PROGRESS TOWARDS A NONINTEGRALITY CONJECTURE

SHANTA LAISHRAM, DANIEL LÓPEZ-AGUAYO, CARL POMERANCE AND
THOTSAPHON THONGJUNTHUG

ABSTRACT. Given $r \in \mathbb{N}$, define the function $S_r : \mathbb{N} \to \mathbb{Q}$ by

$$S_r(n) = \sum_{k=0}^{n} \frac{k}{k+r} \binom{n}{k}.$$

In 2015, the second author conjectured that there are infinitely many $r \in \mathbb{N}$ such that $S_r(n)$ is nonintegral for all $n \geq 1$, and proved that $S_r(n)$ is not an integer for $r \in \{2, 3, 4\}$ and for all $n \geq 1$. In 2016, Florian Luca and the second author raised the stronger conjecture that for any $r \geq 1$, $S_r(n)$ is nonintegral for all $n \geq 1$. They proved that $S_r(n)$ is nonintegral for $r \in \{5, 6\}$ and that $S_r(n)$ is not an integer for any $r \geq 2$ and $1 \leq n \leq r - 1$. In particular, for all $r \geq 2$, $S_r(n)$ is nonintegral for at least $r - 1$ values of n. In 2018, the fourth author gave sufficient conditions for the nonintegrality of $S_r(n)$ for all $n \geq 1$, and derived an algorithm to sometimes determine such nonintegrality; along the way he proved that $S_r(n)$ is nonintegral for $r \in \{7, 8, 9, 10\}$ and for all $n \geq 1$. By improving this algorithm we prove the conjecture for $r \leq 22$. Our principal result is that $S_r(n)$ is usually nonintegral in that the upper asymptotic density of the set of integers n with $S_r(n)$ integral decays faster than any fixed power of r^{-1} as r grows.

1. INTRODUCTION

In 2014, Marcel Chiriță [2] asked to show that $\sum_{k=0}^{n} \frac{k}{k+1} \binom{n}{k}$ is nonintegral for all integers $n \geq 1$. This is true and one can prove it as follows: the given sum is equal to $2^n - \frac{2^{n+1}-1}{n+1}$ and $\frac{2^{n+1}-1}{n+1}$ is never an integer due to the fact that for every integer $r \geq 2$, $2^r \not\equiv 1 \pmod{r}$. Given $r \in \mathbb{N}$, define the function $S_r : \mathbb{N} \to \mathbb{Q}$ by $S_r(n) = \sum_{k=0}^{n} \frac{k}{k+r} \binom{n}{k}$. Motivated by [2], the second author [4] raised the question whether there are infinitely many $r \in \mathbb{N}$ such that $S_r(n)$ is nonintegral for all $n \geq 1$, and proved that $S_r(n)$ is not an integer for $r \in \{2, 3, 4\}$ and for all $n \geq 1$. These results also hinge on the fact that for $r \geq 2$, $2^r \not\equiv 1 \pmod{r}$.

In 2016, Florian Luca and the second author [5] conjectured that for any $r \geq 2$, $S_r(n)$ is always nonintegral for all $n \geq 1$. They proved that $S_r(n)$ is nonintegral for $r \in \{5, 6\}$ and that $S_r(n)$ is not an integer for any $r \geq 2$ and $1 \leq n \leq r - 1$. In particular, for all $r \geq 2$, $S_r(n)$ is nonintegral for at least $r-1$ values of n. The proof of this fact relies heavily on the following theorem of Sylvester [3]: every product of

k consecutive integers larger than k is divisible by a prime larger than k. In 2018, the fourth author [9] gave sufficient conditions for the nonintegrality of $S_r(n)$ for all $n \geq 1$, and derived an algorithm to sometimes determine such nonintegrality; along the way he proved that $S_r(n)$ is nonintegral for $r \in \{7, 8, 9, 10\}$ and for all $n \geq 1$.

Our first result is a simplified version of [9, Theorem 4.1].

Theorem 1.1. *Given an integer $r \geq 2$, the sum $S_r(n)$ is nonintegral for all $n \in \mathbb{N}$ if the following condition holds. With m_r the product of all primes up to r, each integer $n \in \{1, \ldots, m_r\}$ satisfies at least one of the following:*
  (1) *There exists $i \in \{1, \ldots, r\}$ such that $\gcd(n + i, m_r) = 1$.*
  (2) *There exist $i, j \in \{1, \ldots, r\}$ such that $\gcd(n + i, m_r) = \gcd(n + j, m_r) = 2$ and $0 < |i - j| < 8$.*

We discuss the more complicated version of this result from [9] in Section 5. With our simpler criterion, plus some other ideas presented below, we are able to prove the following result.

