

In Section 2 we characterize the property " $\gamma(a) \geq r$ " by r linear equations in a_0, \dots, a_n . Then we exhibit a family of vectors with gap one and note that only an exponentially small fraction of vectors has positive gap. The upper bound $\Gamma(n) < n/2$ is trivial. We show that $\Gamma(n) = 0$ if $n+1$ is prime; using a bound on the gap between consecutive primes, we then obtain $\Gamma(n) = O(n^{-548})$ in general. We conjecture that, in fact, $\Gamma(n) = O(1)$.

In Section 3 we introduce the notion of a *folded vector*. The characterization of the gap then leads to several infinite families of vectors, again with gap one, via the solution of certain Diophantine equations. In Section 4 we extend and combine these examples to obtain families with gap two or three, and then give further examples with gap one; the latter do not, however, give any new information about Γ . In Section 5 we report on a computer search which determined all vectors with positive gap for $n \leq 128$; the largest gap is 3.

This research was motivated by work of Nisan and Szegedy [3]. They investigate the degree of polynomials in $\mathbf{R}[x_1, \dots, x_n]$ that interpolate (or approximate) a given Boolean function $g: \{0, 1\}^n \rightarrow \{0, 1\}$. The smallest degree of such interpolating polynomials is the *Fourier degree* of g . If g is symmetric, there is an associated function $f: \{0, \dots, n\} \rightarrow \{0, 1\}$ whose interpolation problem is equivalent to the original one. Bounds on $\Gamma(n)$ are thus equivalent to bounds on the Fourier degree of symmetric Boolean functions.

2. Bounds on the Gap

We begin by recalling a few basic facts from the theory of difference equations; see Graham, Knuth, and Patashnik [1] for a more complete discussion. Given $g \in \mathbf{R}[x]$, we define its *discrete derivative* Dg by

$$(Dg)(x) = (D^1g)(x) = g(x) - g(x-1).$$

For $i \geq 2$, we define the discrete derivative $D^i g$ of order i inductively by

$$D^i g = D(D^{i-1}g).$$

Proposition 2.1. *For $g \in \mathbf{R}[x]$ and $m \geq 1$, the following hold:*

- (i) $(D^m g)(x) = \sum_{0 \leq j \leq m} (-1)^j \binom{m}{j} g(x-j)$.
- (ii) If $\deg(g) = m$, then $\deg(Dg) = m-1$.
- (iii) If g is constant, then $Dg = 0$.
- (iv) If $\deg(g) = m$, then $D^m g$ is a nonzero constant.
- (v) $D^m g = 0 \iff \deg(g) < m$.

Proof. (i) follows by induction on m ; (ii) and (iii) follow immediately when g is written out as a sum of monomials; and (iv) and (v) follow from (ii) and (iii). ■

Now we characterize the gap $\gamma(a)$ in terms of binomial sums.

Theorem 2.2. Let $n \geq 1$, $a = (a_0, \dots, a_n) \in \mathbf{Q}^{n+1}$, and let $f \in \mathbf{Q}[x]$ be the unique interpolating polynomial with $\deg(f) \leq n$ and $f(j) = a_j$ for $0 \leq j \leq n$. Then for $0 \leq r \leq n$, the following are equivalent:

- (i) $\gamma(a) \geq r$,
- (ii) $\deg(f) \leq n - r$,
- (iii) For $n - r < m \leq n$, we have

$$(2.1) \quad \sum_{0 \leq j \leq m} (-1)^j \binom{m}{j} a_j = (-1)^m (D^m f)(m) = 0.$$

Proof. We recall that $\gamma(a) = n - \deg(f)$, note that the first equation in (2.1) holds by Proposition 2.1(i), and use induction on r . The case $r=0$ is trivial, and the case $r=1$ follows from Proposition 2.1(v).

For the induction step, we assume that the theorem holds for $r=t-1$, where $t \geq 2$, and show that the theorem follows for $r=t$. By definition, f is the minimal interpolating polynomial for a_0, \dots, a_n . We note that

$$\deg(f) \leq n - t + 1 \iff f \text{ is also the (unique) minimal} \\ \text{interpolant for } a_0, \dots, a_{n-t+1}.$$

Now using the case $r=1$ for interpolating at $0, \dots, n-t+1$ and the induction hypothesis for interpolating at $0, \dots, n$, we find that

$$\begin{aligned} \deg(f) \leq n - t &\iff \deg(f) \leq n - t + 1 \text{ and} \\ &\quad (2.1) \text{ holds for } m = n - t + 1 \\ &\iff (2.1) \text{ holds for } n - t < m \leq n. \end{aligned}$$

The following theorem demonstrates that $\Gamma(n) \geq 1$ for all odd $n \geq 3$.

Theorem 2.3. Let $n \geq 3$ be odd and $a \in A_n$ with $a_j = a_{n-j}$ for $0 \leq j \leq n$. Then

$$\gamma(a) \geq 1.$$

Proof. This follows from Theorem 2.2, with $r=1$.

Observation 2.4. Since the sum in (2.1) is linear in the a_j 's, it is clear that any two (or more) solution vectors can be added componentwise to yield another vector with positive gap. The only restriction is that the new vector must also have its components in $\{0, 1\}$. Thus the set of $2^{(n+1)/2} - 2$ nontrivial vectors a in Theorem 2.3 can be described as the set of all admissible sums of the $(n+1)/2$ "basis" vectors

$$(1, 0, 0, \dots, 0, 0, 1), (0, 1, 0, 0, \dots, 0, 0, 1, 0), \dots, (0, \dots, 0, 1, 1, 0, \dots, 0).$$

Although there are many vectors with positive gap, we now show that most choices of a have gap equal to zero.

