On the non-monotonicity of the denominator of generalized harmonic sums

Wouter van Doorn

Abstract

Let $\sum_{i=a}^{b} \frac{1}{i} = \frac{u_{a,b}}{v_{a,b}}$ with $u_{a,b}$ and $v_{a,b}$ coprime. In their influential monograph [1, p. 34], Erdős and Graham ask, among many other questions, the following: Does there, for every fixed a, exist a b such that $v_{a,b} < v_{a,b-1}$? If so, what is the least such $b = b(a)$? In this paper we will investigate these problems in a more general setting, answer the first question in the affirmative and obtain the bounds $a + 0.54 \log(a) < b(a) < 4.38a$, which hold for all large enough a.

Contents

1 Introduction

1.1 Introduction

Let $\{r_i\}_{i\in\mathbb{N}}$ be a fixed periodic sequence of integers, not all equal to 0, with period t. That is, for every $i \in \mathbb{N}$ it holds true that $r_{i+t} = r_i$ and for at least one (and therefore for infinitely many) i, $r_i \neq 0$. Let a be a given positive integer. In this paper we shall be concerned with sums of the form $\sum_{i=a}^{b} \frac{r_i}{i}$. More precisely, if $u_{a,b} \in \mathbb{Z}$ and $v_{a,b} \in \mathbb{N}$ are coprime integers for which $\frac{u_{a,b}}{v_{a,b}} = \sum_{i=a}^{b} \frac{r_i}{i}$, we will be interested in whether $v_{a,b} < v_{a,b-1}$ holds for some b.

Paul Erdős and Ronald Graham asked this question in [1] for the case where $r_i = 1$ for all i and this was solved independently by Peter Shiu in [2] and in unpublished work (predating the current manuscript) by the author. Even though the pre-print [2] only explicitly deals with $a = 1$, their methods can be used for arbitrary $a \in \mathbb{N}$ as well. In personal communication Ernie Croot then asked about the far more general result where $r_i \in A$ for some fixed finite set A. This generalization turns out to be false, however. So it seems natural to ask for a reasonable condition on the r_i that does guarantee that $v_{a,b} < v_{a,b-1}$ holds for some b, and it will turn out that periodicity is sufficient.

Note that, in common vernacular, $v_{a,b} < v_{a,b-1}$ means that the fraction was simplified. Since a fraction can be simplified precisely when both numerator and denominator share a prime divisor, we would like to get a handle on the prime factorizations of $u_{a,b}$ and $v_{a,b}$. However, even in the special case of the harmonic numbers H_n, where $r_i = 1$ for all i, $a = 1$ and $b = n$, surprisingly little is known about this.

For example, in [3] it was conjectured that for every prime p the numerator of H_n is only finitely often divisible by p, and this is still unsolved. In the other direction, we have a well-known eponymous theorem by Wolstenholme ([4]) stating that for any prime number $p \geq 5$, the numerator of H_{p-1} is divisible by p^2. Various generalizations and extensions of this result are known and can be found in [5]. Let L_n be the least common multiple of $1, 2, \ldots, n$. In [2] it is claimed that for every sequence of odd primes $p_1, p_2, .., p_k$ there exists a positive integer n such that the denominator of H_n is a divisor of $\frac{L_n}{p_1 p_2 \cdots p_k}$. Unfortunately the proof depends on the linear independence of the terms $\theta_i = \frac{\log(p_1)}{\log(p_i)}$ for $1 \leq i \leq k$, which is not known for $k \geq 3$. The proof is valid for $k = 2$ and follows for general k from conjectures like Schanuel's Conjecture. Finally, it is often conjectured (see e.g. [1], [2] and [12]) that there exist infinitely many n for which the denominator of H_n is equal to L_n, and this too is not yet solved.

1.2 Overview of results

The main theorem we obtain in Section 2 is that for every $a \in \mathbb{N}$ there exist infinitely many integers $b > a$ for which $v_{a,b} < v_{a,b-1}$. Furthermore, if we denote by $b(a)$ the smallest such b, then there exists an effective constant c, which only depends on the sequence $\{r_i\}_{i\in\mathbb{N}}$, such that $b(a) < ca$. For example, in the original case $r_1 = t = 1$ we have the upper bound $b(a) \leq \frac{162}{37}(a-1) < 4.38a$, which is true for all $a \geq 6$. This section is (by far) the longest and most technical part of this paper and comprises the most important and interesting ideas and results.

In Section 3 we will look at lower bounds and prove that $b(a) > a + (\frac{1}{2} - \epsilon)\log(a)$ holds for all $\epsilon > 0$ and all large enough a. This lower bound turns out to be close to optimal, because if $r_i \neq 0$ for all i, then we will show the existence of infinitely many a for which $b(a) < a + (2+\epsilon)\log(a)$. We may therefore deduce that the lower limit $\liminf_{a\to\infty} \left(\frac{b(a)-a}{\log a}\right)$ exists when $r_i \neq 0$ and is bounded between $\frac{1}{2}$ and 2. We will then end this section with some improvements when $r_i = 1$ for all i, and this gives us $0.54 < \liminf_{a\to\infty} \left(\frac{b(a)-a}{\log a}\right) < 0.61$ in that case.

Finally, in Section 4 we look at two possible generalizations. First we look at what happens when the sequence $\{r_i\}_{i\in\mathbb{N}}$ is no longer assumed to be periodic. For example, if we only assume $r_i = \pm 1$, then it is possible that $v_{a,b}$ is a monotone increasing function of b. In fact, we will see that there are very few results in this paper that generalize to the non-periodic case. A nice exception to this will be a theorem stating that if the r_i are non-zero and remain bounded, then a function similar to $u_{1,b}$ will have arbitrarily large prime divisors. Secondly we look at sums of the form $\dfrac{u_{a,b}}{v_{a,b}} = \displaystyle\sum_{i=a}^{b} \frac{r_i}{i^d}$, where d is a positive integer and we define $b_d(a)$ to be the smallest positive integer b for which $v_{a,b} < v_{a,b-1}$. We will then show that, if at least two out of r_1, r_2, r_3, r_4, r_5 are non-zero and d is large enough, then $b_d(a)$ is finite for all a. To finish it all up, we will focus on the case where all r_i are equal to 1 and prove that there exists a constant $c_d = O(\log^{10}(d))$ so that for every a, $b_d(a) \leq c_d a$. We will furthermore calculate this constant c_d for all $d < 120$.

1.3 Notation

Instead of directly dealing with the sequence $v_{a,b}$, we shall instead work with the more robust sequence $L_{a,b}$, defined as the least common multiple of all integers $i \in \{a, a+1, .., b\}$ for which $r_i \neq 0$. We then define $X_{a,b}$ as $X_{a,b} = L_{a,b} \sum_{i=a}^{b} \frac{r_i}{i}$ and will abbreviate $L_{1,n}$ and $X_{1,n}$ to L_n and X_n respectively. The letters p and q are reserved for prime numbers and most other (Roman) letters will generally denote integers, often non-negative. The integer a should be viewed as fixed, but arbitrary, and $b(a)$ denotes the smallest integer $b > a$ such that $v_{a,b} < v_{a,b-1}$. All these values clearly depend on the sequence of r_i, and this dependence is always implicit; the sequence of r_i should be viewed as fixed but arbitrary as well.

Whenever we say that p^k *exactly divides* an integer n, we mean that n is divisible by p^k, but not by p^{k+1}. In other words, p^k is the largest power of p that divides n, and whenever p does not divide n at all, then this number k equals 0. If the prime p is fixed or understood, then $e(n)$ will often denote the number k such that p^k exactly divides n. When confusion might arise we sometimes use a subscript like $e_p(n)$ to emphasize the dependence on the prime p.

$O(f(x))$ and $o(f(x))$ are the familiar Big-O and Little-o notations, while $x|y$ reads 'x divides y'. The symbols $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$ represent the set of real numbers, the set of integers and the set of positive integers respectively. The greek letter $\lambda = \lambda(t)$ will be the Carmichael function; the smallest positive integer such that $p^\lambda \equiv 1 \pmod t$ for all p with $\gcd(p,t) = 1$. The dependence of λ on t will always be implicit and we have $\lambda|\varphi(t)$, where φ is Euler's totient function. The number of primes smaller than or equal to n is denoted by $\pi(n)$ and we often make use of the prime number theorem which states that $\lim_{n\to\infty} \frac{\pi(n)\log(n)}{n} = 1$. We will refer to both the prime number theorem and its generalization to arithmetic progressions by the acronym PNT. Finally, ϵ will denote a small, positive real number.

2 Upper bounds

2.1 Proof strategy

Our goal in this section is to prove that $b(a)$ is finite and, moreover, that there exists a constant c such that for every a we have $b(a) < ca$. For pedagogical purposes we will first prove this in Section 2.2 if we assume a certain large prime p divides X_n for some $n \in \mathbb{N}$. This furthermore motivates the rest of the proof; trying to find such a large prime divisor p of X_n. That such a prime exists is immediate when $r_1 = t = 1$, initially leading to a bound of $b(a) \leq 6a$ in that case. In Section 2.3 we will look at some examples and prove that when $r_i = 1$ and $a \geq 6$, we can tighten the bound to $b(a) \leq \frac{162}{37}(a-1)$.

To find this large prime divisor of X_n, we first have to show a lower bound on X_n itself. We will do this in Section 2.4 where we first show that there exists a constant c_0 such that $|X_n| > c_0^n$ holds for all large enough n. This follows from some estimates on $\frac{X_n}{L_n}$ and the fact that L_n grows exponentially fast. However, in the end we not only would like to prove $b(a) < ca$, we actually want to give an explicit value for this constant c as well. So phrases like 'for large enough n' will generally not suffice. Therefore, we take some time to find an interval that we can write down explicitly, where $|X_n|$ is large enough for our purposes for sufficiently many n in that interval.

Section 2.5 is then aimed at proving that the prime divisors of X_n can get arbitrarily large. If we let $r = \max_i |r_i|$ and define $m = 1 + \max(r, t)$ (although any integer larger than $\max(r, t)$ works), then our proof will actually show that for every interval I of length at least e^{7m}, there exists an $n \in I$ for which X_n is divisible by a prime $p \geq m$.

The proof of that statement goes through a few steps; first we split up the primes into three subsets Σ_1, Σ_2 and Σ_3. The first subset contains the primes larger than or equal to m, so it would suffice to find an $n \in I$ for which the largest divisor of X_n containing only primes from Σ_2 or Σ_3 is smaller than $|X_n|$. Then we will see that the largest divisor of X_n containing only primes from Σ_3 is always small in a certain congruence class. And finally, let $2 \leq p_1 < p_2 < \ldots < p_y < m$ be the primes in Σ_2. We will construct a nesting sequence of intervals $I \supset I_1 \supset I_2 \supset \ldots \supset I_y$, for which the power of $p_{\sigma(j)}$ that divides X_n is small for $n \in I_j$, where $\sigma : \{1, 2, \ldots, y\} \to \{1, 2, \ldots, y\}$ is a permutation. And so for all $n \in I_y$ the power of any prime in Σ_2 dividing X_n is small. Combining these estimates on the powers of primes from Σ_2 and Σ_3 that divide X_n then implies that X_n must have a prime divisor from Σ_1 as well.

Write $n = lp^k$ with $\gcd(l, p) = 1$ and $p \geq m$ a prime that divides X_n. By Section 2.5 such n and p exist. Then by setting $b = np^{\lambda k_1}$ for some suitable k_1, it turns out that in order to show $v_{a,b} < v_{a,b-1}$, we need to check that $\gcd(l, X_{a,b-1}) < p$. Now, in the case that $r_i \neq 0$ for all i with $\gcd(i, t) = 1$, we have $l < p$, so this condition is trivially satisfied. This will allow us to calculate an explicit upper bound in Section 2.6 for the constant c for which $b(a) < ca$ holds for all a, when $\gcd(i,t) = 1$ implies $r_i \neq 0$. This c will turn out to grow doubly exponential in m.

In the general case it is possible that $l > p$, which can make it more difficult to check the condition $\gcd(l, X_{a,b-1}) < p$. So our goal is to make sure that $\gcd(l, X_{a,b-1})$ is small and we therefore need some information on the prime divisors of l and $X_{a,b-1}$. Section 2.7 is then dedicated to proving that for every prime $q \in \Sigma_1 \cup \Sigma_2$ there are certain intervals such that for all n inside those intervals, the power of q that divides X_n is bounded.

In Section 2.8 we then pick a prime $q \in \Sigma_1 \cup \Sigma_2$ for which, if q^y exactly divides l, then q^y is large. Using results from Section 2.7 we can ensure that, if $b-1$ is contained in a certain interval, then the power of q that divides $X_{a,b-1}$ is small, so that $\gcd(l, X_{a,b-1})$ is small as well, which was our goal. These intervals are of the form $[c_q q^{\lambda k_2}, (c_q+1)q^{\lambda k_2})$, where c_q is a constant and k_2 can be any integer. So when we now choose $b = np^{\lambda k_1}$, for some k_1, then we need the inequalities $c_q q^{\lambda k_2} < np^{\lambda k_1} \leq (c_q+1)q^{\lambda k_2}$ to hold. When we take logarithms, we end up with a linear form in logarithms and, using a well-known Diophantine approximation result by Dirichlet, these inequalities can be satisfied infinitely often, implying that $b(a)$ is finite.

Finally, by using an extension of a result by Baker, we also have a lower bound for the linear form in logarithms that we encountered in Section 2.8. In Section 2.9 we then use this lower bound to give an explicit linear upper bound for $b(a)$. In this general case the constant c grows triply exponential in m.

2.2 Under the assumption of a large prime divisor

Let $r = \max_i |r_i|$ and define c_1 to be the smallest positive integer such that $r_{c_1} \neq 0$. Now let $p > \max(r,t)$ be a prime number that divides X_i for some integer $i \geq c_1$ and let $n = n(p)$ be the smallest such i. In Section 2.5 we will prove that such a prime p actually exists, but for now we will simply assume we have one at our disposal.

Necessarily we see that p does not divide X_{n-1} and $r_n \neq 0$. Since $p > \max(r,t) \geq r_n$ this implies $0 < |r_n| < p$. Write $n = lp^k$ with $\gcd(l,p) = 1$ and let λ be such that $q^\lambda \equiv 1 \pmod t$, whenever $\gcd(q,t) = 1$. Now we set $b = np^{\lambda k_1} = lp^{\lambda k_1 + k}$, where k_1 is an integer for which $p^{\lambda k_1 + k} \geq \max(a, 2t)$. We then have the following theorem.

Theorem 1. *If* $\gcd(l, X_{a,b-1}) < p$, *then* $v_{a,b} < v_{a,b-1}$. *Furthermore, if the condition* $\gcd(l, X_{a,b-1}) < p$ *is satisfied for the smallest* k_1 *such that* $p^{\lambda k_1 + k} \geq \max(a, 2t)$, *then* $b(a) \leq \max(a-1, 2t-1)lp^\lambda$.