Theorem 1.2. *For $1 \leq r \leq 22$ we have $S_r(n)$ nonintegral for all $n \in \mathbb{N}$.*

We also prove the following result.

Theorem 1.3. *For each $k > 0$ there exists a constant $c_k > 0$, such that for each integer $r > 1$, the upper asymptotic density of $\{n \in \mathbb{N} : S_r(n) \in \mathbb{N}\}$ is at most c_k/r^k.*

The proof uses a theorem by Montgomery-Vaughan [6].

This paper is organized as follows. In Section 2 we give the necessary preliminaries. In Section 3 we prove Theorems 1.1 and 1.2, in Section 4 we prove 1.3, and in Section 5 we discuss the original version of Theorem 1.1 from [9].

2. PRELIMINARIES

Throughout this section, for a prime p and an integer u with $p \nmid u$, we shall let $\mathrm{ord}_p u$ denote the least positive integer k such that $u^k \equiv 1 \pmod{p}$.

Lemma 2.1. *Suppose that p is an odd prime dividing $n \in \mathbb{N}$ and $p \mid 2^n - 1$. If a is the largest divisor of n composed of primes smaller than p, we have $p \mid 2^a - 1$.*

Proof. Since $\mathrm{ord}_p 2 \mid p - 1$, all of the primes dividing $\mathrm{ord}_p 2$ are smaller than p. If $p \mid 2^n - 1$, then $\mathrm{ord}_p 2 \mid n$, so that $\mathrm{ord}_p 2 \mid a$. $\qquad\square$

Let ϕ denote Euler's function from elementary number theory.

Definition 2.2. Let $q \in \mathbb{N}$ and let $\alpha \geq 1$ be a real number. Following [6], we define
$$V_\alpha(q) = \sum_{i=1}^{\phi(q)} (a_{i+1} - a_i)^\alpha$$
where $1 = a_1 < a_2 \cdots < a_{\phi(q)+1}$ are the integers in $[1, q+1]$ that are coprime to q.

The following proposition is crucial for the proof of our main theorem.

Proposition 2.3. ([6, Corollary 1]) *Let $q \in \mathbb{N}$. For any fixed real number $\alpha \geq 1$, there is a positive number $c(\alpha)$ such that*

$$V_\alpha(q) \leq c(\alpha)\phi(q)(\phi(q)/q)^{-\alpha}.$$

We have the following inequality, which follows from [8, (3.30)] and a short calculation:

$$\frac{n!}{\phi(n!)} < 3\ln n, \quad \text{for all } n \geq 2. \tag{1}$$

Following [5], we define

$$S(r,n) := \sum_{k=0}^{n} \frac{r}{k+r}\binom{n}{k}.$$

It is clear that

$$S_r(n) + S(r,n) = \sum_{k=0}^{n}\binom{n}{k} = 2^n, \tag{2}$$

so that $S_r(n)$ is integral if and only if $S(r,n)$ is integral.

The following result is shown in [5].

Lemma 2.4. *We have*

$$S(r,n) = \sum_{j=1}^{r}(-1)^{r-j}r\binom{r-1}{j-1}\frac{2^{n+j}-1}{n+j}.$$

3. The search to $r = 22$

Proposition 3.1. *Let $n, r \in \mathbb{N}$ and suppose for some integer $j \in \{1, \ldots, r\}$ we have $n + j = ab$ where $a, b \in \mathbb{N}$, $b > 1$, such that*

- *each prime dividing b is greater than r,*
- *and each prime dividing $2^a - 1$ is at most r.*

Then $S(r,n)$ is nonintegral.

Proof. Let p be the least prime factor of b, so that $p > r$. Note that $p \mid n+j$, but p does not divide any other member of $\{n+1, \ldots, n+r\}$. Suppose that $p \mid 2^{n+j}-1$. By Lemma 2.1, $p \mid 2^a - 1$. But by assumption, all prime factors of $2^a - 1$ are at most r, a contradiction. Thus, in lowest terms, the fraction $(2^{n+j}-1)/(n+j)$ has at least one factor p in the denominator. The term corresponding to j in Lemma 2.4 is, up to sign,

$$r\binom{r-1}{j-1}\frac{2^{n+j}-1}{n+j}.$$

So, since $p > r$, we see that this term, when reduced to its lowest terms, has at least one factor p in the denominator. However, no other term in the sum in Lemma 2.4 has a factor p in the denominator, so that in the full sum $S(r,n)$, there is a factor p in the denominator. That is, $S(r,n)$ is not an integer. This completes the proof. $\square$

Remark 1. We note that for every integer $a \geq 2$, there is a prime $p \equiv 1 \pmod{a}$ that divides $2^a - 1$, see [1]. This implies that with a, b as in Proposition 3.1 we have $a \leq r - 1$, so that a and b are coprime.