Theorem 2.5. For $n \geq 2$ and randomly chosen $a \in A_n$,

$$\text{Prob} \{ \gamma(a) \geq 1 \} < 2^{-n/3}.$$

Proof. Let $s = \lfloor n/3 \rfloor + 1$. With arbitrary values a_j for $s \leq j \leq n$, we show that there is at most one choice of the remaining a_j 's that satisfies the condition in Theorem 2.2 for $r=1$. This will follow because the sequence of binomial coefficients is superincreasing in an appropriate sense. Since $s \geq (n+1)/3$, the desired probability is at most

$$\frac{2^{n+1-s}}{\#A_n} \leq \frac{2^{(2n+2)/3}}{2^{n+1}-2} = 2^{-n/3} \frac{2^{2/3}}{2-2^{-n+1}} < 2^{-n/3}$$

for $n \geq 3$. For $n=2$, the probability is 0, by Table 3.

So now suppose that there are two distinct vectors a and b such that $a_j = b_j$ for $n/3 < j \leq n$ and

$$\sum_{0 \leq j \leq n} (-1)^j \binom{n}{j} a_j = \sum_{0 \leq j \leq n} (-1)^j \binom{n}{j} b_j = 0.$$

Let m be the maximum value of j for which $a_j \neq b_j$, so that $m \leq n/3$. Then

$$\sum_{0 \leq j \leq m} (-1)^j \binom{n}{j} (a_j - b_j) = 0.$$

For each j , let $u_j = (-1)^j (a_j - b_j) / (a_m - b_m)$. Then

$$u_j \in \{-1, 0, 1\}, \quad u_m = (-1)^m, \quad \text{and} \quad \sum_{0 \leq j \leq m} \binom{n}{j} u_j = 0.$$

Thus

$$(2.2) \quad \binom{n}{m} = (-1)^{m+1} \sum_{0 \leq j < m} \binom{n}{j} u_j \leq \sum_{0 \leq j < m} \binom{n}{j}.$$

But for $0 \leq j < m$, we have

$$\binom{n}{j} \binom{n}{j+1}^{-1} = \frac{j+1}{n-j} \leq \frac{m}{n-m+1} \leq \frac{n/3}{\frac{2n}{3}+1} < \frac{1}{2}.$$

Therefore

$$\sum_{0 \leq j < m} \binom{n}{j} < \binom{n}{m} \sum_{0 \leq j < m} (1/2)^{m-j} < \binom{n}{m} \cdot 1,$$

contradicting (2.2). ■

For $n \geq 2$, we have the following trivial bounds on the maximal gap:

$$0 \leq \Gamma(n) \leq \frac{n-1}{2},$$

since for each $a \in A_n$, with interpolating polynomial $f \in \mathbf{Q}[x]$, either f or $f-1$ has at least $(n+1)/2$ zeros in $\{0, 1, \dots, n\}$; since f is not constant, we have $\deg(f) = \deg(f-1) \geq (n+1)/2$. In the following theorems we obtain better upper bounds on $\Gamma(n)$.

Theorem 2.6. *If $n+1$ is prime, then $\Gamma(n) = 0$.*

Proof. By Theorem 2.2, $\Gamma(n) > 0$ if and only if

$$(2.3) \quad \sum_{0 \leq j \leq n} (-1)^j \binom{n}{j} a_j = 0$$

for some $a \in A_n$. Writing $p = n+1$ and using the fact that

$$\binom{p-1}{j} \equiv \frac{(-1)(-2)\cdots(-j)}{1 \cdot 2 \cdots j} \equiv (-1)^j \pmod{p}$$

for $0 < j < p$, we obtain

$$\sum_{0 \leq j < p} (-1)^j \binom{p-1}{j} a_j \equiv \sum_{0 \leq j < p} 1 \cdot a_j \pmod{p}.$$

This is congruent to 0 mod p only if the a_j 's are all 0 or all 1. Since these two cases are excluded (they give constant interpolating polynomials), (2.3) cannot be satisfied. Thus $\Gamma(n) = 0$. ■

Together with Theorem 2.6, any upper bound on the gaps between consecutive prime numbers gives an upper bound on Γ .

Theorem 2.7. $\Gamma(n) = O(n^{.548})$.

Proof. From Theorems 2.6 and 2.2 it follows that $\Gamma(n) \leq n - (p-1)$, where p is the largest prime less than or equal to $n+1$. By Mozzochi's [2] theorem on prime number gaps, we have $p \geq n - O(n^{.548})$, and thus $\Gamma(n) = O(n^{.548})$. ■

The table of prime numbers gives $m = \max\{\Gamma(n) : 1 \leq n \leq 128\} \leq 13$, while the true value is $m = 3$, by Theorem 5.2. Indeed, we conjecture that the true upper bound is a constant.

Szegedy observed that one can apply the results of Nisan and Szegedy [3] to obtain the following.

Theorem 2.8. *Let $g: \{0, 1\}^n \rightarrow \{0, 1\}$ be a nonconstant symmetric Boolean function of n variables, d its Fourier degree, and $a = (a_0, \dots, a_n) \in A_n$ with $a_i = g(1^i 0^{n-i})$ for $0 \leq i \leq n$. Then $d = n - \gamma(a)$, and hence*

$$n - O(n^{.548}) \leq n - \Gamma(n) \leq d \leq n. \quad \blacksquare$$

3. Folded Vectors and Examples with Gap One

In this section and the next we exhibit several infinite families of vectors with positive gap. Three of the families have gap two, and two families have gap three.