Proof. Let us first remark that the second part is easy to see, because for the smallest possible k_1, we have $p^{\lambda k_1+k} > \max(a-1, 2t-1) \geq p^{\lambda(k_1-1)+k}$, implying $b = lp^{\lambda k_1+k} \leq \max(a-1, 2t-1)lp^{\lambda}$. Now, in general we have $\frac{u_{a,b}}{v_{a,b}} = \frac{X_{a,b}}{L_{a,b}}$, so if we define $g_{a,b}$ to be the greatest common divisor of $X_{a,b}$ and $L_{a,b}$, then $v_{a,b} = \frac{L_{a,b}}{g_{a,b}}$. And thus, if $L_{a,b} = L_{a,b-1}$, then $v_{a,b} < v_{a,b-1}$ holds true, precisely when $g_{a,b} > g_{a,b-1}$. We claim that, indeed, $L_{a,b}$ and $L_{a,b-1}$ are equal while $g_{a,b}$ is larger than $g_{a,b-1}$. We start with the first part of this claim, but before we do so, we need some properties.

Lemma 1. *There exists an $l_1 \in \mathbb{N}$ with $1 \leq l_1 < l$ for which $r_{l_1 p^k} \neq 0$. Furthermore, p^k exactly divides L_n and $p^{\lambda k_1+k}$ exactly divides $L_{a,b}$.*

Proof. As we will do a lot in this paper, we look at $X_n \pmod p$ and remove the terms in the sum which are divisible by p. In general, L_n, which is divisible by $n = lp^k$, must be divisible by p^k. In other words, if $\frac{L_n r_i}{i}$ does not vanish modulo p, then p^k divides i. But when no other $l_1 < l$ with $r_{l_1 p^k} \neq 0$ exists, there is only one i for which $\frac{L_n r_i}{i}$ does not vanish modulo p, namely $i = lp^k = n$ itself. So we would then get:

$$\begin{aligned} X_n &= L_n \sum_{i=1}^{n} \frac{r_i}{i} \\ &= \sum_{i=1}^{n} \frac{L_n r_i}{i} \\ &\equiv \frac{L_n r_n}{n} && \pmod p \\ &\equiv \frac{L_n r_n}{lp^k} && \pmod p \\ &\not\equiv 0 && \pmod p \end{aligned}$$

And this would contradict the assumption that p divides X_n. Note that the last inequality uses the fact that $0 < |r_n| < p$. So this proves the first property and remark that this property implies that p does not divide $\frac{L_n}{L_{n-1}}$.

For the other two properties, since L_n is divisible by p^k and $b = np^{\lambda k_1} \equiv n \pmod t$, we see that $r_b = r_n \neq 0$, which implies that $L_{a,b}$ is divisible by $p^{\lambda k_1+k}$. To prove that these are also the largest powers of p dividing L_n and $L_{a,b}$, assume by contradiction that $p^{\lambda k_1+k+1}$ divides $L_{a,b}$. We will show that this implies that L_n is divisible by p^{k+1}, which will lead to a contradiction. If $p^{\lambda k_1+k+1}$ divides $L_{a,b}$, then there exists a positive integer g with $a \leq g \leq b$ such that g is divisible by $p^{\lambda k_1+k+1}$ and for which $r_g \neq 0$. Now we can choose $h = gp^{-\lambda k_1} \leq bp^{-\lambda k_1} = n$ and note that $h \equiv g \pmod t$ by definition of λ, so $r_h = r_g$, which we assumed to be non-zero. Furthermore, h would be divisible by p^{k+1} and, since $r_h \neq 0$, so would L_n. However, $\frac{L_n r_n}{n}$ would then vanish modulo p and we would get

$X_n = \frac{L_n}{L_{n-1}} X_{n-1} + \frac{L_n r_n}{n} \equiv \frac{L_n}{L_{n-1}} X_{n-1} \pmod p$. This is impossible, since it contradicts the assumption that n is the smallest i for which p divides X_i. $\square$

With the proof of Lemma 1 out of the way, let us focus on our (intermediate) goal again; proving that $L_{a,b}$ and $L_{a,b-1}$ are equal to each other, in which case $v_{a,b} < v_{a,b-1}$ is equivalent with $g_{a,b} > g_{a,b-1}$.

Lemma 2. $L_{a,b} = L_{a,b-1}$.

Proof. Since $L_{a,b} = \mathrm{lcm}(b, L_{a,b-1}) = \mathrm{lcm}(lp^{\lambda k_1 + k}, L_{a,b-1})$ with $\gcd(l, p^{\lambda k_1 + k}) = 1$, it suffices to show that both l and $p^{\lambda k_1 + k}$ divide $L_{a,b-1}$.

We observe that $l | (b - lt)$ and we claim that this implies that $l | L_{a,b-1}$. To see this, first note that $r_{b-lt} = r_b = r_n \neq 0$. Secondly, $b > b - lt \geq l \max(a, 2t) - lt = l \max(a - t, t) \geq 2 \max(a - t, t) \geq a$, since $l \geq 2$ by Lemma 1. And so we conclude that $b - lt$, which is a multiple of l, lies in the interval $[a, b-1]$ and must therefore divide $L_{a,b-1}$.

To show that $p^{\lambda k_1 + k}$ divides $L_{a,b-1}$, we use the existence of a positive integer $l_1 < l$ for which $r_{l_1 p^k} \neq 0$, as guaranteed by Lemma 1. Using this l_1, then $r_{l_1 p^{\lambda k_1 + k}} \neq 0$ as well, while $a \leq p^{\lambda k_1 + k} \leq l_1 p^{\lambda k_1 + k} < l p^{\lambda k_1 + k} = b$. And so $L_{a,b-1}$ will be divisible by $l_1 p^{\lambda k_1 + k}$, and in particular by $p^{\lambda k_1 + k}$. $\square$

Now it suffices to show that $g_{a,b} > g_{a,b-1}$. Morally, this holds because p dividing X_n implies that p divides $X_{a,b}$ as well. While, on the other hand, p does not divide $X_{a,b-1}$, since $X_{a,b-1} \not\equiv X_{a,b} \pmod p$. Let us make this into a lemma.

Lemma 3. *p divides $X_{a,b}$, while p does not divide $X_{a,b-1}$.*

Proof. Let us take a look at $X_n \pmod p$ again.

$$\begin{aligned} X_n &= L_n \sum_{i=1}^{n} \frac{r_i}{i} \\ &\equiv L_n \sum_{i=1}^{l} \frac{r_{ip^k}}{ip^k} && \pmod p \\ &\equiv \frac{L_n}{p^k} \sum_{i=1}^{l} \frac{r_{ip^k}}{i} && \pmod p \\ &\equiv 0 && \pmod p \end{aligned}$$

By Lemma 1, p^k exactly divides L_n, so for this sum to be congruent to $0 \pmod p$ we must have that $\sum_{i=1}^{l} \frac{r_{ip^k}}{i} \equiv 0 \pmod p$. Now let us use this knowledge in the analogous sum for $X_{a,b}$.

$$\begin{aligned} X_{a,b} &= L_{a,b} \sum_{i=a}^{b} \frac{r_i}{i} \\ &\equiv L_{a,b} \sum_{i=1}^{l} \frac{r_{ip^{\lambda k_1+k}}}{ip^{\lambda k_1+k}} && \pmod p \\ &\equiv \frac{L_{a,b}}{p^{\lambda k_1+k}} \sum_{i=1}^{l} \frac{r_{ip^k}}{i} && \pmod p \\ &\equiv 0 && \pmod p \end{aligned}$$

And indeed we see that p divides $X_{a,b}$ as well. On the other hand, note that p does not divide $\frac{L_{a,b} r_b}{lp^{\lambda k_1+k}}$ by Lemma 1. From this observation it follows that $X_{a,b-1} = X_{a,b} - \frac{L_{a,b} r_b}{lp^{\lambda k_1+k}} \not\equiv X_{a,b} \pmod p$, and we conclude that p does not divide $X_{a,b-1}$. $\square$

Now we are almost ready to finish up our proof, but before we do so, we need one last lemma.

Lemma 4. *Let w, x, y, z be any non-zero integers such that y divides wx. Then $\gcd(w, \frac{wx}{y} + z)$ is an integer multiple of $\frac{\gcd(w,z)}{\gcd(y, \frac{z}{\gcd(z, \frac{w}{y})})}$. In particular, $\gcd(w, \frac{wx}{y} + z) \geq \frac{\gcd(w,z)}{\gcd(y,z)}$.*

Proof. To prove this, let q be any prime dividing w and let $q^\alpha, q^\beta, q^\gamma, q^\delta$ be the largest powers of q dividing w, x, y, z respectively. If we let q^μ be the largest power of q dividing $\gcd(w, \frac{wx}{y} + z)$ and let q^ν be the largest power of q dividing $\gcd(y, \frac{z}{\gcd(z, \frac{w}{y})})$, then what we want to prove is equivalent to $\mu + \nu \geq \min(\alpha, \delta)$.

Note that $\mu \geq \min(\alpha, \alpha + \beta - \gamma, \delta)$, so that $\mu + \nu \geq \min(\alpha, \alpha + \beta - \gamma + \nu, \delta)$. It therefore suffices to show that $\alpha + \beta - \gamma + \nu \geq \min(\alpha, \delta)$. Now we remark that ν is equal to $\min(\gamma, \delta - \min(\delta, \alpha - \gamma))$, so there are two possibilities. If $\nu = \gamma$, then we see $\alpha + \beta - \gamma + \nu = \alpha + \beta$. On the other hand, if $\nu = \delta - \min(\delta, \alpha - \gamma) \geq \delta - \alpha + \gamma$, then $\alpha + \beta - \gamma + \nu \geq \beta + \delta$. In both cases $\alpha + \beta - \gamma + \nu \geq \min(\alpha, \delta)$, and we are done. $\square$

We are going to apply Lemma 4 with $w = L_{a,b} p^{-\lambda k_1 - k}$, $x = r_b$, $y = l$ and $z = X_{a,b-1}$. The integers w and y are trivially non-zero, the fact that $x \neq 0$ was mentioned in the proof of Lemma 1 and z is non-zero by Lemma 3. We will now calculate $g_{a,b}$ to finish the proof of Theorem 1.

$$
\begin{aligned}
g_{a,b} &= \gcd(L_{a,b}, X_{a,b}) \\
&= \gcd(p^{\lambda k_1+k}, X_{a,b}) \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b}\right) \\
&\geq p \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b}\right) \\
&= p \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b-1} + \frac{L_{a,b} r_b}{b}\right) \\
&= p \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b-1} + \frac{L_{a,b} r_b}{l p^{\lambda k_1+k}}\right) \\
&\geq \frac{p}{\gcd(l, X_{a,b-1})} \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b-1}\right) \\
&> \gcd\left(\frac{L_{a,b}}{p^{\lambda k_1+k}}, X_{a,b-1}\right) \\
&= \gcd\left(\frac{L_{a,b-1}}{p^{\lambda k_1+k}}, X_{a,b-1}\right) \\
&= \gcd(L_{a,b-1}, X_{a,b-1}) \\
&= g_{a,b-1}
\end{aligned}
$$

□

2.3 Some examples and a return to the classical case

In this subsection we would like to look at a couple of examples, in particular the case where $r_i = 1$ for all i. The ideas and proofs in this section are not needed for the rest of this chapter, but seeing what happens in a few explicit examples, could prove useful in building some intuition.

Since $\gcd(l, X_{a,b-1}) \leq l \leq n$, it is worth pointing out that as soon as we find an integer n and a prime $p > \max(r, t, n)$ such that p divides X_n, then the condition in Theorem 1 is satisfied and $b(a)$ is finite for all a. In practice in turns out that, regardless of the sequence of $r_1, r_2, \ldots$ that is chosen, one very often quickly finds such positive integers n for which X_n is divisible by a prime $p > \max(r, t, n)$. As an instructive example, let us look at all possible sequences of r_i for which $\max(r, t) \leq 2$. Since we care about the divisors of X_n, we should view two sequences $r_1, \ldots, r_t$ and $r'_1, \ldots, r'_t$ as morally the same if $r_i = -r'_i$ for all i, as X_n and $-X_n$ have the same divisors. We will therefore assume that the first non-zero r_i is actually positive. Furthermore, for $t = 2$ we may, without loss of generality, assume that $r_1 \neq r_2$. With these assumptions there are 12 distinct sequences with $\max(r, t) \leq 2$. We have tabulated these sequences, together with an n and a prime $p > \max(r, t, n)$ such that X_n is divisible by p. And one can see that for all 12 sequences, one only has to add two or three (non-zero) fractions together to find such an n. Moreover, for the 8 sequences with $r_i \neq 0$ for all i we would, with the use of Theorem 1, obtain $b(a) \leq 21a$.

t	r_1	r_2	n	p
1	1	-	2	3
1	2	-	2	3
2	1	-2	2	3
2	1	-1	3	5
2	1	0	5	23
2	1	2	3	7
2	2	-2	3	5
2	2	-1	2	3
2	2	0	5	23
2	2	1	2	5
2	0	1	6	11
2	0	2	6	11

Specifying to $r_i = 1$ for the rest of this section, we obtain the following corollary of Theorem 1.

Corollary 1. *In the case where $r_i = 1$ for all i, we have $b(a) \leq 6(a-1)$, for all $a > 1$.*

It is however possible to improve upon this corollary. Recall that, if k is such that $3^k < a \leq 3^{k+1}$, then the proof of Theorem 1 shows that with $f(a) = 2 \cdot 3^{k+1}$ one has $v_{a,f(a)} < v_{a,f(a)-1}$. So for all $a \in (3^k, 3^{k+1}]$ the same value of $f(a)$ is chosen. To improve upon Corollary 1, for $k \geq 4$ we are going to split up the interval $(3^k, 3^{k+1}]$ into five sub-intervals and let the value of $f(a)$ depend on the sub-interval that contains a. First, let us state our improvement.