Proof of Theorem 1.1. If $\{n + 1, \ldots, n + r\}$ contains some $n + i$ coprime to m_r, then we can apply Proposition 3.1 to $n + i$ with $b = n + i$ and $a = 1$. On the other hand, if $\{n + 1, \ldots, n + r\}$ contains two even numbers less than 8 apart that are not divisible by any odd prime up to r, at least one of them is not divisible by 8, say it is $n + j$. We apply Proposition 3.1 with $a = 2$ or 4 and $b = (n + j)/a$. Since we may assume that $n \geq r > 6$ from [5], Proposition 3.1 applies.   □

Using Theorem 1.1 and with some additional help from Proposition 3.1 we can prove the conjecture for $r \leq 22$.

Proof of Theorem 1.2. We first handle the cases $r = 11, 12$. The product of the primes to 11 is $m_{11} = 2310$, so it suffices to show that every interval of 11 consecutive integers either contains a member coprime to 2310 or contains two even members less than 8 apart that are coprime to 1155. Since the problem is symmetric about 1155, we only need to search to this level. There are precisely 7 intervals of 11 consecutive integers in this range which do not contain a number coprime to 2310; these are the intervals starting at

$$2, 114, 115, 116, 200, 468, 510.$$

We check that in each of them there are two even numbers less than 8 apart which are coprime to 1155.

The calculation for $r = 13, 14, 15, 16$ is somewhat more extensive. Here we show that every interval of 13 consecutive integers contains either one that is coprime to $m_{13} = 30030$ or contains two even members less than 8 apart that are coprime to 15015. We need only check intervals whose first element is in $[1, 15015]$. All but 76 of them have a member coprime to 30030. Each of these 76 intervals contains two even members less than 8 apart that are coprime to 15015.

This plan breaks down for $r = 17, 18$. For example, when $n = 60462$, the interval $[n, n + 16]$ has each member with a nontrivial gcd with $m_{17} = 510510$, so that condition (1) does not apply. In addition, the only members of this set with no odd prime factors at most 17 are $n + 2$ and $n + 10$. So, condition (2) does not apply either. However, for $r \geq 17$, we can strengthen (2) to

(2') There exist $i, j \in \{1, \ldots, r\}$ such that $0 < |i - j| < 16$ and both $n + i$ and $n + j$ are even but not divisible by any odd prime up to r.

In addition, we have found it easier at higher levels to use Proposition 3.1 directly for those intervals that do not have a member coprime to m_r. When $r = 17$, there are 498 such intervals $[n, n + 16]$ below 255255. For each such interval I, we examine I translated by $j \times 510510$ for $j = 0, 1, 2, 3$, searching in each for a member of the form ab where the primes in b are greater than r, and with $a = 2$ or 4. All but two intervals had this property, namely the length 17 intervals

starting at $n = 60462$ and at $n = 97590$ shifted by 3×510510. However, in these intervals, property (2') applies.

In continuing on to $r = 19, 20, 21, 22$ it turns out that conditions (1) and (2') are not sufficient for all cases. We can supplement with a new condition which works for $r \geq 13$:

(3) There exist $i, j \in \{1, \ldots, r\}$ such that $9 \nmid j - i$, both $n + i$ and $n + j$ are multiples of 3, they are not divisible by any prime in $[5, r]$, and they are not divisible by 8.

The sufficiency of condition (3) follows from Poposition 3.1 with $a \mid 12$, using that the largest prime factor of $2^{12} - 1$ is 13. Condition (3) works well in conjunction with condition (2') since to apply them one needs to to translate the interval by jm_r, for $j = 0, 1, 2, 3$. In examining length 19 intervals, all but 8439 of them satisfy condition (1). Looking up to 4 times m_{19}, all but a handful of these 4×8439 intervals satisfy the hypothesis of Proposition 3.1 with $a = 2$ or 4. This handful is settled using conditions (2') and (3). This completes the proof. $\square$