Definition 3.1. We have four natural operations on A_n , namely, for $a = (a_0, \dots, a_n) \in A_n$:

- the reverse of a : $\bar{a} = (a_n, \dots, a_0) \in A_n$,
- the complement of a : $e_n - a = (1 - a_0, \dots, 1 - a_n) \in A_n$,
- the prefix of a : $\pi a = (a_0, \dots, a_{n-1}) \in A_{n-1} \cup \{(0, \dots, 0), (1, \dots, 1)\}$,
- the suffix of a : $\sigma a = (a_1, \dots, a_n) \in A_{n-1} \cup \{(0, \dots, 0), (1, \dots, 1)\}$,

where $e_n = (1, \dots, 1) \in \mathbf{N}^{n+1}$. If $f \in \mathbf{Q}[x]$ interpolates a , then $f(n-x)$ and $1-f$ interpolate \bar{a} and $e_n - a$, respectively. In particular, $\gamma(a) = \gamma(\bar{a}) = \gamma(e_n - a)$.

Because of the symmetry

$$(3.1) \quad \binom{n}{j} = \binom{n}{n-j},$$

it is convenient to "fold" each vector $a \in A_n$ back on itself and consider the shortened vector $b = (b_0, \dots, b_s) = \varphi(a)$ defined below.

Definition 3.2. For $n \geq 2$ and $a \in A_n$, we define the *folded vector* $b = (b_0, \dots, b_s) = \varphi(a)$ as follows:

$$\varphi(a) = \begin{cases} (a_0 - a_n, a_1 - a_{n-1}, \dots, a_{(n-1)/2} - a_{(n+1)/2}) & \text{if } n \text{ is odd,} \\ (a_0 + a_n, a_1 + a_{n-1}, \dots, a_{(n-2)/2} + a_{(n+2)/2}) & \text{if } n \text{ is even,} \end{cases}$$

where $s = \lfloor (n-1)/2 \rfloor$. We denote this mapping by the *folding operator* $\varphi: A_n \rightarrow B_n \subset \mathbf{Z}^{s+1}$, where

$$B_n = \varphi(A_n) = \begin{cases} \{-1, 0, 1\}^{(n+1)/2} & \text{if } n \text{ is odd,} \\ \{0, 1, 2\}^{n/2} & \text{if } n \text{ is even.} \end{cases}$$

Observation 3.3. If n is even, we have an additional component $a_{n/2}$ in a , which does not appear in $\varphi(a)$. However, we may usually assume without loss of generality that $a_{n/2} = 0$. Otherwise, the complementary vector $e_n - a$ has $(e_n - a)_{n/2} = 0$ and $\varphi(e_n - a) = 2e_{(n-2)/2} - \varphi(a)$.

In this notation, Theorem 2.2 with $r=1$ says that $\gamma(a) \geq 1$ for an $a \in A_n$ (with $a_{n/2} = 0$ if n is even) if and only if

$$(3.2) \quad \sum_{0 \leq j \leq s} (-1)^j \binom{n}{j} b_j = 0,$$

where $b = \varphi(a)$ and $s = \lfloor (n-1)/2 \rfloor$. This is the property that we will work with in the sequel, and if $b \in B_n$ satisfies (3.2), we say that b has *positive gap*. In particular, if $b \in B_n$ has positive gap, then $\gamma(a) \geq 1$ for every $a \in \varphi^{-1}(\{b\})$. (If n is even, we require $a_{n/2} = 0$.) On the other hand, for any $a \in A_n$ with $\gamma(a) > 0$, either $b = \varphi(a)$ has positive gap or, if n is even and $a_{n/2} = 1$, $b = \varphi(e_n - a)$ has positive gap.

Observation 3.4. If n is odd, $a \in A_n$, and $b = \varphi(a) \in B_n$, then $-b = \varphi(\bar{a}) = \varphi(e_n - a) \in B_n$. Thus if n is odd, b has positive gap if and only if $-b$ does.

Observation 3.5. For $i \in \{0, 1\}$ and $b \in B_n$, we let $w_i(b)$ be the number of components in b that equal i . If n is odd, then $\#\varphi^{-1}(\{b\}) = 2^{w_0(b)}$, and if n is even, then $\#\varphi^{-1}(\{b\}) = 2^{w_1(b)+1}$, because $a_{n/2}$ can be 0 or 1. (3.2) will guide our search for nontrivial gaps. However, the actual gap $\gamma(a)$ cannot be determined from $\varphi(a)$ alone. This is illustrated by the vectors $a^{(1)}, a^{(2)}, a^{(3)}$, defined after Theorem 4.2 below, and πa from Theorem 4.2, which have the same folded vector but gaps 2, 1, 1, 2, respectively. Given $b \in B_n$, it is easy to construct $\varphi^{-1}(\{b\})$. This may, however, be a large set, and working with the single folded vector b is advantageous, both conceptually (see the theorems below) and computationally (see Section 5). Note in particular that

$$\#B_n \approx 3^{n/2} \ll 2^n \approx \#A_n.$$

We now use our “folded vector” notation to describe several more infinite families of vectors with gap one. Our basic idea is to take $b \in B_n$ with $b_j = 0$ for all but a small number of consecutive values of j . Then (3.2) is equivalent to a Diophantine equation in n and k of small degree whose solutions we can determine. Our examples in Section 4 of gap two and three are based on these families.

In the simplest nontrivial case, we have $b_j = 0$ for all values of j but k and $k+1$, where $0 \leq k \leq s-1 = \lfloor (n-1)/2 \rfloor - 1$. Then we can rewrite (3.2) as follows:

$$\begin{aligned} 0 &= \sum_{0 \leq j \leq s} (-1)^j \binom{n}{j} b_j = (-1)^k \binom{n}{k} b_k + (-1)^{k+1} \binom{n}{k+1} b_{k+1} \\ &= (-1)^{k+1} \frac{n!}{(k+1)!(n-k)!} ((n-k)b_{k+1} - (k+1)b_k). \end{aligned}$$

If n is odd, then $b_k, b_{k+1} \in \{-1, 1\}$, and the only possible solution has $b_k = b_{k+1}$ and $n = 2k + 1$. But this violates the requirement that $k \leq s - 1$.