Theorem 2. *In the case where $r_i = 1$ for all i, we have $b(a) \leq \frac{162}{37}(a-1) < 4.38a$, for all $a \geq 6$.*

Proof. To start off, for $6 \leq a \leq 81$, we define $f(a)$ as follows:

$$f(a) = \begin{cases} 18 & \text{if } 6 \leq a \leq 9 \\ 35 & \text{if } 10 \leq a \leq 14 \\ 54 & \text{if } 15 \leq a \leq 27 \\ 75 & \text{if } 28 \leq a \leq 50 \\ 162 & \text{if } 51 \leq a \leq 81 \end{cases}$$

Then one can check that $f(a) < 4(a-1) < \frac{162}{37}(a-1)$ holds for all these a and, possibly with the help of a computer, one can also check that in each case $v_{a,f(a)} < v_{a,f(a)-1}$, proving that Theorem 2 is true for all $a \leq 81$. In fact, for the first, third and fifth intervals of a, $v_{a,f(a)} < v_{a,f(a)-1}$ follows from the proof of Theorem 1 and for the other two intervals one can prove $v_{a,f(a)} < v_{a,f(a)-1}$ without too much trouble by looking at $X_{a,f(a)} \pmod{p}$ for the primes p dividing $f(a)$. So pen and paper should suffice and the help of a computer is not even needed. We should point out, by the way, that $f(a)$ is generally not equal to $b(a)$. That is, for most a in the above range, there exists a b which is smaller than $f(a)$, such that already $v_{a,b} < v_{a,b-1}$. In any case, we may assume $a \geq 82$.

For $a \geq 82$, there exists a $k \geq 4$ such that $3^k < a \leq 3^{k+1}$. We will partition $I := (3^k, 3^{k+1}]$ into the following five intervals:

$$
\begin{aligned}
I_1 &= (3^k, 10 \cdot 3^{k-2}] \\
I_2 &= (10 \cdot 3^{k-2}, 11 \cdot 3^{k-2}] \\
I_3 &= (11 \cdot 3^{k-2}, 4 \cdot 3^{k-1}] \\
I_4 &= (4 \cdot 3^{k-1}, 37 \cdot 3^{k-3}] \\
I_5 &= (37 \cdot 3^{k-3}, 3^{k+1}]
\end{aligned}
$$

We now define $f(a)$ as follows:

$$
f(a) = \begin{cases} 5 \cdot 3^{k-1} & \text{if } a \in I_1 \\ 16 \cdot 3^{k-2} & \text{if } a \in I_2 \\ 5 \cdot 3^{k-1} & \text{if } a \in I_3 \\ 14 \cdot 3^{k-2} & \text{if } a \in I_4 \\ 2 \cdot 3^{k+1} & \text{if } a \in I_5 \end{cases}
$$

Again it is straight-forward to check that for all $a \in I$, $f(a)$ is smaller than or equal to $\frac{162}{37}(a-1)$, so it suffices to prove $v_{a,f(a)} < v_{a,f(a)-1}$.

For all $a \in I_5$, the proof of Theorem 1 tells us that $v_{a,f(a)} < v_{a,f(a)-1}$. For a in the other four intervals, Theorem 1 does not directly help, but we will follow its proof quite closely with $p = 3$.

First, analogously to Lemma 2, we remark that in all cases $L_{a,f(a)} = L_{a,f(a)-1}$. To see this, write $f(a) = l \cdot 3^{k_1}$ with $\gcd(l, 3) = 1$ and recall that $L_{a,f(a)} = \operatorname{lcm}(f(a), L_{a,f(a)-1})$. Since $l | f(a) - l$ and $3^{k_1} | f(a) - 3^{k_1}$ and, in all cases, $a \leq \min(f(a) - l, f(a) - 3^{k_1})$, we indeed get $L_{a,f(a)} = \operatorname{lcm}(f(a), L_{a,f(a)-1}) = \operatorname{lcm}(l \cdot 3^{k_1}, L_{a,f(a)-1}) = L_{a,f(a)-1}$. The main difference with the proof of Theorem 1 lies in the fact that $X_{a,f(a)}$ is not just divisible by 3; we actually claim that 9 divides $X_{a,f(a)}$ for all a in the first four intervals. We will then make use of the following very slight extension of Theorem 1, the proof of which is basically the same.

Theorem (Extension of Theorem 1)**.** *Let e and $f > e$ be such that p^e exactly divides $X_{a,f(a)-1}$ and p^f divides both $X_{a,f(a)}$ and $L_{a,f(a)}$. If $\gcd(l, X_{a,f(a)-1}) < p^{f-e}$, then $v_{a,f(a)} < v_{a,f(a)-1}$.*

To show that 9 does indeed divide $X_{a,f(a)}$ for all $a \in \bigcup_{1 \leq i \leq 4} I_i$, we use the fact that if $L_{a,f(a)}$ is exactly divisible by 3^{k_1}, then $\frac{L_{a,f(a)}}{i} \equiv 0 \pmod 9$, unless

3^{k_1-1} divides i. So to calculate $X_{a,f(a))} \pmod 9$ the only terms $\frac{L_{a,f(a)}}{i}$ that we have to add are the ones where 3^{k_1-1} divides i. Note that in all the four intervals we consider, we have $3^k < a < f(a) < 2\cdot 3^k$, so that k_1 is at most $k-1$.

1. For $a \in I_1$ we have chosen $f(a) = 5 \cdot 3^{k-1}$, so that $L_{a,f(a)}$ is exactly divisible by 3^{k-1}. This means that, modulo 9, the only non-zero terms in the sum for $X_{a,f(a)}$ are the ones where i is divisible by 3^{k-2}. We will now calculate $X_{a,f(a)} \pmod 9$ by rearranging those terms and then taking consecutive terms together.

$$\begin{aligned} X_{a,f(a)} &\equiv \frac{L_{a,f(a)}}{10\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{11\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{4\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{5\cdot 3^{k-1}} &\pmod 9 \\ &\equiv \left(\frac{L_{a,f(a)}}{10\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{11\cdot 3^{k-2}}\right) + \left(\frac{L_{a,f(a)}}{4\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{5\cdot 3^{k-1}}\right) + \left(\frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}}\right) &\pmod 9 \\ &\equiv \frac{21L_{a,f(a)}}{110\cdot 3^{k-2}} + \frac{9L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{27L_{a,f(a)}}{182\cdot 3^{k-2}} &\pmod 9 \\ &\equiv 9\left(\frac{7L_{a,f(a)}}{110\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{9L_{a,f(a)}}{182\cdot 3^{k-1}}\right) &\pmod 9 \\ &\equiv 0 &\pmod 9 \end{aligned}$$

2. For $a \in I_2$, we have chosen $f(a) = 16\cdot 3^{k-2}$. And since $a < 5\cdot 3^{k-1} < f(a)$, $L_{a,f(a)}$ is again exactly divisible by 3^{k-1}.

$$\begin{aligned} X_{a,f(a)} &\equiv \frac{L_{a,f(a)}}{11\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{4\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{5\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{16\cdot 3^{k-2}} &\pmod 9 \\ &\equiv \left(\frac{L_{a,f(a)}}{11\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{16\cdot 3^{k-2}}\right) + \left(\frac{L_{a,f(a)}}{4\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{5\cdot 3^{k-1}}\right) + \left(\frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}}\right) &\pmod 9 \\ &\equiv \frac{27L_{a,f(a)}}{176\cdot 3^{k-2}} + \frac{9L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{27L_{a,f(a)}}{182\cdot 3^{k-2}} &\pmod 9 \\ &\equiv 9\left(\frac{9L_{a,f(a)}}{176\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{9L_{a,f(a)}}{182\cdot 3^{k-1}}\right) &\pmod 9 \\ &\equiv 0 &\pmod 9 \end{aligned}$$

3. The calculation for $a \in I_3$ is very similar to the one for the first interval, except that it does not contain the two terms corresponding to $10 \cdot 3^{k-2}$ and $11 \cdot 3^{k-2}$.

$$
\begin{aligned}
X_{a,f(a)} &\equiv \frac{L_{a,f(a)}}{4\cdot 3^{k-1}} + \frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{5\cdot 3^{k-1}} &&\pmod 9\\
&\equiv \frac{9L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{27L_{a,f(a)}}{182\cdot 3^{k-2}} &&\pmod 9\\
&\equiv 9\left(\frac{L_{a,f(a)}}{20\cdot 3^{k-1}} + \frac{9L_{a,f(a)}}{182\cdot 3^{k-1}}\right) &&\pmod 9\\
&\equiv 0 &&\pmod 9
\end{aligned}
$$

4. Finally, for $a \in I_4$ we have $4\cdot 3^{k-1} < a < f(a) < 5\cdot 3^{k-1}$, so $L_{a,f(a)}$ is exactly divisible by 3^{k-2} and $\frac{L_{a,f(a)}}{i} \equiv 0 \pmod 9$, unless 3^{k-3} divides i.

$$
\begin{aligned}
X_{a,f(a)} &\equiv \frac{L_{a,f(a)}}{37\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{38\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{40\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{41\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}} &&\pmod 9\\
&\equiv \left(\frac{L_{a,f(a)}}{37\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{38\cdot 3^{k-3}}\right) + \left(\frac{L_{a,f(a)}}{13\cdot 3^{k-2}} + \frac{L_{a,f(a)}}{14\cdot 3^{k-2}}\right) + \left(\frac{L_{a,f(a)}}{40\cdot 3^{k-3}} + \frac{L_{a,f(a)}}{41\cdot 3^{k-3}}\right) &&\pmod 9\\
&\equiv \frac{75L_{a,f(a)}}{1406\cdot 3^{k-3}} + \frac{27L_{a,f(a)}}{182\cdot 3^{k-2}} + \frac{81L_{a,f(a)}}{1640\cdot 3^{k-3}} &&\pmod 9\\
&\equiv 9\left(\frac{25L_{a,f(a)}}{1406\cdot 3^{k-2}} + \frac{3L_{a,f(a)}}{182\cdot 3^{k-2}} + \frac{27L_{a,f(a)}}{1640\cdot 3^{k-2}}\right) &&\pmod 9\\
&\equiv 0 &&\pmod 9
\end{aligned}
$$

For $a \in I_1 \cup I_3 \cup I_4$, we see that 3 does not divide $\frac{L_{a,f(a)}}{f(a)}$, while for $a \in I_2$, 9 does not divide $\frac{L_{a,f(a)}}{f(a)}$. Since $X_{a,f(a)-1} = X_{a,f(a)} - \frac{L_{a,f(a)}}{f(a)}$, this implies (compare with Lemma 3) that for $a \in I_1 \cup I_3 \cup I_4$, 3 does not divide $X_{a,f(a)-1}$, while for $a \in I_2$, 9 does not divide $X_{a,f(a)-1}$.

Using the aforementioned extension of Theorem 1, it suffices to show that $\gcd(l, X_{a,f(a)-1}) < 9$ for $a \in I_1 \cup I_3 \cup I_4$ and $\gcd(l, X_{a,f(a)-1}) < 3$ for $a \in I_2$. Since $l = 5, 16, 5, 14$ for $i = 1, 2, 3, 4$ respectively, this at once follows from the following well-known proposition.

Lemma 5. *For all a and $b \geq a$, $X_{a,b}$ is odd.*

Proof. Let m be such that $L_{a,b}$ is exactly divisible by 2^m, and let $i \in [a,b]$ be an integer divisible by 2^m. Then we claim that this i is unique; if $i' \neq i$ is also divisible by 2^m, then $i' \notin [a,b]$. To see this, first note that if i' is divisible by 2^m, then either $i' \leq i - 2^m$ or $i' \geq i + 2^m$. Secondly note that, since i is exactly divisible by 2^m, it must be an odd multiple of 2^m. This implies that $i - 2^m$ and $i + 2^m$ are both even multiples of 2^m, which means they are divisible by 2^{m+1}. Since $L_{a,b}$ is not divisible by 2^{m+1}, this then shows that both $i - 2^m$ and $i + 2^m$ have to be outside of the interval $[a,b]$, so i' cannot be contained in $[a,b]$ either. Since we have shown that this i is unique, this implies $X_{a,b} \equiv \frac{L_{a,b}}{i} \equiv 1 \pmod 2$. □

2.4 Exponential growth

In Section 2.2 we used a prime $p > \max(r,t)$ dividing X_n, for some $n \in \mathbb{N}$. We will now start to concern ourselves with proving the existence of such a p. In order to do this, the first thing we need to find are lower bounds on the growth of X_n itself. For whomever just wants an exponential lower bound that works for all large enough n, we will prove that first. However, in this paper we would love to end up with explicit bounds, and for that we need to work a bit harder, which we shall do right after.

Lemma 6. *For all $n > 3t^2$ we have $L_n > 2^{\frac{n}{t}-2}$.*

Proof. Let c_1 be the smallest positive integer such that $r_{c_1} \neq 0$. Then

$$\begin{aligned} L_n &\geq \operatorname{lcm}(c_1, c_1+t, c_1+2t, \ldots, c_1+At) \\ &\geq \operatorname{lcm}\left(\frac{c_1}{\gcd(c_1,t)}, \frac{c_1+t}{\gcd(c_1,t)}, \frac{c_1+2t}{\gcd(c_1,t)}, \ldots, \frac{c_1+At}{\gcd(c_1,t)}\right) \end{aligned}$$

where $A = \left\lfloor \left(\frac{n-c_1}{t}\right) \right\rfloor > \frac{n}{t} - 2$. By Theorem 1.1 from [6, p. 2] [1] we obtain

$$\begin{aligned} L_n &\geq \left(\frac{c_1}{\gcd(c_1,t)}\right)\left(\frac{t}{\gcd(c_1,t)}\right)\left(\frac{t}{\gcd(c_1,t)}+1\right)^A \\ &> 2^{\frac{n}{t}-2} \end{aligned}$$

which holds when $A > \frac{t}{\gcd(c_1,t)}$. And when $n > 3t^2$, then $A > \frac{n}{t} - 2 > 3t - 2 \geq t \geq \frac{t}{\gcd(c_1,t)}$. □

We will now use Lemma 6 to prove lower bounds on $|X_n|$.

Lemma 7. *There exists a constant $c > 1$ such that for all large enough integers n, $|X_n| > c^n$. In fact, any c smaller than $2^{1/t}$ will work.*

Proof. By Lemma 6 L_n grows exponentially fast with base at least $2^{1/t}$. It therefore suffices to show that $\left|\frac{X_n}{L_n}\right|$ cannot go to 0 exponentially fast. And indeed, we will see that $\left|\frac{X_n}{L_n}\right| > c_0 n^{-t}$ for large enough n and some constant c_0.

