n \leq 2m$, it is straightforward to count the words $w \in H(m, n, q)$ in which the maximal such subword is of length $b \in \{m, \ldots, n\}$. When $b = n$ there is exactly one such word. When $b = n - 1$ there are two types of such words, xa_q^b and $a_q^b x$, $x \in A \backslash \{a_q\}$, so the number of such words is $2(q-1)$. When $m \leq b \leq n-2$ the words are in 3 categories: $a_q^b yv$, or uxa_a^b, in which x, $y \in A \backslash \{a_q\}$ and u, $v \in A^{n-b-1}$, or $uxa_q^b yv$, with x, $y \in A \backslash \{a_q\}$ and u, v are possibly empty words such that $uv \in A^{n-b-2}$. The total number of such words, when $m \leq b \leq n-2$, is $2(q-1)q^{n-b-1} + (q-1)^2(n-b-1)q^{n-b-2}$. Summing over b we have, when $m + 2 \leq n \leq 2m$,

$$\begin{aligned}
h(m,n,q) &= 1 + 2(q-1) + \sum_{b=m}^{n-2} [2(q-1)q^{n-b-1} + (q-1)^2(n-b-1)q^{n-b-2}] \\
&= 2q - 1 + 2(q-1) \sum_{j=1}^{n-m-1} q^j + (q-1)^2 \sum_{j=1}^{n-m-1} jq^{j-1} \\
&= 2q - 1 + 2(q-1)q\frac{q^{n-m-1}-1}{q-1} + (q-1)^2 \frac{d}{dq} \sum_{j=0}^{n-m-1} q^j \\
&= 2q - 1 + 2q(q^{n-m-1}-1) + (n-m-1)q^{n-m} - (n-m)q^{n-m-1} + 1 \\
&= (n-m+1)q^{n-m} - (n-m)q^{n-m-1}
\end{aligned}$$

The conclusion now follows from

$$g(m,n,q) = q^n - h(m,n,q).$$

□

Convention: $g(m,0,q) = 1$ for all positive integers $q \geq 2$ and m. (The only element of $G(m,0,q)$ is the empty word.)

Lemma 2.4. *For integers $q \geq 2$, $m \geq 1$, n, a, $b \geq 0$ such that $a + b = n$, $g(m,n,q) \leq g(m,a,q)\, g(m,b,q)$.*

Proof. Clearly $G(m,n,q) \subseteq \{w = xy \mid x \in G(m,a,q), y \in G(m,b,q)\}$. □

Theorem 2.5. *For integers $q \geq 2$ and $n > m > 0$,*
$\max\,[(q-1)^n, q^n - (n-m+1)q^{n-m}] \leq g(m,n,q)$. If, in addition, $n > 2m$, then

$$g(m,n,q) \leq (q^{2m} - (m+1)q^m + mq^{m-1})^{\lfloor \frac{n}{2m} \rfloor} g(m, n - 2m\lfloor \frac{n}{2m} \rfloor, q).$$

Proof. Using notation from the proof of Proposition 2.2, clearly $H(m,n,q) \subseteq H(m-1,n,q)$ if $1 < m$. Therefore $h(m,n,q)$ is non-increasing with m, so $g(m,n,q)$ is non-decreasing with m. Therefore, $g(m,n,q) \geq g(1,n,q) = (q-1)^n$.

There are $n - m + 1$ blocks of m consecutive indices in $\{1,\ldots,n\}$, and for each such block B there are q^{n-m} words $w \in A^n$ with a_q's in every position indicated by indices in B. Since every $w \in H(m,n,q)$ has a_q's in positions indicated by indices in at least one of those blocks, we have

$q^n - g(m,n,q) = h(m,n,q) \leq (n-m+1)q^{n-m}$, so
$q^n - (n-m+1)q^{n-m} \leq g(m,n,q)$.

Now suppose that $n > 2m$ and let z and r be integers such that $n = 2mz + r$ and $0 \leq r < 2m$; that is, $z = \lfloor \frac{n}{2m} \rfloor$ and $r = n - 2mz$. Applying Lemma 2.4 z times, we have $g(m,n,q) \leq g(m,2m,q)^z g(m,r,q)$, which, in view of Proposition 2.2, is the final conclusion of the theorem. □

The lower bound on $g(m,n,q)$ in Theorem 2.5 is underwhelming, in view of the fact that

$$q^n - (n-m+1)q^{n-m} = q^n[1 - \frac{n-m+1}{q^m}] \to -\infty \text{ as } n \to \infty,$$

for fixed m and q. However, for large n the lower bound of $(q-1)^n$ is not bad, in view of the upper bound.