If n is even, then $b_k, b_{k+1} \in \{1, 2\}$. But $k+1 \leq s$, so that $\binom{n}{k} < \binom{n}{k+1}$ and the only possible solution is $b_k = 2, b_{k+1} = 1$. We find all solutions to this equation in the next theorem.

Theorem 3.6. Let n be even, $0 \leq k \leq (n-4)/2$, and $b \in B_n$ with $b_k = 2, b_{k+1} = 1$, and $b_j = 0$ otherwise. Then b has positive gap if and only if

$$n = 6t + 2 \text{ and } k = 2t \text{ for some } t \geq 1.$$

Proof. By the above, it is sufficient to find all solutions to the equation

$$2(k+1) = n - k.$$

Then $n = 3k + 2$. Since n is even, we have $k = 2t$ and $n = 6t + 2$ for some $t \geq 0$. This solution satisfies $k \leq (n-4)/2$ if and only if $t \geq 1$. ■

Corollary 3.7. For each $b \in B_n$ as in Theorem 3.6, we have $\#\varphi^{-1}(\{b\}) = 4$, and for each of the two $a \in \varphi^{-1}(\{b\})$ with $a_{n/2} = 0$, we have $\gamma(a) = 1$.

Proof. Observation 3.5 implies the first claim. It is straightforward to check that neither of the two vectors $a \in \varphi^{-1}(\{b\})$ with $a_{n/2} = 0$ satisfies (2.1) with $r = 2$; thus $\gamma(a) < 2$. ■

If we allow $b \in B_n$ to have *three* consecutive nonzero components, we obtain a second-order Diophantine equation in n and k . So suppose that all components of b except b_{k-1} , b_k , b_{k+1} are zero. Then (3.2) is equivalent to

$$(-1)^{k-1} \binom{n}{k-1} b_{k-1} + (-1)^k \binom{n}{k} b_k + (-1)^{k+1} \binom{n}{k+1} b_{k+1} = 0,$$

or, equivalently,

$$(n-k+1)(n-k)b_{k-1} - (n-k+1)(k+1)b_k + (k+1)k b_{k+1} = 0.$$

Writing $m = n - k$, (3.2) is equivalent to

$$(3.3) \quad (m+1)m b_{k-1} - (m+1)(k+1) b_k + (k+1)k b_{k+1} = 0.$$

Because $k < m$, we can quickly rule out most values of (b_{k-1}, b_k, b_{k+1}) as solutions. We find that if n is even, then (b_{k-1}, b_k, b_{k+1}) equals $(1, 2, 1)$, and that if n is odd, then it equals $(-1, 1, 1)$ or $(1, -1, -1)$. The two possibilities for odd n correspond to pairs of unfolded vectors in A_n that are reverses of each other, as discussed in Observation 3.4. The complete solutions are given in the next two theorems.

Theorem 3.8. Let n be even, $1 \leq k \leq (n-4)/2$, and $b \in B_n$ with $b_{k-1} = b_{k+1} = 1$, $b_k = 2$, and $b_j = 0$ otherwise. Then b has positive gap if and only if

$$n = 4t^2 - 2 \quad \text{and} \quad k = 2t^2 - t - 1 \quad \text{for some} \quad t \geq 2.$$

Proof. Letting $m = n - k$ and using (3.3), we have to find all integral solutions with $1 \leq k < m$ of the equation

$$(m+1)m - 2(m+1)(k+1) + (k+1)k = 0.$$

Solving for m , we obtain

$$m = \frac{(2k+1) \pm \sqrt{8k+9}}{2}.$$

Thus $8k+9=u^2$ for some $u \in \mathbf{N}$, from which it follows that

$$m = \frac{u^2 \pm 4u - 5}{8}.$$

Since $m \in \mathbf{N}$, it follows that u is odd, say $u=2v-1$ for some $v \in \mathbf{N}$. Then

$$k = \frac{v^2 - v - 2}{2} \quad \text{and} \quad m = \frac{(v^2 - v - 1) \pm (2v - 1)}{2}.$$

Since $m > k$, we have

$$m = \frac{(v^2 - v - 1) + (2v - 1)}{2} = \frac{v^2 + v - 2}{2}.$$

Then $n = k + m = v^2 - 2$. Since n is even, we have $v = 2t$ for some $t \in \mathbf{N}$. Thus we obtain

$$n = 4t^2 - 2, \quad k = 2t^2 - t - 1, \quad m = n - k = 2t^2 + t - 1.$$

One checks that this solution is valid for all $t \geq 2$, but not for $t = 1$. ■

The Fibonacci numbers are given by $F_0 = 0$, $F_1 = 1$, and $F_i = F_{i-1} + F_{i-2}$ for $i \geq 2$.

Theorem 3.9. *Let n be odd, $1 \leq k \leq (n-3)/2$, and $b \in B_n$ with $b_{k-1} = -1$, $b_k = b_{k+1} = 1$, and $b_j = 0$ otherwise. Then b has positive gap if and only if*

$$n = F_{2i+2}F_{2i+3} - 1 \quad \text{and} \quad k = F_{2i}F_{2i+3} \quad \text{for some } i \geq 2 \text{ with } i \not\equiv 1 \pmod{3}.$$

Proof. We seek all solutions to the equation

$$\binom{n}{k+1} = \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

This problem has been solved by several authors using standard techniques for Pell equations; see, e.g., Singmaster [4] and Tovey [6]. In Singmaster [4] it is shown that (n, k) is a solution if and only if

$$n = F_{2i+2}F_{2i+3} - 1 \quad \text{and} \quad k = F_{2i}F_{2i+3} \quad \text{for some } i \geq 0.$$

In addition, we require that n be odd and that $k \geq 1$, so $i \geq 1$ and either F_{2i+2} or F_{2i+3} is even. Since F_j is even if and only if $j \equiv 0 \pmod{3}$, it follows that the general solution is valid for our purposes if and only if $i \geq 1$ and $i \not\equiv 1 \pmod{3}$. ■

Theorem 3.9 shows that $\Gamma(n) \geq 1$ for $n = 103, 731, \dots$. However, we already know from Theorem 2.3 that $\Gamma(n) \geq 1$ for all odd $n \geq 3$.