Fix a residue class $n \pmod t$ and note that $\dfrac{X_{n+t}}{L_{n+t}} - \dfrac{X_n}{L_n} = \displaystyle\sum_{i=n+1}^{n+t} \frac{r_i}{i}$ can be written as $\frac{f(n)}{g(n)}$, where $f(n)$ and $g(n)$ are polynomials with integral coefficients and degree at most t. If the leading coefficients of $f(n)$ and $g(n)$ have the same sign, then $\frac{f(n)}{g(n)}$ is positive for all large n, and if the leading coefficients of $f(n)$ and $g(n)$ differ in sign, then $\frac{f(n)}{g(n)}$ is negative for all large n. Either way, this

[1] With $\alpha = 1$, their n is our A, their r is our $\frac{t}{\gcd(c_1,t)}$ and their u_0 is our $\frac{c_1}{\gcd(c_1,t)}$.

implies that the sequence $\frac{X_n}{L_n}, \frac{X_{n+t}}{L_{n+t}}, \frac{X_{n+2t}}{L_{n+2t}}, ..$ is monotonic, for large enough n. If this sequence does not converge to zero, we are clearly done. If it does converge to zero, we have (for some constant c' and large enough n):

$$\begin{aligned} \left|\frac{X_n}{L_n}\right| &= \left|\frac{X_n}{L_n} - 0\right| \\ &> \left|\frac{X_n}{L_n} - \frac{X_{n+t}}{L_{n+t}}\right| \\ &= \left|\frac{f(n)}{g(n)}\right| \\ &> c'n^{-t} \end{aligned}$$

We can now take c_0 to be the minimum value of c' over all residue classes modulo t, and we are done. □

Like we mentioned before however, we would like to find explicit bounds. And to that end, we introduce some notation. Define $m = \max(r+1, t+1)$ and note that by the table in Section 2.3, we may assume $m \geq 4$. Let z be the number of primes strictly below m and define $\tilde{m}$ to be the smallest integer larger than $e^{6.7m}$ with $\tilde{m} \equiv c_1 \pmod{t}$ and such that $\tilde{m}$ has a prime divisor q_0 larger than e^{5m}. Finally, we define the half-open interval $I = [\tilde{m} - m^{3z+7}, \tilde{m} + m^{3z+7})$ and subdivide it into the sub-intervals $J_1 = [\tilde{m} - m^{3z+7}, \tilde{m})$ and $J_2 = [\tilde{m}, \tilde{m} + m^{3z+7})$.

We will prove that either $|X_n| \geq m^2 n^z$ for all $n \in J_1$ or $|X_n| \geq m^2 n^z$ for all $n \in J_2$. Without loss of generality we may assume that there exists an integer $w \in J_1$ with $|X_w| < w^z m^2$. Let $w + k$ be an integer in J_2 and note that $\tilde{m} - w \leq k < \tilde{m} - w + m^{3z+7} \leq 2m^{3z+7}$. Our goal then is to prove that $|X_{w+k}| > (w+k)^z m^2$, but we first need a few technical lemmas.

Lemma 8. *For all $m \geq 2$ we have $z = \pi(m-1) < \left(\frac{m}{\log(m)}\right) \min\left(1.25, 1 + \frac{3}{2\log(m)}\right)$.*

Lemma 9. *For all $m \geq 1$ we have $m^{3z+7} < e^{5m}$.*

Lemma 10. *For all $k \in \mathbb{N}$ with $w + k \in J_2$ we have the lower bound* $\left|\sum_{i=w+1}^{w+k} \frac{r_i}{i}\right| \geq \frac{1}{(w+k)^k}$.

Lemma 11. *For all $k \in \mathbb{N}$ with $w + k \in J_2$ the inequality* $\frac{2^{\frac{w+k}{t}-2}}{(w+k)^k} - (w+k)^k w^{z+1} > (w+k)^{z+1}$ *holds.*

Proof of Lemma 8. These are the statements of Theorem 1 and Corollary 2 of [7, p. 69]. □

Proof of Lemma 9. For $m < 16$ this can be checked by hand or computer. for $m \geq 16$ we get $7 \leq z$ and Lemma 8 gives $m^{3z+7} \leq m^{4z} < e^{5m}$. □

Proof of Lemma 10. The sum $\sum_{i=w+1}^{w+k} \frac{r_i}{i}$ can be written as a fraction with denominator equal to the least common multiple of $w+1, w+2, .., w+k$, which is at most their product, which is trivially upper bounded by $(w+k)^k$. So to prove that the estimate we want to show holds, it suffices to show that the left-hand side is non-zero. Note that $\tilde{m} \leq w+k < \tilde{m} + m^{3z+7} < \tilde{m} + e^{5m} < \tilde{m} + q_0$. So if we multiply $\sum_{i=w+1}^{w+k} \frac{r_i}{i}$ by the least common multiple of $w+1, w+2, .., w+k$, then every term in the sum is divisible by q_0, except for the term corresponding to $i = \tilde{m}$. The term corresponding to $i = \tilde{m}$ is not divisible by q_0 as $0 < |r_{\tilde{m}}| < q_0$. Since the sum is then not divisible by q_0, it is certainly non-zero, which means $\sum_{i=w+1}^{w+k} \frac{r_i}{i}$ is non-zero as well. □

Proof of Lemma 11. We calculate, using the inequalities $(w+k)^{\frac{1}{4}} > 7t\log(w+k)$ and $4.1(w+k)^{\frac{1}{4}} \geq 4.1e^{5m} > 4e^{5m} + m > 2k + z + 1$, which both follow from the fact that $w + k \geq e^{6.7m} \geq e^{26.8}$.

$$
\begin{aligned}
w + k &= (w+k)^{\frac{1}{4}}(w+k)^{\frac{3}{4}} \\
&> 7t\log(w+k)(w+k)^{\frac{3}{4}} \\
&> 3t + 6t\log(w+k)(w+k)^{\frac{3}{4}} \\
&> 3t + \frac{4.1t\log(w+k)(w+k)^{\frac{3}{4}}}{\log(2)}
\end{aligned}
$$

When we subtract $2t$ from both sides, divide by t and then take 2 to the power of both sides, we obtain:

$$
\begin{aligned}
2^{\frac{w+k}{t}-2} &> 2(w+k)^{4.1(w+k)^{\frac{3}{4}}} \\
&> (w+k)^{4.1(w+k)^{\frac{3}{4}}} + (w+k)^{4.1(w+k)^{\frac{3}{4}}} \\
&> (w+k)^{2k+z+1} + (w+k)^{k+z+1} \\
&> (w+k)^{2k}w^{z+1} + (w+k)^{k+z+1}
\end{aligned}
$$

Dividing by $(w+k)^k$ and rearranging gives the desired inequality. □

Combining all these lemmas lets us finish the proof that $|X_{w+k}| > (w+k)^{z+1}$ for an arbitrary integer $w + k \in J_2$. We calculate:

$$
\begin{aligned}
|X_{w+k}| &= \left| L_{w+k} \sum_{i=1}^{w+k} \frac{r_i}{i} \right| \\
&= \left| \frac{L_{w+k}}{L_w} X_w + L_{w+k} \sum_{i=w+1}^{w+k} \frac{r_i}{i} \right| \\
&\geq L_{w+k} \left| \sum_{i=w+1}^{w+k} \frac{r_i}{i} \right| - \frac{L_{w+k}}{L_w} |X_w| \\
&\geq \frac{2^{\frac{w+k}{t}-2}}{(w+k)^k} - (w+k)^k w^{z+1} \\
&\geq (w+k)^{z+1}
\end{aligned}
$$

Since $n \in I$ implies $n \geq \tilde{m} - m^{3z+7} > e^{7.6m} - e^{5m} > m^2$, we conclude the following:

Lemma 12. *Either* $|X_n| > m^2 n^z$ *for all* $n \in J_1 = [\tilde{m} - m^{3z+7}, \tilde{m})$ *or* $|X_n| > m^2 n^z$ *for all* $n \in J_2 = [\tilde{m}, \tilde{m} + m^{3z+7})$.

2.5 Large prime divisors exist

With the notation of Lemma 12, set $I_0 = J_1$ if $|X_n| \geq m^2 n^z$ holds true for all $n \in J_1$, or else set $I_0 = J_2$. This section will then be devoted to proving the following theorem.

Theorem 3. *There exists an integer* $n \in I_0$ *for which* X_n *is divisible by a prime larger than or equal to* m.

The idea behind the proof of Theorem 3 is that for every prime $p < m$ we search for integers n for which the power of p that divides X_n is small. If we find an n that simultaneously works for all $p < m$, we must conclude that X_n must have a prime divisor larger than or equal to m. For this idea to work, we need to take special care of the primes $p < m$ for which $r_{ip^{e(t)}} = 0$ for all i, where $e(t) = e_p(t)$ is the largest power of p that divides t. It turns out that when p is such a prime, then there is an infinite arithmetic progression of n for which the power of p that divides X_n is small.

Recall that c_1 was defined as the smallest positive integer for which $r_{c_1} \neq 0$ and let $\Sigma_1, \Sigma_2, \Sigma_3$ be defined as follows:

1. $\Sigma_1 = \{p : p \geq m\}$
2. $\Sigma_2 = \{p : p < m, \text{ and } r_{ip^{e(t)}} \neq 0 \text{ for some } i\}$
3. $\Sigma_3 = \{p : p < m, \text{ and } r_{ip^{e(t)}} = 0 \text{ for all } i\}$

Note that with these definitions, proving Theorem 3 is equivalent to proving that there exists an $n \in I_0$ and a $p \in \Sigma_1$ for which $p|X_n$. We will prove this by finding an $n \in I_0$ for which the largest divisor of X_n that is composed solely of primes from $\Sigma_2 \cup \Sigma_3$ is strictly smaller than $|X_n|$ and let us start by focusing our attention on the primes from Σ_3. First note that, since $r_{ip^{e(t)}} = 0$ for all i when $p \in \Sigma_3$, p must divide t. Because otherwise, $e(t)$ would, by assumption, equal 0, which would imply $r_i = 0$ for all i. To state and prove the following two lemmas, let us put $f_p = e(t) + e(r_{c_1})$ for the moment.

Lemma 13. *If $p \in \Sigma_3$, then for all $n \in \mathbb{N}$ and all $i \in \mathbb{N}$ with $i + tp^{f_p} \le n$ we have $\frac{L_n r_i}{i} \equiv \frac{L_n r_{i+tp^{f_p}}}{i + tp^{f_p}} \pmod{p^{f_p}}$.*

Proof. Clearly, $r_i = r_{i+tp^{f_p}}$. So when $r_i = 0$, Lemma 13 follows immediately. When $r_i \neq 0$, let p^α be the largest power of p that divides i and let p^β be the largest power of p that divides L_n. Moreover, define $L'_n = \frac{L_n}{p^\beta}$, $i' = \frac{i}{p^\alpha}$ and $t' = \frac{t}{p^\alpha}$.

$$\frac{L_n r_i}{i} - \frac{L_n r_{i+tp^{f_p}}}{i + tp^{f_p}} = p^{\beta-\alpha}\left(\frac{L'_n r_i}{i'} - \frac{L'_n r_{i+tp^{f_p}}}{i' + t'p^{f_p}}\right)$$

Now, $i' \equiv i' + t'p^{f_p} \pmod{p^{f_p}}$ and this residue class is invertible, since p does not divide i'. Let i^* be its inverse. We then get that the right-hand side of the above equation is congruent to $p^{\beta-\alpha}\left(L'_n r_i i^* - L'_n r_i i^*\right) \equiv 0 \pmod{p^{f_p}}$.

□

For $p \in \Sigma_3$ we can use Lemma 13 to bound the largest power of p that divides X_n, if n lies in a certain residue class.

Lemma 14. *If $p \in \Sigma_3$, then for all $n \in \mathbb{N}$ with $n \equiv c_1 \pmod{t^3 r_{c_1}^2}$, we have that p^{f_p} does not divide X_n.*

Proof. If we let $n \equiv c_1 \pmod{t^3 r_{c_1}^2}$, then a non-negative integer c_2 exists such that $n = c_1 + c_2 tp^{2f_p}$.

By Lemma 13 we know that $\frac{L_n r_i}{i}$ and $\frac{L_n r_{i+tp^{f_p}}}{i + tp^{f_p}}$ differ by a multiple of p^{f_p}. We can use this to split up the sum $\sum_{i=1}^{n} \frac{L_n r_i}{i}$ into parts that are all congruent modulo p^{f_p}. Writing $x_j = c_1 + jtp^{f_p}$, this yields:

$$
\begin{aligned}
X_n &= \sum_{i=1}^{n} \frac{L_n r_i}{i} \\
&= \frac{L_n r_{c_1}}{c_1} + \sum_{j=0}^{c_2 p^{f_p}-1} \sum_{i=x_j+1}^{x_{j+1}} \frac{L_n r_i}{i} \\
&\equiv \frac{L_n r_{c_1}}{c_1} + c_2 p^{f_p} \sum_{i=x_0+1}^{x_1} \frac{L_n r_i}{i} && \pmod{p^{f_p}} \\
&\equiv \frac{L_n r_{c_1}}{c_1} && \pmod{p^{f_p}} \\
&\not\equiv 0 && \pmod{p^{f_p}}
\end{aligned}
$$

□

Remark: Lemma 14 implies that for $n \equiv c_1 \pmod{t^3 r_{c_1}^2}$, the largest divisor of X_n composed solely of primes from Σ_3 is at most:

$$
\begin{aligned}
&\prod_{p \in \Sigma_3} p^{f_p} \\
&\leq t r_{c_1} \\
&< m^2
\end{aligned}
$$

Assume for the moment that $n \in I_0$ and $n \equiv c_1 \pmod{t^3 r_{c_1}^2}$. Since $|X_n| \geq m^2 n^z$ by Lemma 12 and since the largest divisor of X_n composed solely of primes from Σ_3 is at most m^2, it follows that if the largest divisor of X_n composed solely of primes from Σ_2 is smaller than n^z, then X_n must have a prime divisor from Σ_1, which is exactly what we want.

So let $p_1 < p_2 < \ldots < p_y < m$ be the sequence of primes in Σ_2 with $y \leq z$ and let $e_i(x)$ denote the largest power of p_i that divides x, where $e_i(0) = \infty$. With this notation, $p_1^{e_1(X_n)} \cdots p_y^{e_y(X_n)}$ is the prime decomposition of the largest divisor $d(n)$ of X_n which consists only of primes contained in Σ_2. The goal is to find an n with $d(n) < n^z$. We define m_0 to be the smallest integer in I_0 that is congruent to $c_1 \pmod{t^3 r_{c_1}^2}$ and note that such an integer $m_0 \in I_0$ exists, since $|I_0| = m^{3z+7} > m^5 > t^3 r_{c_1}^2$. Moreover, if m' is the smallest integer in I_0, then $m_0 \leq m' + t^3 r_{c_1}^2$.

We shall then construct a sequence $m_0 = n_1 < n_2 < \ldots < n_{y+1}$ of integers lying in I_0, such that all these n_j are congruent to $c_1 \pmod{t^3 r_{c_1}^2}$ and such that either $d(n_j) < n_j^y \leq n_j^z$ for some j with $1 \leq j \leq y$, or for n_{y+1} we have that $p_i^{e_i(X_{n_{y+1}})} < n_{y+1}$ holds for all i with $1 \leq i \leq y$, implying $d(n_{y+1}) < n_{y+1}^y \leq n_{y+1}^z$.