Remark 2. One possible route to proving that $S_r(n)$ is always nonintegral is to show that one of (1), (2) in Theorem 1.1 always occurs. In fact, in the next section we show that condition (1) holds when r is large for most of the intervals $\{n + 1, \ldots, n + r\}$. However, there are exceptional intervals where (1) does not hold, and we have already seen that there can be intervals where neither (1) nor (2) hold. In addition, one can show that for all sufficiently large numbers r there is some n such that $\{n + 1, \ldots, n + r\}$ has each member with an odd prime factor at most r. For example, a short argument shows this is the case for $r = 103$. So, replacing condition (2') with higher powers of 2 does not always work either. It is conceivable that for every r and every interval of r consecutive integers at least r there is a member for which the hypothesis of Proposition 3.1 holds, but we are not sure if this is so. Complicating things, one has for all sufficiently large numbers r an interval of r consecutive integers each divisible by a prime p in the range $\log_2 r < p \leq r$, see [7, equation (3)].

4. Density

In order to prove Theorem 1.3, we first show that condition (1) from Theorem 1.1 usually holds.

Theorem 4.1. *For each $k > 0$ there exists a constant $c_k > 0$, such that for each integer $r > 1$, the asymptotic density of those integers n such that $\{n+1, \ldots, n+r\}$ contains no number coprime to m_r is at most c_k/r^k.*

Proof. Let $\alpha \geq 1$ be a real number to be determined. Recall that

$$V_\alpha(q) = \sum_{i=1}^{\phi(q)} (a_{i+1} - a_i)^\alpha$$

where $1 = a_1 < a_2 \cdots < a_{\phi(q)+1}$ are the integers in $[1, q+1]$ that are coprime to q. By Proposition 2.3, it follows that

$$V_\alpha(q) < c(\alpha)\phi(q)(\phi(q)/q)^{-\alpha} = c(\alpha)q(q/\phi(q))^{\alpha-1}$$

for some constant $c(\alpha) > 0$. Applying this with $q = m_r$, together with (1), yields

$$V_\alpha(q) < c(\alpha)q(3\ln r)^{\alpha-1}. \tag{3}$$

Let $N := \sum_{\{i:\ a_{i+1}-a_i \geq r\}} (a_{i+1} - a_i)$. Then

$$Nr^{\alpha-1} = \sum_{\{i:\ a_{i+1}-a_i \geq r\}} (a_{i+1} - a_i)r^{\alpha-1} \leq \sum_{i=1}^{\phi(q)} (a_{i+1} - a_i)^\alpha = V_\alpha(q),$$

so that $N \leq V_\alpha(q)/r^{\alpha-1}$. Using inequality (3) we obtain

$$N \leq c(\alpha)q\left(\frac{3\ln r}{r}\right)^{\alpha-1}. \tag{4}$$

Now, we note that if an interval $I \subset [1, m_r]$ of integers does not contain any member of $\{a_1, \ldots, a_{\phi(q)+1}\}$, then since it is an interval, it must lie completely between two consecutive members of this set. So, if $\{n+1, \ldots, n+r\}$ contains no number coprime to $m_r = q$, there exists $w \in [1, \phi(q)]$ such that $(n, n+r+1) \subseteq (a_w, a_{w+1})$. Then n may be any of the numbers $a_w, a_w+1, \ldots, a_{w+1}-r-1$, that is the interval (a_w, a_{w+1}) gives rise to exactly $a_{w+1}-a_w-r$ intervals $\{n+1, \ldots, n+r\}$. Therefore

$$\#\{1 \leq n \leq q : \gcd(n+i, q) \neq 1 \text{ for } i = 1, \ldots, r\} \leq \sum_{\{i:a_{i+1}-a_i>r\}} (a_{i+1} - a_i - r) \leq N.$$

It remains to note that the integers n where $\{n+1, \ldots, n+r\}$ has no element coprime to m_r form a periodic set mod m_r. That is, if n has this property, so does every positive integer $m \equiv n \pmod{r!}$. Hence by (4) the density of the set of such numbers is at most $c(\alpha)\left(\frac{3\ln r}{r}\right)^{\alpha-1}$. Then let $\alpha = k+2$, and the result follows with c_k the maximal value of $c(\alpha)(3\ln r)^{\alpha-1}/r$. $\square$

Theorem 1.3 now follows as a corollary.

Proof of Theorem 1.3. It follows from Theorem 1.1 that if some member of $\{n+1, \ldots, n+r\}$ is coprime to m_r, then $S_r(n)$ is nonintegral. Thus, the theorem follows immediately from Theorem 4.1. $\square$

5. THONGJUNTHUG'S THEOREM

In [9], the fourth author proved the following theorem.