So far we have considered the first- and second-order Diophantine equations that arise if all but two or three consecutive components of the folded vector b are zero. We could continue in this vein by letting four consecutive components be nonzero, but third-order Diophantine equations do not usually yield infinite families of solutions. One exception is given by the following theorem.

Theorem 3.10. *Let n be odd, $2 \leq k \leq (n-3)/2$, and $b \in B_n$ with $b_{k-2} = b_{k-1} = 1$, $b_k = b_{k+1} = -1$, and $b_j = 0$ otherwise. Then b has positive gap if and only if*

$$n = 4t^2 - 3 \quad \text{and} \quad k = 2t^2 - t - 1 \quad \text{for some } t \geq 2.$$

Proof. We have

$$\begin{aligned} (-1)^k \sum_{0 \leq j \leq (n-1)/2} (-1)^j \binom{n}{j} b_j &= \binom{n}{k-2} - \binom{n}{k-1} - \binom{n}{k} + \binom{n}{k+1} \\ &= \binom{n}{k-2} + \binom{n}{k-1} - 2 \left(\binom{n}{k-1} + \binom{n}{k} \right) + \binom{n}{k} + \binom{n}{k+1} \\ &= \binom{n+1}{k-1} - 2 \binom{n+1}{k} + \binom{n+1}{k+1}. \end{aligned}$$

By Theorem 3.8, this expression is zero only if

$$n+1 = 4t^2 - 2 \quad \text{and} \quad k = 2t^2 - t - 1 \quad \text{for some } t \geq 1.$$

Moreover, this solution is valid for all $t \geq 2$. ■

4. Families with Gap Two or Three

In this section we use Theorem 2.2 with $r \geq 2$ to find families with gap two or three. First we convert the condition in Theorem 2.2 into a more convenient condition.

We claim that condition (iii) in Theorem 2.2 entails the following relations:

$$(4.1) \quad \forall m, i \quad n-r < m \leq n, \quad 0 \leq i \leq n-m \quad \sum_{0 \leq j \leq m} (-1)^j \binom{m}{j} a_{j+i} = 0,$$

which are easily proved by induction, using Pascal's triangle. (4.1), together with Theorem 2.2, implies the following propagation of the gap to and from prefix and suffix vectors, which will be useful in establishing the theorems of this section.

Proposition 4.1. *Let $n \geq 2, r \geq 2$ and $a \in \mathbf{Q}^{n+1}$.*

- (i) $\gamma(a) \geq r \iff \gamma((a_0, \dots, a_{n-i})) \geq r-i$ for $0 \leq i \leq r$.
- (ii) $\gamma(a) \geq r \iff \gamma(a) \geq 1$ and $\gamma(\pi a) \geq r-1$.
- (iii) $\gamma(a) \geq r \iff \gamma(a) \geq 1$ and $\gamma(\sigma a) \geq r-1$.
- (iv) $\gamma(a) \geq r \iff \gamma(\pi a) \geq r-1$ and $\gamma(\sigma a) \geq r-1$.
- (v) $\Gamma(n-1) \geq \Gamma(n) - 1$.

Proof. We denote by $\Sigma(m, a)$ the sum in (2.1). For $0 \leq i < r$ and $n-r < m \leq n-i$, we have $\Sigma(m, a) = \Sigma(m, (a_0, \dots, a_{n-i}))$. Then (i) and (ii) follow from Theorem 2.2.

Next, (iii) follows from (ii) applied to \bar{a} , since $\gamma(a) = \gamma(\bar{a})$ and $\gamma(\sigma a) = \gamma(\overline{\sigma a}) = \gamma(\pi \bar{a})$. In (iv), " \implies " follows from (ii) and (iii). For " \impliedby ", we have a unique polynomial of degree at most $n-2$ interpolating a_1, \dots, a_{n-1} ; by assumption, it also interpolates a_0 and a_n . For (v), let $a \in A_n$ with $\Gamma(n) = \gamma(a) = r$. Then $\pi a \in A_{n-1}$ or $\sigma a \in A_{n-1}$, and in either case $\Gamma(n-1) \geq r-1$, by (ii) or (iii). ■

The propagation in (v) is illustrated in Table 3 for $n = 15, 35, 63, 99, 105$.

Proposition 4.1 (v) says that for gap two we need positive gaps for two consecutive values of n . This occurs, e.g., in Theorems 3.8 and 3.10. The next theorem gives the gaps of the corresponding vectors. We use the notation introduced in Definition 3.1 to represent prefix and suffix vectors.

Theorem 4.2. *Let $t \geq 2$, $n = 4t^2 - 1$, $k = 2t^2 - t - 1$, and $a \in A_n$ with $a_k = a_{k+1} = a_{n-k-1} = a_{n-k} = 1$, and $a_j = 0$ otherwise. Then*

- (i) $\gamma(a) = 3$,
- (ii) $\gamma(\pi a) = 2$,
- (iii) $\gamma(\pi \pi a) = 1$.