Proof of Theorem 3. To start off, choose $n_1 = m_0$. Now, once we have defined n_j for some j with $1 \le j \le y$, if $d(n_j) < n_j^y$, we are done, Theorem 3 is proved and we can stop. So for the rest of this proof we are free to assume that $d(n_j) \ge n_j^y$ holds for all j with $1 \le j \le y$. This implies in particular that there exists a $\sigma(j) \in \{1, 2, \dots, y\}$ such that for $p_{\sigma(j)}$ we have $p_{\sigma(j)}^{e_{\sigma(j)}(X_{n_j})} \ge n_j$.[2] Then let $p_{\sigma(j)}^{k_j}$ be the largest power of $p_{\sigma(j)}$ smaller than $m^{3y+8-3j}$, set $\widetilde{n}_{j+1}$ equal to the smallest integer larger than n_j such that $e_{\sigma(j)}(\widetilde{n}_{j+1}) - e_{\sigma(j)}(r_{\widetilde{n}_{j+1}}) \ge k_j$ and set n_{j+1} equal to the smallest integer greater than or equal to $\widetilde{n}_{j+1}$ that lies in the residue class $c_1 \pmod{t^3 r_{c_1}^2}$. Further, define the half-open interval $I_j = [n_{j+1}, n_{j+1} + p_{\sigma(j)}^{k_j} - m^5)$.

Now, before we go on, notice that the smallest integer in I_j is n_{j+1} and not n_j. Although this might look confusing at first, think about it as follows. We know that n_j is such that X_{n_j} is divisible by a large power of $p_{\sigma(j)}$. And $\widetilde{n}_{j+1}$ is then defined as also being divisible by some large (suitable) power of that same prime $p_{\sigma(j)}$. So the j in I_j does not signal that I_j contains n_j (because it does not), but it means that I_j has something to do with $p_{\sigma(j)}$. And indeed, we shall shortly prove that for integers $n \in I_j$, the largest power of $p_{\sigma(j)}$ that divides X_n is small. But first we will prove a crucial property of the intervals I_j we defined.

Lemma 15. *The intervals I_j form a decreasing sequence. That is, $I_0 \supset I_1 \supset I_2 \supset \ldots \supset I_y$.*

Proof. Since $I_j = [n_{j+1}, n_{j+1} + p_{\sigma(j)}^{k_j} - m^5)$ and m^5 is just a constant independent of j, we note that the statement $I_{j-1} \supset I_j$ for $j \ge 2$ is equivalent to the two inequalities

$$n_j \le n_{j+1}$$
$$n_{j+1} + p_{\sigma(j)}^{k_j} \le n_j + p_{\sigma(j-1)}^{k_{j-1}}$$

While for $I_0 \supset I_1$ the second inequality gets replaced by $n_2 + p_{\sigma(1)}^{k_1} - m^5 \le m' + m^{3z+7}$, where m' is the smallest integer in I_0. And since $n_1 = m_0 \le m' + m^5$, for $I_0 \supset I_1$ it suffices to prove $n_2 + p_{\sigma(1)}^{k_1} \le n_1 + m^{3z+7}$.

So we would like to get some upper and lower bounds on n_{j+1} and $p_{\sigma(j)}^{k_j}$, and all we need to use are their definitions. First of all, as n_{j+1} is defined as the smallest integer *larger* than or equal to $\widetilde{n}_{j+1}$ for which something holds, while $\widetilde{n}_{j+1}$ is defined as the smallest integer *larger* than or equal to n_j with some property, the inequality $n_j \le n_{j+1}$ is trivial.

Secondly, for an upper bound on n_{j+1}, we need a small lemma.

[2] Of course, there can be more than one such prime. Just pick, say, the smallest.

Lemma 16. *If $p \notin \Sigma_3$ and $A \in \mathbb{N}$ is such that* $\gcd(A,t) = p^{e(t)}$*, then for every $i \in \mathbb{N}$, there is an $i' \in \{iA, (i+1)A, \ldots, (i + \frac{t}{p^{e(t)}} - 1)A\}$ for which $r_{i'} \neq 0$.*

Proof. There are exactly $\frac{t}{p^{e(t)}}$ distinct residue classes $i' \pmod t$ that are divisible by $p^{e(t)}$, and all of them are represented in $\{iA, (i+1)A, \ldots, (i+\frac{t}{p^{e(t)}}-1)A\}$, since $i_1 A \equiv i_2 A \pmod t$ implies $i_1 \equiv i_2 \pmod{\frac{t}{p^{e(t)}}}$. For at least one of them we must have $r_{i'} \neq 0$ by the fact that $p \notin \Sigma_3$. □

Now, the inequality $n_{j+1} \leq \widetilde{n}_{j+1} + m^5$ is trivial, by definition of n_{j+1}. Furthermore, recall that $\widetilde{n}_{j+1}$ is such that $e_{\sigma(j)}(\widetilde{n}_{j+1}) - e_{\sigma(j)}(r_{\widetilde{n}_{j+1}}) \geq k_j$. And because $e_{\sigma(j)}(r_{\widetilde{n}_{j+1}}) \leq \left\lfloor \frac{\log(m-1)}{\log(p_{\sigma(j)})} \right\rfloor$ if $r_{\widetilde{n}_{j+1}} \neq 0$, we have that $\widetilde{n}_{j+1}$ is smaller than or equal to g, where g is defined as the smallest integer larger than n_j for which $e_{\sigma(j)}(g) \geq \left\lfloor \frac{\log(m-1)}{\log(p_{\sigma(j)})} \right\rfloor + k_j$ and $r_g \neq 0$. Let x be equal to $\left\lfloor \frac{\log(m-1)}{\log(p_{\sigma(j)})} \right\rfloor + k_j$ and let us call an integer h with $e_{\sigma(j)}(h) \geq x$ and $r_h \neq 0$ *good* for the moment. By Lemma 16, for every $i \in \mathbb{N}$ there is an $i' \in \{ip^x_{\sigma(j)}, (i+1)p^x_{\sigma(j)}, \ldots, (i+m-1)p^x_{\sigma(j)}\}$ such that $r_{i'} \neq 0$. This implies that we can find a good integer in every interval of $mp^x_{\sigma(j)} \leq (m-1)mp^{k_j}_{\sigma(j)}$ consecutive integers. In conclusion we can say that $n_{j+1} \leq \widetilde{n}_{j+1} + m^5 \leq n_j + (m-1)mp^{k_j}_{\sigma(j)} + m^5$.

Lastly, we look for bounds on $p^{k_j}_{\sigma(j)}$. Again we have a trivial bound $p^{k_j}_{\sigma(j)} < m^{3y+8-3j}$ because $p^{k_j}_{\sigma(j)}$ is defined as the largest power of $p_{\sigma(j)}$ *smaller* than $m^{3y+8-3j}$. On the other hand, there is always a power of $p_{\sigma(j)}$ between two consecutive powers of m since $p_{\sigma(j)} < m$. So $p^{k_j}_{\sigma(j)}$ must be larger than $m^{3y+7-3j}$. Putting all these inequalities together we can prove $I_{j-1} \supset I_j$, for all $j \in \{2, \ldots, y\}$:

$$\begin{aligned} n_{j+1} + p^{k_j}_{\sigma(j)} &< n_j + (m-1)mp^{k_j}_{\sigma(j)} + m^5 + p^{k_j}_{\sigma(j)} \\ &< n_j + (m-1)mp^{k_j}_{\sigma(j)} + 2p^{k_j}_{\sigma(j)} \\ &\leq n_j + m^2 p^{k_j}_{\sigma(j)} \\ &< n_j + m^{3y+10-3j} \\ &= n_j + m^{3y+7-3(j-1)} \\ &< n_j + p^{k_{j-1}}_{\sigma(j-1)} \end{aligned} \tag{1}$$

To prove $I_0 \supset I_1$, use the above reasoning up to and including (1) with $j = 1$. □

Lemma 17. *For all $n \in I_j$ we have $p^{e_{\sigma(j)}(X_n)}_{\sigma(j)} < n$.*

Proof. Let n be any integer in $I_j \subset I_0$ and we will take a look at X_n. Let us first note that $k_j \leq e_{\sigma(j)}(L_{\widetilde{n}_{j+1}}) = e_{\sigma(j)}(L_n)$, where the first inequality follows from the fact that $L_{\widetilde{n}_{j+1}}$ is divisible by $\widetilde{n}_{j+1}$ and therefore by $p_{\sigma(j)}^{k_j}$, and the equality follows since $n < \widetilde{n}_{j+1} + p_{\sigma(j)}^{k_j}$. Now let us dissect X_n;

$$\begin{aligned} X_n &= L_n \sum_{i=1}^{n} \frac{r_i}{i} \\ &= \underbrace{\sum_{i=1}^{n_j} \frac{L_n r_i}{i}}_{S_1} + \underbrace{\sum_{i=n_j+1}^{\widetilde{n}_{j+1}-1} \frac{L_n r_i}{i}}_{S_2} + \frac{L_n r_{\widetilde{n}_{j+1}}}{\widetilde{n}_{j+1}} + \underbrace{\sum_{i=\widetilde{n}_{j+1}+1}^{n} \frac{L_n r_i}{i}}_{S_3} \end{aligned}$$

Using Lemma 9 and the definitions of $\tilde{m}$, J_1 and J_2 from Lemma 12, we have the inequalities $n_j > e^{6.7m} - e^{5m} > e^{5m} > |I_0| \geq n - n_j$, and this implies $n_j > \frac{n}{2}$. Now, by assumption, S_1 is divisible by a power of $p_{\sigma(j)}$ that is at least as large as $n_j > \frac{n}{2}$, hence we obtain $e_{\sigma(j)}(S_1) > e_{\sigma(j)}(L_n) - 1 \geq e_{\sigma(j)}(L_n) - k_j$.

By the definition of $\widetilde{n}_{j+1}$ we know that for every i in S_2 we have $e_{\sigma(j)}(\frac{L_n r_i}{i}) \geq e_{\sigma(j)}(L_n) - k_j + 1$, while $e_{\sigma(j)}(\frac{L_n r_{\widetilde{n}_{j+1}}}{\widetilde{n}_{j+1}}) \leq e_{\sigma(j)}(L_n) - k_j$.

Finally, since $e_{\sigma(j)}(\widetilde{n}_{j+1}) \geq k_j$ and $n < \widetilde{n}_{j+1} + p_{\sigma(j)}^{k_j}$, we have that for every i in S_3 that $e_{\sigma(j)}(i) < k_j$ and so $e_{\sigma(j)}(\frac{L_n r_i}{i}) \geq e_{\sigma(j)}(L_n) - k_j + 1$ for all i in S_3.

Combining the above estimates we see that there is exactly one term in the sum for X_n that is not divisible by $p_{\sigma(j)}^{e_{\sigma(j)}(L_n)-k_j+1}$, and we conclude that the largest power of $p_{\sigma(j)}$ that divides X_n is at most $p_{\sigma(j)}^{e_{\sigma(j)}(L_n)-k_j} < n$. □

Now we may finish the proof of Theorem 3. First off, all the $p_{\sigma(j)}$ have to be distinct, since $p_{\sigma(i)}^{e_{\sigma(i)}(X_{n_i})} > n_i$, while Lemma 17 shows that if $i > j$, then for all $n \in I_{i-1} \subset I_j$ it holds true that $p_{\sigma(j)}^{e_{\sigma(j)}(X_{n_j})} < n$. In other words, $(\sigma(1), \sigma(2), \ldots, \sigma(y))$ is a permutation of $(1, 2, \ldots, y)$. Secondly, since our intervals form a nesting sequence, for $n_{y+1} \in I_y \subset I_j$ we have $p_{\sigma(j)}^{e_{\sigma(j)}(X_{n_{y+1}})} < n_{y+1}$ for all j with $1 \leq j \leq y$. We conclude that $d(n_{y+1}) = \prod_{j=1}^{y} p_j^{e_j(X_{n_{y+1}})} = \prod_{j=1}^{y} p_{\sigma(j)}^{e_{\sigma(j)}(X_{n_{y+1}})} < \prod_{j=1}^{y} n_{y+1} = n_{y+1}^{y}$ and the theorem is proved.

□

2.6 Explicit bounds for non-zero sequences and Dirichlet characters

Let p be a prime bigger than m and let $n = lp^k$ be the smallest positive integer for which X_n is divisible by p. If we could force $\gcd(l, X_{a,b-1})$ to be smaller than p (as is the condition in Theorem 1), then we can straightaway combine Theorems 1 and 3. We claim that this can be done when $r_i \neq 0$ for all i with $\gcd(i, t) = 1$. Because in that case, it is not hard to see that l will always be smaller than p, so the condition $\gcd(l, X_{a,b-1}) < p$ is fulfilled automatically. Indeed, by Lemma 1 p^k exactly divides L_n. But if $l > p$, then $n = lp^k > p^{k+1}$, while $r_{p^{k+1}} \neq 0$, so p^{k+1} should divide L_n as well; contradiction.

Recall that in the proofs of Lemma 12 and Theorem 3, I_0 was a sub-interval of I, where I was defined as $[\tilde{m} - 2m^{3z+7}, \tilde{m} + m^{3z+7})$ and $\tilde{m}$ was the smallest integer larger than $e^{6.7m}$ with $\tilde{m} \equiv c_1 \pmod{t}$ and such that $\tilde{m}$ has a prime divisor q_0 larger than e^{5m}. To find an upper bound on $\tilde{m}$ we use Theorem 1 in [9] which immediately implies that $\tilde{m} < e^{6.7m} + m + me^{5m}$. Now Theorem 3 gives us an integer $n < e^{6.7m} + m + me^{5m} + e^{5m} < e^{6.71m}$ for which X_n is divisible by a prime p larger than m. Since p divides X_n, and $L_n < e^{1.04n}$ by Theorem 12 in [7, p. 71], we can find an upper bound on p.

$$
\begin{aligned}
p &\leq |X_n| \\
&= L_n \left| \sum_{i=1}^{n} \frac{r_i}{i} \right| \\
&\leq L_n \sum_{i=1}^{n} \frac{|r_i|}{i} \\
&< 3m \log(n) L_n \\
&< 21m^2 e^{1.04 e^{6.71m}} \\
&< e^{e^{6.72m}}
\end{aligned}
$$

Now we can bound the quantity $\max(a-1, 2t-1)lp^\lambda$ that appears in Theorem 1 as follows:

$$
\begin{aligned}
\max(a-1, 2t-1)lp^\lambda &< 2amp^{m-1} \\
&< 2ame^{(m-1)e^{6.72m}} \\
&< ae^{e^{7m}}
\end{aligned}
$$

And in conclusion we may say the following.