Theorem 5.1. ([9, Theorem 4.1]) *Given an integer $r \geq 5$, the sum $S_r(n)$ is not a positive integer for all $n \geq 2$ if the following two conditions hold:*
(a) *For all $l \in \{1, 2, \ldots, r\}$, we have*

$$r(F_r(n) - (-1)^{r-1}(r-1)!) \equiv \pm r!(2^{n+l} - 1) \pmod{(n+l)}$$

where $F_r(n) = \sum_{i=0}^{r-1} \sum_{k=0}^{n+r} s(r,i+1)k^i \binom{n+r}{k}$, *and* $s(r,i+1)$ *is the signed Stirling number of the first kind.*

(b) *Each integer* $n \in \{0, 1, \ldots, P-1\}$, *where* P *is the product of all primes up to* r, *satisfies at least one of the following:*

(b1) *There exists* $i \in \{1, \ldots, r\}$ *such that* $p \nmid (n+i)$ *for all primes* $p \leq r$.

(b2) *There exist* $i, j \in \{1, \ldots, r\}$ *such that* $0 < |i - j| < 8$ *and both* $n+i$ *and* $n+j$ *are even but not divisible by any odd prime up to* r.

We have seen in Theorem 1.1 that this result holds without condition (a), so that condition is superfluous. However, condition (a) is harmless, in that it always holds. We now prove this assertion.

Proposition 5.2. *The condition (a) in Theorem 5.1 holds for all* $r, n \in \mathbb{N}$.

Proof. Using [9, Lemma 3.2] and (2) one has

$$r(F_r(n) - (-1)^{r-1}(r-1)!) = (n+1) \cdots (n+r)S(r,n)$$

$$= \sum_{i=1}^{r} (-1)^{r-i} r \binom{r-1}{i-1}(2^{n+i} - 1) \prod_{j=1, j \neq i}^{r} (n+j).$$

Thus, for each $i = 1, \ldots, r$ we have

$$r(F_r(n) - (-1)^{r-1}(r-1)!) \equiv (-1)^{r-i} r \binom{r-1}{i-1}(2^{n+i} - 1) \prod_{j=1, j \neq i}^{r} (n+j) \pmod{n+i}.$$
(5)

Since $n + j \equiv j - i \pmod{n+i}$, we have

$$\prod_{j=1, j \neq i}^{r} (n+j) \equiv \prod_{j=1, j \neq i}^{r} (j-i) = \pm(r-i)!(i-1)! \pmod{n+i}.$$

Multiplying this by $r\binom{r-1}{i-1}$ one gets $\pm r!$, and substituting into (5) gives

$$r(F_r(n) - (-1)^{r-1}(r-1)!) \equiv \pm r!(2^{n+i} - 1) \pmod{n+i},$$

which completes the proof. □

References

[1] Bang, A. S., Taltheoretiske Undersøgelser, Math. Scand. **4**, 70–80 (1886).

[2] Chiriță, M. Problem 1942. Math. Mag **87**(2), 151 (2014).

[3] Erdős, P. A theorem of Sylvester and Schur J. London Math. Soc **9**(4), 282–288 (1934).

[4] López-Aguayo, D. Non-integrality of binomial sums and Fermat's little theorem. Math Mag. **88**(3), 231–234 (2015).

[5] López-Aguayo, D., Luca, F. Sylvester's theorem and the non-integrality of a certain binomial sum. Fibonacci Quart. **54**(1), 44–48 (2016).

[6] Montgomery, H. L., Vaughan, R. C., On the distribution of reduced residues. Annals of Mathematics, **123**(2), 311–333 (1986).

[7] Pomerance, C., A note on the least prime in an arithmetic progression. J. Number Theory **12** (1980), 218–223.

[8] Rosser, J. B. and Schoenfeld, L. Approximate formulas for some functions of prime numbers. Illinois J. Math. **6**, 64–94 (1962).

[9] Thongjunthug, T. Nonintegrality of certain binomial sums. European Journal of Mathematics. Advanced online publication. DOI: `https://doi.org/10.1007/s40879-018-0260-y`

Stat-Math Unit, India Statistical Institute, 7, S.J.S Sansanwal Marg, New Delhi, 110016, India.
   E-mail address: `shanta@isid.ac.in`

Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Monterrey, Nuevo León, México.
   E-mail address: `dlopez.aguayo@tec.mx`

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA.
   E-mail address: `carlp@math.dartmouth.edu`

Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand.
   E-mail address: `thotho@kku.ac.th`