Proof. Let $a' = \pi \pi a = (a_0, \dots, a_{n-2}) \in A_{n-2}$. Then $b = -\varphi(a') \in B_{n-2}$ satisfies Theorem 3.10, and thus $\gamma(a') \geq 1$. Proposition 4.1 (ii) and Theorem 3.8 imply the lower bound on $\gamma(\pi a)$ in (ii). By Theorem 2.3, we have $\gamma(a) \geq 1$, and hence $\gamma(a) \geq 3$, using the lower bound in (ii) and Proposition 4.1 (ii). Finally, the sum (2.1) for $m = n - 3$ is

$$(-1)^k \cdot \left(-\binom{n-3}{k-3} + \binom{n-3}{k-2} + \binom{n-3}{k} - \binom{n-3}{k+1} \right).$$

Substituting for n and k , one finds that the sum is nonzero for all t ; this implies all the upper bounds, by Proposition 4.1 (ii). ■

This approach suggests considering some other vectors, with $n = 4t^2 - 2$ and $k = 2t^2 - t - 1$:

$$\begin{aligned} a^{(1)} &= (\dots \underset{k-1}{1} \ \underset{k}{1} \ \underset{k+1}{0} \ \dots \ \underset{n-k-1}{1} \ \underset{n-k}{1} \ \underset{n-k+1}{0} \ \dots) \in A_n, \\ a^{(2)} &= (\dots \ 0 \ \underset{k}{1} \ \underset{k+1}{0} \ \dots \ \underset{n-k-1}{1} \ \underset{n-k}{1} \ \underset{n-k+1}{1} \ \dots) \in A_n, \\ a^{(3)} &= (\dots \ \underset{k-1}{1} \ \underset{k}{1} \ \underset{k+1}{1} \ \dots \ \underset{n-k-1}{0} \ \underset{n-k}{1} \ \underset{n-k+1}{0} \ \dots) \in A_n, \\ a^{(4)} &= (\dots \ \underset{k-1}{1} \ \underset{k}{1} \ \underset{k+1}{0} \ \dots \ \underset{n-k-1}{1} \ \underset{n-k}{1} \ \underset{n-k+1}{0} \ \dots \ 0) \in A_{n+1}, \\ a^{(5)} &= (\dots \ \underset{k-1}{0} \ \underset{k}{1} \ \underset{k+1}{1} \ \dots \ \underset{n-k-1}{0} \ \underset{n-k}{1} \ \underset{n-k+1}{1} \ \dots \ 0 \ 0) \in A_{n+2}, \end{aligned}$$

Now $a^{(1)} = \pi a^{(4)}$, $a^{(2)} = \overline{a^{(3)}}$, and $a = \pi a^{(5)}$ is as in Theorem 4.2. Then one checks, using the method of Theorem 4.2 and additional calculations for the last four upper bounds, that

$$\gamma(a^{(1)}) = 2, \ \gamma(a^{(2)}) = \gamma(a^{(3)}) = 1, \ \gamma(a^{(4)}) = \gamma(a^{(5)}) = 0.$$

Just as we found vectors with gap three from the basic solution of Theorem 3.8, we now extend Theorem 3.9 in the same way.

From Proposition 4.1 (iv) it follows that in order to have $\gamma(a) \geq 2$, we must have $\gamma(\pi a) \geq 1$ and $\gamma(\sigma a) \geq 1$. One can show that there is essentially just one way to choose $\pi a, \sigma a$ satisfying Theorem 3.9 and/or Theorem 2.3 to create $a \in A_{n+1}$ with $\gamma(a) \geq 2$. This solution is given in the following theorem.

The proof below will use the fact that alternating sums of binomial coefficients can be collapsed as follows, for any $0 \leq k \leq \ell \leq n$ with $n \geq 1$:

$$(4.2) \quad \sum_{k \leq j \leq \ell} (-1)^j \binom{n}{j} = \sum_{k \leq j \leq \ell} (-1)^j \left(\binom{n-1}{j-1} + \binom{n-1}{j} \right) = (-1)^k \binom{n-1}{k-1} + (-1)^\ell \binom{n-1}{\ell}.$$

As usual, $\binom{n-1}{-1}$ and $\binom{n-1}{n}$ have to be interpreted as zero, and $\binom{0}{0}$ as one.

Theorem 4.3. *Let $i \geq 2$ with $i \not\equiv 1 \pmod 3$, $n = F_{2i+2}F_{2i+3} + 1$, $k = F_{2i}F_{2i+3}$, and let $a \in A_n$ be given by*

$$a = \begin{matrix} \{ (0, 1, 0, 1, \dots, 1, 0, 1, 1, 0, 0, \dots, 0, 0, & 1, & 1, & 0, 1, \dots, 1, 0), \\ (1, 0, 1, 0, \dots, 1, 0, 1, 1, 0, 0, \dots, 0, 0, & 1, & 1, & 0, 1, \dots, 0, 1), \\ \underbrace{\hspace{1.5cm}}_0 & \underbrace{\hspace{1.5cm}}_{k \ k+1} & \underbrace{\hspace{1.5cm}}_{n-k-1 \ n-k} & \underbrace{\hspace{1.5cm}}_n \end{matrix}$$

where the top line is used for odd k , and the bottom one for even k . Then n is odd, and the following hold.

- (i) $\gamma(a) = 3$.
- (ii) $\gamma(\pi a) = \gamma(\sigma a) = 2$.

Proof. Both a and $\sigma \pi a = \pi \sigma a$ are symmetric, hence have gap at least one by Theorem 2.3. If k is odd, then $b = \varphi(\sigma \sigma a)$ satisfies Theorem 3.9, so that $\gamma(\sigma \sigma a) \geq 1$ as well, and then $\gamma(\sigma a) \geq 2$ and $\gamma(a) \geq 3$, by Proposition 4.1 (iii) and (iv). If k is even, then $b = \varphi(\pi \pi a)$ satisfies Theorem 3.9, so that $\gamma(\pi a) \geq 2$ and $\gamma(a) \geq 3$, again by Proposition 4.1. This shows all lower bounds.

Finally, it is sufficient to show that $\gamma(\sigma \pi \pi a) = 0$; this implies all claimed upper bounds by Proposition 4.1 (iv). For odd k , Theorem 2.2 implies that $\gamma(\sigma \pi \pi a) > 0$ if and only if the following sum is nonzero:

$$\begin{aligned} \sum_{0 \leq j \leq n-3} (-1)^j \binom{n-3}{j} a_{j+1} &= \sum_{0 \leq j \leq k} (-1)^j \binom{n-3}{j} + \binom{n-3}{k-1} \\ &= \binom{n-4}{k-2} + \binom{n-4}{k-1} - \binom{n-4}{k}, \end{aligned}$$

using (3.1) and (4.2). The last sum is nonzero by Theorem 3.9.