Theorem 4. *If $r_i \neq 0$ for all i with $\gcd(i, t) = 1$, then for all a there exists a $b < ca$ for which $v_{a,b} < v_{a,b-1}$, where $c = e^{e^{7m}}$.*

2.7 Bounding prime divisors

We could combine Theorems 1 and 3 in Section 2.6 when $\gcd(i, t) = 1$ implies $r_i \neq 0$, because in that case we always have $l < p$. However, in general this is not true. Consider for example $t = 2$, $r_1 = 0$, $r_2 = 1$. Then $p = 3$ divides $X_4 = 4(\frac{1}{2} + \frac{1}{4}) = 3$ and $n = l = 4 > 3 = p$. Luckily, we do not need $l < p$ to invoke Theorem 1; all we need is $\gcd(l, X_{a,b-1}) < p$. So we need to be able to bound prime divisors of either l or $X_{a,b-1}$. In a sense we have done this already for primes from Σ_3 in Lemma 14. In this section we will do something similar for the primes in Σ_1 and Σ_2 and we will find out how exactly this knowledge can be used in the next section.

Lemma 18. *If $p \in \Sigma_1$, then there exists a positive integer $c_p < m$ with $r_{c_p} \neq 0$, such that for all $k \in \mathbb{N}$ and all $n \in \mathbb{N}$ with $c_p p^{\lambda k} \leq n < (c_p + 1)p^{\lambda k}$, we have that p does not divide X_n.*

Proof. Assume $p \in \Sigma_1$, and set $c_p = c_1 < m$, the smallest positive integer i for which $r_i \neq 0$. If we choose k arbitrary and n such that $c_p p^{\lambda k} \leq n < (c_p+1)p^{\lambda k}$, then in the sum for X_n we have that p divides $\frac{L_n r_i}{i}$, unless $p^{\lambda k}|i$. But there is only one such i for which $r_i \neq 0$, and that is $i = c_p p^{\lambda k}$. Therefore everything except for the term corresponding to that i vanishes modulo p. And we conclude $X_n \equiv \frac{L_n r_{c_p p^{\lambda k}}}{c_p p^{\lambda k}} \not\equiv 0 \pmod{p}$. □

A nice way of looking at the statement of Lemma 18 is that X_n is not divisible by $p \in \Sigma_1$ if, when n is written in base p, the number of digits of n is congruent to $1 \pmod{\lambda}$ and its first digit equals c_p.

Lemma 19. *If $p \in \Sigma_2$, then there exists a positive integer $c_p < m^3$ with $r_{c_p} \neq 0$, such that for all $k \in \mathbb{N}$ and all $n \in \mathbb{N}$ with $c_p p^{\lambda k} \leq n < (c_p + 1)p^{\lambda k}$, we have that the largest power of p that divides X_n is smaller than m^2.*

Proof. The proof of this lemma is in spirit very similar to the proof of Lemma 18, but a few complications arise, since the r_i are allowed to be divisible by p. The core idea of choosing c_p in such a way that, once we look modulo a power of p, only one term in the sum for X_n will remain, is however completely the same.

So let us fix $p \in \Sigma_2$ for this proof and define sequences $\{e_i\}_{i\in\mathbb{N}}$ and $\{f_i\}_{i\in\mathbb{N}}$ such that p^{e_i} exactly divides i and p^{f_i} exactly divides r_i. We take $f_i = \infty$ if $r_i = 0$. Recall that $e(t)$ is such that $p^{e(t)}$ divides t exactly, $\mu_p = \left\lfloor \frac{\log(m)}{\log(p)} \right\rfloor$ and note that this implies that $f_i \leq \mu_p$ for all i for which $r_i \neq 0$. Now we define c_p to be the smallest integer i for which the maximum of $e_i - f_i$ is attained, where i runs from 1 to $tp^{2\mu_p}$. That is, $\max\limits_{1\leq i\leq tp^{2\mu_p}} (e_i - f_i) \leq e_{c_p} - f_{c_p}$, with strict inequality for all $i < c_p$. Clearly $c_p \leq tp^{2\mu_p} < m^3$ and we aim to show that this c_p will do the trick.

Firstly we can find a lower bound on $e_{c_p} - f_{c_p}$ as follows. By Lemma 16 we have that for at least one value of $i \in \{1, 2, \ldots, t\}$ it must hold that $r_{ip^{2\mu_p}} \neq 0$ and using this i we have $e_{c_p} - f_{c_p} \geq e_{ip^{2\mu_p}} - f_{ip^{2\mu_p}} \geq 2\mu_p - \mu_p = \mu_p \geq e(t)$. Since $e_{c_p} - f_{c_p} \geq \mu_p$ is positive, $f_{c_p} \neq \infty$, so as promised $r_{c_p} \neq 0$.

Now let us take a look at $X_{c_p} \pmod{p^{x+1+f_{c_p}-e_{c_p}}}$, where x is such that p^x divides L_{c_p} exactly. We claim that the only term $\frac{L_{c_p} r_i}{i}$ in the sum for X_{c_p} which is not divisible by $p^{x+1+f_{c_p}-e_{c_p}}$ is the term corresponding to $i = c_p$. Indeed, by the definition of c_p, for all $i < c_p$ we have that $f_i - e_i \geq 1 + f_{c_p} - e_{c_p}$, implying that the largest power of p that divides $\frac{L_{c_p} r_i}{i}$ will be $p^{x+f_i-e_i} \geq p^{x+1+f_{c_p}-e_{c_p}}$. On the other hand, the largest power of p that divides $\frac{L_{c_p} r_{c_p}}{c_p}$ is equal to $p^{x+f_{c_p}-e_{c_p}}$. In conclusion we can say that $X_{c_p} \not\equiv 0 \pmod{p^{x+1+f_{c_p}-e_{c_p}}}$.

Let k now be given and let n be such that $c_p p^{\lambda k} \leq n < (c_p+1)p^{\lambda k}$. Then let us take a look at $X_n \pmod{p^{x+1+f_{c_p}-e_{c_p}}}$ this time and note that L_n will now be exactly divisible by $p^{x+\lambda k}$. Again we claim that only one term $\frac{L_n r_i}{i}$ does not vanish modulo $p^{x+1+f_{c_p}-e_{c_p}}$, namely the term corresponding to $i = c_p p^{\lambda k}$. This would give us that X_n is not divisible by $p^{x+1+f_{c_p}-e_{c_p}}$.

If $i = c_p p^{\lambda k}$, then $e_i = e_{c_p} + \lambda k$ and we see that the largest power of p that divides $\frac{L_n r_i}{i}$ equals $p^{x+\lambda k+f_i-e_i} = p^{x+f_{c_p}-e_{c_p}}$, if we can show that $f_i = f_{c_p p^{\lambda k}} = f_{c_p}$, which we will do down below.

When $i \leq n$ is different from $c_p p^{\lambda k}$, the largest power of p that divides $\frac{L_n r_i}{i}$ will still be $p^{x+\lambda k+f_i-e_i}$. If this is to be at most $p^{x+f_{c_p}-e_{c_p}}$, then $x + \lambda k + f_i - e_i \leq x + f_{c_p} - e_{c_p}$ or, equivalently, $e_i - \lambda k - f_i \geq e_{c_p} - f_{c_p}$, which will lead to a contradiction. Since the right-hand side of this inequality is at least $e(t)$, we can define $j = ip^{-\lambda k} < c_p$ and note that $p^{e(t)}|j$. Clearly $e_j = e_i - \lambda k$ and we claim that $f_i = f_j$, analogous to the case where $i = c_p p^{\lambda k}$. Assuming this claim for the moment, we then have $e_j - f_j = e_i - \lambda k - f_i \geq e_{c_p} - f_{c_p}$, contrary to our definition of c_p.

To prove our claim that f_i and f_j are equal, it suffices to show that $r_j = r_i = r_{jp^{\lambda k}}$. Or in other words, $j \equiv jp^{\lambda k} \pmod{t}$. But this is equivalent to $(jp^{-e(t)}) \equiv (jp^{-e(t)})p^{\lambda k} \pmod{tp^{-e(t)}}$, which is true as $p^{\lambda k} \equiv 1 \pmod{tp^{-e(t)}}$ by the property of the Carmichael function that $d|t$ implies $\lambda(d)|\lambda(t)$.

We have shown that X_n is not divisible by $p^{x+1+f_{c_p}-e_{c_p}}$, but what is x? Well, p^x is the largest power of p that divides L_{c_p}, so $p^x \leq c_p \leq tp^{2\mu_p}$. Therefore X_n is at most divisible by $p^{x+f_{c_p}-e_{c_p}} \leq p^{x-\mu_p} \leq tp^{\mu_p} < m^2$.

□

2.8 Diophantine approximation to the rescue

Like we mentioned in the previous section, we do not need $l < p$ to invoke Theorem 1; all we need is $\gcd(l, X_{a,b-1}) < p$. The first thing we should note is that we can prove the weaker estimate $l < pt$, since by Lemma 16 there exists at least one $i \in \{1, 2, \ldots, t\}$ with $r_{ip^{k+1}} \neq 0$. This implies that $\frac{L_n r_n}{lp^k} \equiv 0 \pmod{p}$ if $l > pi$, showing that if $l > pt \geq pi$, then $n = lp^k$ cannot be the smallest integer for which $p|X_n$. So from now on we may safely assume $p < l < pt$. Now, if we have a large prime power divisor q^y of l such that the largest power of q that divides $X_{a,b-1}$ is smaller than $\frac{q^y}{t}$, then $\gcd(l, X_{a,b-1}) < \frac{l}{t} < p$. So this will be our plan; first make sure that l is large enough so that it has a large prime power divisor, and then we define b such that $X_{a,b-1}$ is only divisible by a small power of that prime.

The first part of the plan is the easy part, because we have already done all the hard work in previous sections! Indeed, the m that appears in the statement and proof of Theorem 3 can be taken to be any integer bigger than $\max(r, t)$, because that is the only property of m that we used. In other words, we can guarantee the existence of an integer $n = lp^k$ for which $p|X_n$, where p is some prime larger than or equal to M and where M is any arbitrarily large integer. We can use Lemmas 18 and 19 to then smartly choose this M.

In fact, we can choose $M = \lfloor e^{3.4m} \rfloor$, so that if $l \geq p+1 \geq M+1 > e^{3.4m}$, then there are two possibilities; the first possibility is that l has a prime divisor $q \geq m$. In that case we claim that we are set if we can choose a large enough $b = np^{\lambda k_1}$ such that, for some k_2, we have $c_q q^{\lambda k_2} \leq b-1 < (c_q+1)q^{\lambda k_2}$. Indeed, Lemma 18 implies that q does not divide X_{b-1}, so $\gcd(l, X_{b-1}) \leq \frac{l}{q} \leq \frac{l}{m} < p$ and we will prove shortly that this is sufficient.

The other possibility is that l only has prime divisors smaller than m. In that case we claim that there must be a prime q such that if q^y exactly divides l, then $q^y > m^3$.

Lemma 20. *If for every prime power divisor q^y of l we have $q < m$ and $q \leq m^3$, then $l < e^{3.4m}$.*

Proof. Let l be an integer such that for all prime power divisors p^y of l we have $p < m$ and $p^y \leq m^3$. Then $l \leq \prod_{p<m} p^{\left\lfloor \frac{\log(m^3)}{\log(p)} \right\rfloor}$ and with a computer one can check that for $m < 10^5$, this product is smaller than $e^{3.4m}$. For $m \geq 10^5$, we can bound l as follows:

$$
\begin{aligned}
l &\leq \prod_{p<m} p^{\left\lfloor \frac{\log(m^3)}{\log(p)} \right\rfloor} \\
&= \prod_{p<m^{\frac{3}{4}}} p^{\left\lfloor \frac{\log(m^3)}{\log(p)} \right\rfloor} \prod_{m^{\frac{3}{4}} \leq p<m} p^3 \\
&< \left(m^3\right)^{\pi(m^{\frac{3}{4}})} \prod_{p<m} p^3 \\
&< \left(m^3\right)^{\left(\frac{1.174m^{\frac{3}{4}}}{\log(m^{\frac{3}{4}})}\right)} e^{3.12m} \\
&= e^{4.7m^{\frac{3}{4}}+3.12m} \\
&< e^{3.4m}
\end{aligned}
$$

Where we used Lemma 8 and Theorem 12 in [7, p. 71], and where the last line uses $m \geq 10^5$. □

So let q be a prime divisor of l such that $q^y > m^3$. First we remark that this implies that $q \in \Sigma_2$, as $q \in \Sigma_3$ would imply $r_n = 0$, which contradicts the assumption that n is the smallest integer for which $p|X_n$. Now we can use Lemma 19 which gives us that if $c_q q^{\lambda k_2} \leq b-1 < (c_q+1)q^{\lambda k_2}$, then the largest power of q that divides X_{b-1}, let us call it q^x for the moment, is smaller than m^2. And then $\gcd(l, X_{b-1}) \leq \frac{lq^x}{q^y} < \frac{lm^2}{m^3} = \frac{l}{m} < p$. In both cases a prime q exists so that with $c_q q^{\lambda k_2} \leq b-1 < (c_q+1)q^{\lambda k_2}$ we guarantee that $\gcd(l, X_{b-1}) < p$.

The astute reader might point out that the condition in Theorem 1 is $\gcd(l, X_{a,b-1}) < p$ which is different from $\gcd(l, X_{b-1}) < p$. So it looks like we only guarantee something about the factorization of X_{b-1} instead of $X_{a,b-1}$. However, we claim that with $b-1 \in [c_q q^{\lambda k_2}, (c_q+1)q^{\lambda k_2})$, if $q^{\lambda k_2} \geq am^2$, then $X_{a,b-1} \equiv X_{b-1} \pmod{q^{x+1}}$, which is non-zero by definition. First of all we claim that $L_{a,b-1}$ is equal to L_{b-1}. On the one hand we trivially have $L_{a,b-1}|L_{b-1}$. For the other direction, since $b-1 \geq am^2 > am$, every integer i smaller than a with $r_i \neq 0$ has a multiple of the form $(jt+1)i$ between a and $b-1$, with $r_{(jt+1)i} = r_i \neq 0$. So if i divides L_{b-1}, it will also divide $L_{a,b-1}$, proving $L_{b-1}|L_{a,b-1}$ and therefore $L_{a,b-1} = L_{b-1}$.

Secondly L_{b-1} is divisible by $q^{\lambda k_2}$ since $r_{c_q q^{\lambda k_2}} = r_{c_q} \neq 0$, by Lemmas 18 and 19. Therefore the only terms $\frac{L_{b-1} r_i}{i}$ in the sum for X_{b-1} that are non-zero modulo q^{x+1} are the ones where i is divisible by $q^{\lambda k_2 - x}$. The latter quantity is bigger than a as we assumed $q^{\lambda k_2} \geq am^2$ and $q^x < m^2$. Since all terms that are non-zero modulo q^{x+1} are bigger than a we indeed have $X_{a,b-1} \equiv X_{b-1} \pmod{q^{x+1}}$.

In summary, if $n = lp^k$ with $l > p$ is such that X_n is divisible by a prime $p > e^{3.4m}$ (and n is the smallest positive integer i for which p divides X_i) and q is a prime divisor of l such that either $q \geq m$ or $q^y \geq m^3$, we then have concluded the following.