That upper bound can be more explicit, since for $0 \leq r < 2m$ explicit formulae for $g(m,r,q)$ are available in Propositions 2.1 and 2.2. Also, the upper bound is obviously crude–the inequalities at each iteration of Lemma 2.4 in the proof are strict. However, the upper bound is good enough to verify what one would expect from the discussion in the Introduction, where it is shown that the following holds whenever q is a prime.

Corollary 2.6. *For all integers $m \geq 1$ and $q \geq 2$, $\frac{g(m,n,q)}{q^n} \to 0$ as $n \to \infty$.*

Proof. Suppose that $m \geq 1$ and $n > 2m$. Let $n = 2mz + r$, $0 \leq r < 2m$, as in the proof of Theorem 2.5. Then

$$\begin{aligned} \frac{g(m,n,q)}{q^n} &\leq \frac{(q^{2m} - (m+1)q^m + mq^{m-1})^z}{q^{2mz}} \frac{g(m,r,q)}{q^r} \\ &= (1 - \frac{(m+1)q-m}{q^{m+1}})^z \frac{g(m,r,q)}{q^r} \to 0 \end{aligned}$$

as $z \to \infty$. □

3 A linear homogeneous difference equation for $g(m,n,q)$

Let us suppose that $n \geq m$. Every $w \in G(m,n,q)$ will either end with a letter $x \in A \backslash \{a_q\}$, or xa_q, or xa_qa_q (if $m > 2$), etc. That is to say, $w = uxa_q^k$ for some $k \in \{0, \ldots, m-1\}$, $x \in A \backslash \{a_q\}$ and $u \in G(m, n-k-1, q)$. Letting $f(n) = g(m,n,q)$, for fixed m and q, we therefore have that $f(n) = (q-1)f(n-1) + (q-1)f(n-2) + \cdots + (q-1)f(n-m) = (q-1)\sum_{k=1}^{m} f(n-k)$. Hence, we have

Theorem 3.1. *For all integers $m \geq 1$, $q \geq 2$, $f(n) = g(m, n, q)$ is the unique solution of the order m linear homogeneous difference equation $f(n) = (q-1)\sum_{k=1}^{m} f(n-k)$ satisfying the initial conditions $f(n) = q^n$, $n = 0, \ldots, m-1$.*

The Fibonacci sequence $F_0, F_1, F_2, \ldots$ is defined by $F_0 = 0$, $F_1 = 1$, and, for $n > 1$, $F_n = F_{n-1} + F_{n-2}$.

Corollary 3.2. *For all $n = 0, 1, 2, \ldots, g(2, n, 2) = F_{n+2}$.*

Proof. By Theorem 3.1 and Proposition 2.1, $g(2, 0, 2) = 1 = F_2$, $g(2, 1, 2) = 2 = F_3$, and for $n > 1$, $g(2, n, 2) = (2-1)[g(2, n-1, 2) + g(2, n-2, 2)]$. □

There is a routine for solving linear homogeneouos difference equations with initial conditions. We have only to solve the characteristic equation of the difference equation, which is

$$r^m = (q-1)(r^{m-1} + \cdots + 1)$$

in this case, write the general solution of the difference equation as a linear combination of m special solutions associated with the solutions of the polynomial characteristic equation, and then determine which coefficents in the linear combination will give a solution that satisfies the initial conditions. The main impediment to the execution of this plan is the difficulty of extracting solutions of the characteristic equation. When $m = 2$ the equation is quadratic. We leave it to the reader to verify the claim of the following corollary.

Corollary 3.3. *For integers $q \geq 2$ and $n \geq 0$, $g(2, n, q) =$*
$\frac{(q-1)^{n/2}}{2^{n+1}(\sqrt{(q-1)(q+3)}}[(\sqrt{(q-1)(q+3)}+q+1)(\sqrt{q-1}+\sqrt{q+3})^n+(\sqrt{(q-1)(q-3)}-$
$q-1)(\sqrt{q-1}-\sqrt{q+3})^n]$

References

[1] L. J. Guibas, A. M. Odlyzko, String overlaps, pattern matching, and nontransitive games, *Journal of Combinatorial Theory*, Series A, **30,** (1981), 183–208.

[2] Peter Johnson, Zechun Yang, On the independence numbers of the Van der Waerden and cyclic Van der Waerden hypergraphs $W(k,t)$ and $W_c(k,t)$ when t and k are powers of the same prime, *Geombinatorics,* **28,** (April 2019), 191–200.

[3] M. A. Nyblom, Counting binary strings without r-runs of ones, *International Mathematical Forum,***7,** no. 38, (2012), 1865–1876.

[4] Endre Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arithmetica,* **27,** (1975), 199–254.

[5] B. L. van der Waerden, Beweis einer baudetschen vermutung, *Nieuw Arch. Wisk.,* **15,** (1927), 212–216.