A similar calculation establishes the theorem for even k . ■

$\Gamma(n)$	n	first five values	reference
≥ 1	$2t - 1$	3,5,7,9,11	Thm 2.3
≥ 1	$6t - 4$	8, 14, 20, 26, 32	Thm 3.6
≥ 2	$4t^2 - 2$	14, 34, 62, 98, 142	Thm 4.2 (ii)
≥ 3	$4t^2 - 1$	15, 35, 63, 99, 143	Thm 4.2 (i)
≥ 2	$F_{2i+2}F_{2i+3}$	104, 714, 33552, 229970, 10803704	Thm 4.3 (ii)
≥ 3	$F_{2i+2}F_{2i+3} + 1$	105, 715, 33553, 229971, 10803705	Thm 4.3 (i)

Table 1. Lower bounds on Γ .

We summarize our main results so far, namely the lower bounds on $\Gamma(n)$, in Table 1, where $t \geq 2$, and $i \geq 2$ satisfies $i \not\equiv 1 \pmod 3$.

We conclude this section with three more examples of positive gap which, however, do not increase the lower bounds on Γ that we already proved. For brevity's sake, the proofs are left out. Two key ideas used in discovering the examples are the telescoping trick in (4.1) and the fact that folded solution vectors b can be combined linearly to form new folded solution vectors.

Theorem 4.4. (i) Let $r \geq 2$ be even,

$$n = n_r = \frac{4 + 3\sqrt{2}}{8}(3 + 2\sqrt{2})^r + \frac{4 - 3\sqrt{2}}{8}(3 - 2\sqrt{2})^r,$$

$$k = k_r = \frac{2 + \sqrt{2}}{8}(3 + 2\sqrt{2})^r + \frac{2 - \sqrt{2}}{8}(3 - 2\sqrt{2})^r - \frac{1}{2},$$

and $b \in B_n$ with $b_j = 1$ for $0 \leq j < k$, $b_k = b_{k+1} = -1$, and $b_j = 0$ otherwise. Then $n, k \in \mathbb{N}$, n is odd, and b has positive gap.

(ii) Let $u \geq 4$ with $u \not\equiv 0 \pmod 3$, $n = 4u^2 - 2$, $k = (4u^2 - 4)/3$, $m = 2u^2 - u - 1$, and $b \in B_n$ with $b_k = b_m = 2$, $b_{k+1} = b_{m-1} = b_{m+1} = 1$, and $b_j = 0$ otherwise. Then b has positive gap.

(iii) Let $r, n = n_r$, and $k = k_r$ be as in (i), with $r \equiv 2 \pmod 4$; let $\ell = (n - 2)/3$; and let $b \in B_n$ be defined by

$$b_j = \begin{cases} 1 & \text{if } j \in \{0, \dots, k - 1\} \setminus \{\ell, \ell + 1\}, \\ 0 & \text{if } j \in \{\ell + 1\} \cup \{k + 2, \dots, (n - 1)/2\}, \\ -1 & \text{if } j \in \{\ell, k, k + 1\}. \end{cases}$$

Then $n \equiv 5 \pmod 6$, and b has positive gap.

Remark. The vectors given in Theorem 4.4(iii) have $n = 35, 40389, \dots$, where each value of n is about 577 times as great as the previous value. The example with $n = 35$ can be found in the computer-generated list of folded vectors with positive gap.

5. Computer-Generated Vectors with Positive Gap

For any given $n \in \mathbf{N}$ we can, in principle, apply Theorem 2.2 to all $2^{n+1} - 2$ vectors $a \in A_n$ to find those with positive gap. It is more efficient, however, to do our exhaustive search over *folded* vectors, of which there are only about $3^{n/2}$. The unfolded vectors corresponding to folded vectors with positive gap can then be examined individually to find those with gap equal to two, three, or more.

The search over folded vectors can itself be simplified considerably by making use of the superincreasing nature of the sequence of binomial coefficients $\binom{n}{j}$ for sufficiently small j , much as in Theorem 2.5. This is illustrated by the adaptive algorithm below, for odd n . A similar algorithm works for even n .

Algorithm 5.1. Suppose that n is odd. We seek all vectors

$$b = (b_0, \dots, b_{(n-1)/2}) \in \{-1, 0, 1\}^{(n+1)/2}$$

such that (3.2) holds.

We choose the components in reverse order, beginning with $b_{(n-1)/2}$, which multiplies the largest binomial coefficient in the sum. After choosing each new component b_k , we observe that

$$\begin{aligned} \sum_{0 \leq j \leq (n-1)/2} (-1)^j \binom{n}{j} b_j = 0 &\iff \sum_{k \leq j \leq (n-1)/2} (-1)^j b_j \binom{n}{j} = - \sum_{0 \leq j < k} (-1)^j b_j \binom{n}{j} \\ &\implies \left| \sum_{k \leq j \leq (n-1)/2} (-1)^j b_j \binom{n}{j} \right| \leq \sum_{0 \leq j < k} \binom{n}{j}. \end{aligned}$$

If the last inequality is violated, then our choice of $(b_k, \dots, b_{(n-1)/2})$ is inconsistent, so we can prune 3^k vectors from the search tree and try a different value of b_k . The search ends when we have either pruned or explicitly tried all $3^{(n+1)/2}$ vectors $b \in \{-1, 0, 1\}^{(n+1)/2}$.