Lemma 21. *If k_1 and k_2 are positive integers such that with $b = np^{\lambda k_1}$ the string of inequalities $c_q am^2 \leq c_q q^{\lambda k_2} < b < (c_q + 1)q^{\lambda k_2}$ holds, then $v_{a,b} < v_{a,b-1}$.*

The numbers n, l, p, k and q will now all be fixed for the rest of this section. All we have to do is find k_1 and k_2 so that the above string of inequalities is satisfied. By looking at the inequalities $c_q q^{\lambda k_2} < np^{\lambda k_1} < (c_q + 1)q^{\lambda k_2}$ and taking logarithms, we see that we end up with a linear form in logarithms, so it is no wonder that bounds on such linear forms will turn out to be helpful.

Lemma 22. *There exist positive integers b_1 and b_2 with $b_2 < 2\log(q)m^4$ such that*

$$|b_2 \log(p) - b_1 \log(q)| = \epsilon < \frac{1}{2m^4}$$

Proof. Dirichlet's Approximation Theorem states that for any positive real number $\zeta > 0$ and any $N \in \mathbb{N}$, there exist positive integers b_1 and b_2 with $b_2 \leq N$ such that $|b_2\zeta - b_1| < \frac{1}{N+1}$. Now we apply this with $\zeta = \frac{\log(p)}{\log(q)}$ and $N = \lfloor 2\log(q)m^4 \rfloor$ to obtain $|b_2 \frac{\log(p)}{\log(q)} - b_1| < \frac{1}{2\log(q)m^4}$. Multiplying both sides of the inequality by $\log(q)$ finishes the proof. □

Lemma 23. *Let b_1, b_2 and ϵ be as in Lemma 22. Let $\gamma > 0$ be any positive real number and set $C = \lceil \frac{\gamma}{\epsilon} \rceil$. Then, if $b_2 \log(p) - b_1 \log(q) > 0$, we have*

$$0 \leq Cb_2 \log(p) - Cb_1 \log(q) - \gamma < \frac{1}{2m^4}$$

while if $b_2 \log(p) - b_1 \log(q) < 0$, we have

$$\frac{-1}{2m^4} < Cb_2 \log(p) - Cb_1 \log(q) + \gamma \leq 0$$

Proof. Assume $b_2 \log(p) - b_1 \log(q) > 0$. The other case can be proven in an analogous manner. Then, on the one hand:

$$\begin{aligned} Cb_2 \log(p) - Cb_1 \log(q) - \gamma &= C(b_2 \log(p) - b_1 \log(q)) - \gamma \\ &= \left(\left\lceil \frac{\gamma}{\epsilon} \right\rceil\right) \epsilon - \gamma \\ &\geq \left(\frac{\gamma}{\epsilon}\right) \epsilon - \gamma \\ &= 0 \end{aligned}$$

while on the other hand:

$$
\begin{aligned}
Cb_2\log(p) - Cb_1\log(q) - \gamma &= C(b_2\log(p) - b_1\log(q)) - \gamma \\
&= \left(\left\lceil \frac{\gamma}{\epsilon} \right\rceil\right)\epsilon - \gamma \\
&< \left(\frac{\gamma}{\epsilon} + 1\right)\epsilon - \gamma \\
&= \epsilon \\
&< \frac{1}{2m^4}
\end{aligned}
$$

□

Lemma 24. *Let $D \in \mathbb{N}$ be any integer bigger than or equal to $k+2$ and assume that we choose γ in Lemma 23, equal to*

$$\gamma = \pm\left(\frac{\log(c_q) + \log(c_q+1) - 2\log(n)}{2\lambda}\right) + D\log(p)$$

where plus or minus depends on whether $b_2\log(p) - b_1\log(q)$ is positive or negative, respectively. Then $\gamma > 0$ and $c_q q^{\lambda k_2} < np^{\lambda k_1} < (c_q+1)q^{\lambda k_2}$ holds, with $k_2 = Cb_1$ and $k_1 = Cb_2 - D$.

Proof. Let us first prove that γ is positive. We have to consider the cases where $b_2\log(p) - b_1\log(q)$ is bigger or smaller than 0 separately.

Case I. $b_2\log(p) - b_1\log(q) > 0$.

$$
\begin{aligned}
\gamma &= \frac{\log(c_q) + \log(c_q+1) - 2\log(n)}{2\lambda} + D\log(p) \\
&> D\log(p) - \log(n) \\
&= D\log(p) - \log(lp^k) \\
&> D\log(p) - (k+2)\log(p) \\
&\geq 0
\end{aligned}
$$

Case II. $b_2\log(p) - b_1\log(q) < 0$.

$$
\begin{aligned}
\gamma &= \frac{2\log(n) - \log(c_q) - \log(c_q+1)}{2\lambda} + D\log(p) \\
&> D\log(p) - \log(c_q+1) \\
&\geq D\log(p) - \log(m^3) \\
&> 3.4Dm - 3\log(m) \\
&> 0
\end{aligned}
$$

To prove that $c_q q^{\lambda Cb_1} < np^{\lambda(Cb_2 - D)} < (c_q+1)q^{\lambda Cb_1}$ holds, we also have to handle the two cases separately, but these proofs will be completely analogous to each other. So let us only do the first one and leave the second one as exercise for the reader. Assume $b_2 \log(p) - b_1 \log(q) > 0$ and let us first try to find an upper bound for $np^{\lambda k_1}$, taking Lemma 23 as a starting point.

$$\begin{aligned} Cb_2 \log(p) - Cb_1 \log(q) - \gamma &< \frac{1}{2m^4} \\ &< \frac{\log(c_q + 1) - \log(c_q)}{2\lambda} \end{aligned}$$

Here we used $\lambda < m$ and the fact that $\log(x) - \log(x-1) > \frac{1}{x}$ with $x = c_q + 1 \leq m^3$. Now we multiply by λ, apply the definition of γ, rearrange and exponentiate to finish.

$$\begin{aligned} \lambda(Cb_2 - D)\log(p) + \log(n) &< \lambda C b_1 \log(q) + \log(c_q + 1) \\ np^{\lambda(Cb_2 - D)} &< (c_q + 1)q^{\lambda C b_1} \end{aligned}$$

For a lower bound on $np^{\lambda k_1}$, we use similar ideas.

$$\begin{aligned} Cb_2 \log(p) - Cb_1 \log(q) - \gamma &\geq 0 \\ &> \frac{\log(c_q) - \log(c_q + 1)}{2\lambda} \end{aligned}$$

And once more we multiply by λ, use the definition of γ, rearrange and exponentiate.

$$\begin{aligned} \lambda(Cb_2 - D)\log(p) + \log(n) &> \lambda C b_1 \log(q) + \log(c_q) \\ np^{\lambda(Cb_2 - D)} &> c_q q^{\lambda C b_1} \end{aligned}$$

□

Corollary 2. *For every a there are infinitely many b for which $v_{a,b} < v_{a,b-1}$.*

Proof. The only inequality from Lemma 21 that we have not checked yet is the inequality $c_q am^2 \leq c_q q^{\lambda k_2}$. Choose D from Lemma 24 to be any integer bigger than $am^2 + k + 2$. Then:

$$\begin{aligned} c_q q^{\lambda k_2} &= c_q q^{\lambda C b_1} \\ &= c_q q^{\lambda \left\lceil \frac{\gamma}{\epsilon} \right\rceil b_1} \\ &> c_q q^{\gamma} \\ &> c_q q^{am^2} \\ &> c_q am^2 \end{aligned}$$

□

2.9 Explicit bounds for all sequences

We are now in a position to prove our final theorem on upper bounds.

Theorem 5. *For all a there exists a $b < ca$ for which $v_{a,b} < v_{a,b-1}$, where $c = e^{e^{e^{10m}}}$.*

Proof. To find an upper bound for this constant c, let us recall the chain of dependency. We forced p to be larger than or equal to $M = \lfloor e^{3.4m} \rfloor$ and, with m replaced by M, we claim that Theorem 3 and Lemma 12 are still true when $\tilde{m}$ is instead chosen to be the smallest integer larger than $e^{3.9M}$ with $\tilde{m} \equiv c_1 \pmod{t}$ and such that $\tilde{m}$ has a prime divisor q_0 larger than $e^{3.5M}$.

To demonstrate this claim, first note that with the use of Lemma 8 it can be shown that Lemma 9 is still true with m replaced by M, z the number of primes strictly below M, and 5 replaced by 3.5. Indeed, since $M > 10^5$, Lemma 8 implies $3z + 7 < \frac{3.5M}{\log(M)}$, Moreover, with the aforementioned replacements, the statements and proofs of Lemma 10 and 11 still go through verbatim by replacing every occurrence of 5 by 3.5, every occurrence of 6.7 by 3.9, 26.8 by 100 and every occurrence of $\frac{1}{4}$ and $\frac{3}{4}$ by $\frac{1}{10}$ and $\frac{9}{10}$ respectively.

With these slightly improved constants, we obtain $n < e^{3.9M} + M + Me^{3.5M} + e^{3.5M} < e^{4M} < e^{4e^{3.4m}} < e^{4e^{4m}}$, analogously to how we upper bounded n in Section 2.6. For pm we also get via similar reasoning $pm < m(e^{3.9M} + M + Me^{3.5M} + e^{3.5M}) < e^{4e^{3.4m}} < e^{4e^{4m}}$.

Using the bound for pm, our number b can be upper bounded by $(c_q+1)q^{\lambda C b_1} < m^3 q^{mCb_1} < (pm)^{2mCb_1} < e^{8mCb_1 e^{4m}}$, where the first inequality follows from Lemmas 18 and 19 and the second inequality follows from $q \leq l < pt < pm$ as explained at the start of Section 2.8. So we still need to find upper bounds for C and b_1.

As for b_1, Lemma 22 gives us that it is smaller than $\frac{b_2 \log(p)}{\log(q)} + 1 < 2m^4 \log(p) + 1 < 2m^4 \log(pm) < 8m^4 e^{4m} < e^{6m}$. And finally, we would like to find a bound for $C = \lceil \frac{\gamma}{\epsilon} \rceil < (\gamma + 1)\epsilon^{-1}$. We therefore need to bound both γ and ϵ^{-1} and starting with ϵ^{-1}, we use an effective version of Baker's Theorem on a lower bound on linear forms in logarithms.

Lemma 25. *Let b_1, b_2 and ϵ be as in Lemma 22. Then we have the lower bound*

$$\log(\epsilon) = \log\left(|b_2 \log(p) - b_1 \log(q)|\right) > -e^{9.9m}$$

Proof. We need to take a look at Corollary 2 of [8, p. 288] and the notation they use. In their notation, α_2 equals our p, while α_1 is our prime q. Furthermore, b_1 is our b_1 and $b_2 = b_2$. So D, which is defined in Section 2 of that paper as $[\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}]/[\mathbb{R}(\alpha_1, \alpha_2) : \mathbb{R}]$, simply equals 1. We can let $\log(A_1)$ and $\log(A_2)$ be $\log(q)$ and $\log(p)$ respectively, which makes their $b' = \frac{b_1}{D \log(A_2)} + \frac{b_2}{D \log(A_1)}$ in our case bounded by $2\frac{b_2}{\log q}+1 < 4m^4+1$, so that $\log(b')+0.14 < \log(4m^4+1)+0.14 < 6\log(m)$. And now we may apply Corollary 2 of [8];

$$\begin{aligned}
\log(|b_2\log(p)-b_1\log(q)|) &\geq -24.34\left(\max\left\{6\log(m), 21, \frac{1}{2}\right\}\right)^2\log(q)\log(p)\\
&= -876.24\max\left(\log^2(m), 12.25\right)\log(p)\log(q)\\
&> -876.24\max\left(\frac{12.25\log^2(m)}{\log^2(4)}, 12.25\right)\log(pm)\log(pm)\\
&> -5586\log^2(m)\log^2(pm)\\
&> -89376\log^2(m)e^{6.8m}\\
&> -e^{9.9m}
\end{aligned}$$

□

For $\gamma=\gamma_D$, we use its definition as it was given in Lemma 24;

$$\begin{aligned}
\gamma+1 &= 1\pm\left(\frac{\log(c_q)+\log(c_q+1)-2\log(n)}{2\lambda}\right)+D\log(p)\\
&< 1+\max\bigl(\log(c_q+1), \log(n)\bigr)+D\log(p)\\
&< 1+\max\left(\log(m^3), 4e^{4m}\right)+4De^{4m}\\
&= 1+4(D+1)e^{4m}\\
&< De^{5m}
\end{aligned}$$

Here, by Lemma 24 and the proof of Corollary 2, we have to choose D bigger than or equal to $k+2$ and such that $q^{\lambda Cb_1}\geq am^2$, where C depends on γ, which in turn depends on D. If $D=k+2$ already ensures that $q^{\lambda Cb_1}\geq am^2$, then we choose $D=k+2$ and, by using $k<\lambda\leq m-2$ (otherwise $p|X_{n'}$ with $n'=np^{-\lambda}$, contradicting the definition of n), the upper bound on $\gamma+1$ simplifies to $\gamma+1<De^{5m}\leq me^{5m}<e^{6m}$. In this case we have:

$$\begin{aligned}
b(a) &\leq b\\
&< e^{8mCb_1e^{4m}}\\
&< e^{8me^{e^{9.9m}}e^{6m}e^{6m}e^{4m}}\\
&< e^{e^{17m}e^{e^{9.9m}}}\\
&< e^{e^{e^{10m}}}\\
&= c\leq ca
\end{aligned}$$

In the other case we have to choose D bigger than $k+2$ to make sure that $q^{\lambda Cb_1}\geq am^2$ holds. So then we can choose D in such a way that $q^{\lambda Cb_1}=q^{\lambda b_1\lceil\gamma_D\epsilon^{-1}\rceil}\geq am^2>q^{\lambda b_1\lceil\gamma_{D-1}\epsilon^{-1}\rceil}$ and we get:

$$
\begin{aligned}
b(a) &\leq b \\
&< (c_q+1)q^{\lambda b_1 \lceil \gamma_D \epsilon^{-1} \rceil} \\
&= (c_q+1)q^{\lambda b_1 \lceil \gamma_{D-1} \epsilon^{-1} \rceil} q^{\lambda b_1 \left(\lceil \gamma_D \epsilon^{-1} \rceil - \lceil \gamma_{D-1} \epsilon^{-1} \rceil \right)} \\
&< am^5 q^{\lambda b_1 \epsilon^{-1}((\gamma_D+1)-\gamma_{D-1})} \\
&< a(pm)^{6\lambda b_1 \epsilon^{-1}((\gamma_D+1)-\gamma_{D-1})} \\
&< ae^{24\lambda b_1 \epsilon^{-1}((\gamma_D+1)-\gamma_{D-1})e^{4m}} \\
&< ae^{96me^{6m}e^{e^{9.9m}}e^{4m}e^{4m}} \\
&< ae^{e^{16m}e^{e^{9.9m}}} \\
&< ae^{e^{10m}} \\
&= ca
\end{aligned}
$$

□

3 Lower bounds

3.1 A logarithmic lower bound

So far in this paper, we have proven the upper bound $b(a) < ca$, for some constant c. Or, in other words, we can upper bound the difference $b(a) - a$ by a linear function. In this section we will look at lower bounds, and prove that this difference is at least logarithmic. That is, there exists an absolute constant c_1 such that $b(a) > a + c_1 \log(a)$ holds for all large enough a. In Section 3.2 we will then show that this lower bound is close to optimal when $r_i \neq 0$ for all i, as there exists an absolute constant c_2 such that for infinitely many a we have $b(a) < a + c_2 \log(a)$. In Section 3.3 we will then improve upon these constants c_1 and c_2 in the case where $r_i = 1$ for all i.