Our empirical evidence indicates that the algorithm above requires about $2^{n/4}$ steps to test all $b \in B_n$. A Fortran implementation running for a week stealing unused machine cycles on a 4-processor 33-Megahertz MIPS R3000 system found all folded vectors with positive gap for $2 \leq n \leq 128$. In addition to the 98 solutions obtained from the theorems of the previous sections, we found 29 new solutions.

In Table 2 we list all folded vectors with positive gap for $2 \leq n \leq 128$ found by an exhaustive computer search. To save space, we omit the 63 all-zero vectors, which are solutions for all odd n . We also omit the 21 vectors $(0, \dots, 0, 2, 1, 0, \dots, 0)$,

which occur for all n of the form $6t+2$. Thus of the 98 previously obtained solutions, only 14 are recorded in the table.

For even n , the components are written in the order $b_0, b_1, \dots, b_{(n-2)/2}$. For odd n , we represent -1 and $+1$ by “ $-$ ” and “ $+$ ”, respectively. If $b \in B_n$ for odd n , then $-b \in B_n$ has the same gap, as discussed in Observation 3.4. We list just one of the two in the table.

After each vector given by a previous theorem we write “T” followed by the theorem number. In addition, there are 21 essentially different new solution vectors, which we do not know how to fit into a family, plus 8 other solutions that can be written as simple linear combinations of other solutions.

In Table 3, we list $\Gamma(n)$ for $1 \leq n \leq 128$. We omit the sometimes lengthy calculations that justify these values and yield the following theorem.

Theorem 5.2. *The maximum value of $\Gamma(n)$ for $2 \leq n \leq 128$ is 3.*

Question 1. What is the order of $\Gamma(n)$?

Conjecture. The maximum value of $\Gamma(n)$ for $n \in \mathbb{N}$ is 3.

Question 2. A generalization of our problem is to consider some set $K \subseteq \mathbb{Q}$ and $f \in \mathbb{Q}[x]$ of degree at most n with $\{f(0), \dots, f(n)\} \subseteq K$. What bounds can we give on $\Gamma_K(n) = \max\{n - \deg(f)\}$?

Remark. For $k = \#K \geq 2$, the trivial bound is $\Gamma_K(n) < \frac{k-1}{k} n$.

Some results of this paper are valid more generally than stated; for example, the characterizations of Section 2 obviously hold for fields of characteristic zero or $p > n$. Theorem 2.6 then implies that every two-valued polynomial of degree less than p over a finite prime field \mathbb{F}_p has maximal degree, namely $p-1$. The trace from \mathbb{F}_{2^n} to \mathbb{F}_2 shows that this is not true for all finite fields.

Question 3. Which degrees less than $q-1$ can two-valued polynomials over a finite field \mathbb{F}_q have?

Acknowledgements Thanks go to Andrew Odlyzko for providing the reference on the difference between consecutive primes. Special thanks go to Mario Szegedy, who posed the problem considered in this paper and who greeted our results with gratifying enthusiasm. He also noted Theorem 2.8.

The first author gratefully acknowledges the hospitality of Gaston Gonnet at ETH Zürich, Switzerland, where part of this work was done. It was partially supported by the Natural Sciences and Engineering Research Council of Canada, and the Information Technology Research Centre, Ontario. We are grateful to an anonymous referee who suggested the use of discrete derivatives in Theorem 2.2 and other improvements.

n	$\Gamma(n)$	n	$\Gamma(n)$	n	$\Gamma(n)$	n	$\Gamma(n)$
1	0	33	1	65	1	97	1
2	0	34	2	66	0	98	2
3	1	35	3	67	1	99	3
4	0	36	0	68	1	100	0
5	1	37	1	69	1	101	1
6	0	38	1	70	0	102	0
7	1	39	1	71	1	103	1
8	1	40	0	72	0	104	2
9	1	41	1	73	1	105	3
10	0	42	0	74	1	106	0
11	1	43	1	75	1	107	1
12	0	44	1	76	0	108	0
13	1	45	1	77	1	109	1
14	2	46	0	78	0	110	1
15	3	47	1	79	1	111	1
16	0	48	1	80	1	112	0
17	1	49	1	81	1	113	1
18	0	50	1	82	0	114	0
19	1	51	1	83	1	115	1
20	1	52	0	84	0	116	1
21	1	53	1	85	1	117	1
22	0	54	1	86	1	118	0
23	1	55	1	87	1	119	1
24	1	56	1	88	0	120	0
25	1	57	1	89	1	121	1
26	1	58	0	90	0	122	1
27	1	59	1	91	1	123	1
28	0	60	0	92	1	124	0
29	1	61	1	93	1	125	1
30	0	62	2	94	0	126	0
31	1	63	3	95	1	127	1
32	1	64	0	96	0	128	1

Table 3. Maximal gaps for $n \leq 128$

References

- [1] R. L. GRAHAM, D. E. KNUTH, and O. PATASHNIK: *Concrete Mathematics*, Addison-Wesley, New York, 1989.
- [2] C. J. MOZZOCHI: On the difference between consecutive primes, *J. Number Theory*, **24** (1986), 181–187.
- [3] N. NISAN and M. SZEGEDY: On the degree of Boolean functions as real polynomials, *STOC*, 1992, Victoria, 462–474.
- [4] D. SINGMASTER: Repeated binomial coefficients and Fibonacci numbers, *The Fibonacci Quarterly*, **13** (1975), 295–298.
- [5] B. M. STEWART: *Theory of Numbers*, 2nd edition, The Macmillan Co., New York, 1965.
- [6] C. A. TOVEY: Multiple occurrences of binomial coefficients, *The Fibonacci Quarterly*, **23** (1985), 356–358.

Joachim von zur Gathen

*FB Mathematik-Informatik,
Universität-GH Paderborn
D-33095 Paderborn, Germany
gathen@uni-paderborn.de*

James R. Roche

*Center for Communications Research
Thanet Road, Princeton NJ 08540, USA
roche@ccr-p.ida.org*