Theorem 6. $\liminf_{a \to \infty} \dfrac{b(a) - a}{\log a} \geq \dfrac{1}{2}$.

Proof. If $r_b = 0$, then for sure $b \neq b(a)$. So let ϵ be a given small positive constant and let b be any integer with $r_b \neq 0$ such that $a < b < a + (\frac{1}{2} - \epsilon) \log a$. Then we shall see that $v_{a,b} > v_{a,b-1}$, assuming a is large enough. To achieve this, recall that $v_{a,b} = \frac{L_{a,b}}{g_{a,b}}$, where $g_{a,b} = \gcd(L_{a,b}, X_{a,b})$.

So $v_{a,b} > v_{a,b-1}$ precisely when $\frac{L_{a,b}}{L_{a,b-1}} > \frac{g_{a,b}}{g_{a,b-1}}$. This inequality readily follows by combining the following two lemmas.

Lemma 26. $\dfrac{L_{a,b}}{L_{a,b-1}} > b^{1/2+\epsilon+o(1)}$.

Lemma 27. $\dfrac{g_{a,b}}{g_{a,b-1}} < b^{1/2-\epsilon+o(1)}$.

Proof of Lemma 26. A straightforward calculation:

$$
\begin{aligned}
\frac{L_{a,b-1}}{L_{a,b}} &= \frac{L_{a,b-1}}{\operatorname{lcm}(L_{a,b-1}, b)} \\
&= \frac{L_{a,b-1} \gcd(L_{a,b-1}, b)}{b L_{a,b-1}} \\
&= b^{-1} \gcd(L_{a,b-1}, b) \\
&\le b^{-1} \prod_{\substack{p^k \le b-a < p^{k+1} \\ p^k \text{ a prime power}}} \gcd(p^k, b) \\
&\le b^{-1} \prod_{\substack{p^k \le b-a < p^{k+1} \\ p^k \text{ a prime power}}} p^k \\
&\le b^{-1} e^{(\frac{1}{2} - \epsilon + o(1)) \log a} \qquad (2) \\
&= b^{-1} a^{\frac{1}{2} - \epsilon + o(1)} \\
&= b^{-1/2 - \epsilon + o(1)}
\end{aligned}
$$

Where (2) is obtained as a consequence of PNT. □

Proof of Lemma 27. Let p be any prime and let $e(n) = e_p(n)$ denote the largest power of p that divides n. Furthermore, write $b = xyz$ such that for all primes p dividing x we have $e_p(b) > e_p(L_{a,b-1}) + e_p(r_b)$, for all primes p dividing y we have $e_p(L_{a,b-1}) < e_p(b) \le e_p(L_{a,b-1}) + e_p(r_b)$, and $p|z$ implies $e_p(b) \le e_p(L_{a,b-1})$. Then Lemma 26 shows that $xy \ge b^{1/2+\epsilon+o(1)}$, which implies $z \le b^{1/2-\epsilon+o(1)}$.

So if we now fix a prime p, then our goal is to find a good upper bound for $e(g_{a,b})$. We have to distinguish between three different cases.

Case I. $p|x$.
Then $e(L_{a,b}) = e(b) > e(L_{a,b-1}) + e(r_b)$, so

$$
\begin{aligned}
X_{a,b} &= \frac{L_{a,b}}{L_{a,b-1}} X_{a,b-1} + \frac{L_{a,b} r_b}{b} \\
&\equiv \frac{L_{a,b} r_b}{b} \pmod{p^{e(r_b)+1}} \\
&\not\equiv 0 \pmod{p^{e(r_b)+1}}
\end{aligned}
$$

And this implies $e(g_{a,b}) \le e(r_b)$.

Case II. $p|y$.
Let us first remark that the definition of y implies that in this case we have $e(r_b) \ge 1$, or equivalently, $p|r_b$. Moreover, if $i \in [a, b-1]$ is such that $e(i) = e(L_{a,b-1})$, then both i and b are divisible by $p^{e(i)}$, which means $p^{e(i)} \le b - i < (\frac{1}{2} - \epsilon) \log(a) < \log(a)$, hence we conclude:

$$
\begin{aligned}
e(g_{a,b}) &\le e(L_{a,b}) \\
&= e(b) \\
&\le e(L_{a,b-1}) + e(r_b) \\
&= e(i) + e(r_b) \\
&< \frac{\log\log(a)}{\log(p)} + e(r_b)
\end{aligned}
$$

Case III. $p|z$ or $p \nmid b$.
Our goal in this case is to prove the upper bound $e(g_{a,b}) \le e(g_{a,b-1}) + e(z)$, so without loss of generality we may assume that $e(g_{a,b}) > e(g_{a,b-1})$. Note that $e(L_{a,b}) = e(L_{a,b-1})$, so if $e(g_{a,b}) > e(g_{a,b-1})$, then $e(g_{a,b-1}) = e(X_{a,b-1})$. Also note that $e(X_{a,b}) = e\left(\frac{L_{a,b}}{L_{a,b-1}}X_{a,b-1} + \frac{L_{a,b}r_b}{b}\right) > e(X_{a,b-1})$ is only possible when $e\left(\frac{L_{a,b}}{L_{a,b-1}}X_{a,b-1}\right) = e\left(\frac{L_{a,b}r_b}{b}\right)$ while, in this case, $e(b) = e(z)$. Therefore,

$$
\begin{aligned}
e(g_{a,b}) - e(z) &= e(g_{a,b}) - e(b) \\
&\le e(L_{a,b}) + e(r_b) - e(b) \\
&= e\left(\frac{L_{a,b}r_b}{b}\right) \\
&= e\left(\frac{L_{a,b}}{L_{a,b-1}}X_{a,b-1}\right) \\
&= e(X_{a,b-1}) \\
&= e(g_{a,b-1})
\end{aligned}
$$

Let us take all three cases together and calculate.

$$
\begin{aligned}
\frac{g_{a,b}}{g_{a,b-1}} &= \prod_p p^{e_p(g_{a,b})-e_p(g_{a,b-1})} \\
&\le \prod_{p|x} p^{e_p(g_{a,b})} \prod_{p|y} p^{e_p(g_{a,b})} \prod_{p|z \text{ or } p\nmid b} p^{e_p(g_{a,b})-e_p(g_{a,b-1})} \\
&\le \prod_{p|x} p^{e_p(r_b)} \prod_{p|y} p^{\frac{\log\log(a)}{\log(p)}+e_p(r_b)} \prod_{p|z \text{ or } p\nmid b} p^{e_p(z)} \\
&\le \prod_{p|r_b} p^{\frac{\log\log(a)}{\log(p)}+e_p(r_b)} \prod_{p|z} p^{e_p(z)} \\
&= r_b z \prod_{p|r_b} \log(a) \\
&< r_b z \log(a)^{r_b} \\
&\le b^{1/2-\epsilon+o(1)}
\end{aligned}
$$

□

3.2 A sequence with logarithmic difference

As it turns out, the lower bound from the previous section is close to sharp, at least when $r_i \neq 0$ for all i.

Theorem 7. *If $r_i \neq 0$ for all i, then* $\liminf_{a\to\infty} \frac{b(a)-a}{\log a} \leq 2$.

Proof. By our discussion at the start of Section 2.6, when the r_i are non-zero there exists a prime p dividing X_n which is bigger than $\max(|r_1|, |r_2|, \ldots, |r_t|, t)$. Here, $n = lp^k$ is the smallest $i \in \mathbb{N}$ for which p divides X_i and l is such that $1 < l < p$. Since $l < p$, we see that p^k exactly divides L_n, and therefore $\frac{L_n r_i}{i} \not\equiv 0 \pmod{p}$ if, and only if, $p^k | i$. Writing X_n as a sum modulo p, we obtain:

$$\begin{aligned}
X_n &= L_n \sum_{i=1}^{n} \frac{r_i}{i} \\
&\equiv L_n \sum_{i=1}^{l} \frac{r_{ip^k}}{ip^k} && \pmod{p} \\
&\equiv \frac{L_n}{p^k} \sum_{i=1}^{l} \frac{r_{ip^k}}{i} && \pmod{p} \\
&\equiv 0 && \pmod{p}
\end{aligned}$$

Let k_1 now be large enough such that for every invertible residue class $h \pmod{pt}$ we have the Bertrand's Postulate type result that, for the interval $I = (\frac{1}{2}(l-1)p^{\lambda k_1+k}, (l-1)p^{\lambda k_1+k})$, there is a prime $q \in I$ for which $q \equiv h \pmod{pt}$. The existence of such a q is of course guaranteed by PNT.

Assume that the product of all primes in I is congruent to $h \pmod{pt}$ and let $\tilde{q} \in I$ be a prime with $\tilde{q} \equiv h \pmod{pt}$. Then we define $Q = \tilde{q}^{-1} \prod_{q \in I} q$, where the product is taken over all primes $q \in I$. We now set $b = lp^{\lambda k_1+k}Q$ and $a = b - (l-1)p^{\lambda k_1+k} = p^{\lambda k_1+k}(lQ - l + 1)$. Note that $lQ \equiv l \pmod{pt}$ by the definition of $\tilde{q}$.

By PNT we see that $b = \exp[((l-1)p^{\lambda k_1+k})(\frac{1}{2} + o(1))]$, so $b - a = (2 + o(1))\log(b) = (2 + o(1))\log(a)$. Therefore, Theorem 7 would follow if we could prove $v_{a,b} < v_{a,b-1}$, and this is exactly our plan.

To show this, first observe that every prime power divisor of b is smaller than or equal to $(l-1)p^{\lambda k_1+k}$. And since $b - a = (l-1)p^{\lambda k_1+k}$, every prime power divisor of b is also a prime power divisor of some number between a and $b-1$ (inclusive). This implies that $L_{a,b} = L_{a,b-1}$. So to prove $v_{a,b} < v_{a,b-1}$, it suffices to show that $g_{a,b} > g_{a,b-1}$, and our proof will be reminiscent of our proof of Theorem 1.

First, analogous to Lemma 3, we claim that p divides $X_{a,b}$ while p does not divide $X_{a,b-1}$. Since $L_{a,b-1}$ is exactly divisible by $p^{\lambda k_1+k}$, we have that $\frac{L_{a,b-1}r_i}{i} \not\equiv 0 \pmod{p}$ if, and only if, $p^{\lambda k_1+k}|i$. And because $p^{\lambda k_1+k}|b$, we remark $X_{a,b} = X_{a,b-1} + \frac{L_{a,b}r_b}{b} \not\equiv X_{a,b-1} \pmod{p}$ implying that at most one of $X_{a,b-1}$ and $X_{a,b}$ can be divisible by p. Looking at $X_{a,b} \pmod{p}$ now proves our claim;

$$\begin{aligned}
X_{a,b} &= L_{a,b}\sum_{i=a}^{b}\frac{r_i}{i} \\
&\equiv \frac{L_{a,b}}{p^{\lambda k_1+k}}\sum_{i=1}^{l}\frac{r_{p^{\lambda k_1+k}(lQ-l+i)}}{lQ-l+i} && \pmod{p} \\
&\equiv \frac{L_{a,b}}{p^{\lambda k_1+k}}\sum_{i=1}^{l}\frac{r_{ip^k}}{i} && \pmod{p} \\
&\equiv 0 && \pmod{p}
\end{aligned}$$

Secondly, for a prime q dividing Q, q does not divide $X_{a,b-1}$. Indeed, as both $b-2q$ and b are outside the interval $[a, b-1]$, there is only term in the sum for $X_{a,b-1}$ that does not vanish modulo q, and that is the term corresponding to $i = b-q$.

For our final calculation of $g_{a,b}$ we can almost copy our calculation of $g_{a,b}$ at the end of Theorem 1 verbatim, but now with $b = lp^{\lambda k_1+k}Q$, instead of $b = lp^{\lambda k_1+k}$. This results in $g_{a,b} \geq \frac{p}{\gcd(lQ, X_{a,b-1})}g_{a,b-1}$, which proves the theorem as $l < p$ and $\gcd(Q, X_{a,b-1}) = 1$. □

To deal with the problems that arise in the case where some of the r_i are equal to zero, recall that for a prime divisor q of Q, q divides exactly one term in the interval $[a, b-1]$, namely $i = b-q$. Therefore, q is a divisor of $L_{a,b-1}$ but not a divisor of $X_{a,b-1}$. However, when $r_{b-q} = 0$, then q no longer divides $L_{a,b-1}$, which would be a problem because then we would have $L_{a,b} > L_{a,b-1}$. A possible way to work around this is to only choose certain primes q which are congruent to $1 \pmod{t}$, and such that the interval not only contains $b-q$, but $b-2q, b-3q, \ldots, b-dq$ as well, where d is the smallest positive integer for which $r_{b-dq} = r_{n-d} \neq 0$. Because $r_{b-dq} \neq 0$, we guarantee that q divides $L_{a,b-1}$ and since d was the smallest such integer, we still get that there is only one term in the sum for $X_{a,b-1}$ that does not vanish modulo q, implying that q does not divide $X_{a,b-1}$. Furthermore, since $r_{b-tq} = r_b = r_n \neq 0$, we see that d is smaller than or equal to t.

With the usual definitions for n, l, p, k, λ and with d defined as the smallest positive integer for which $r_{n-d} \neq 0$, let k_1 be a large integer and let S be the set of primes q that are contained in the interval $(\frac{1}{d+1}(l-1)p^{\lambda k_1+k}, \frac{1}{d}(l-1)p^{\lambda k_1+k})$

and for which $q \equiv 1 \pmod{t}$. Assume $\prod_{q \in S} q \equiv h \pmod{p}$ and let $\tilde{q} \in S$ be such that $\tilde{q} \equiv h \pmod{p}$. Just like before, such a $\tilde{q}$ exists whenever k_1 is large enough, by PNT.

Now we define $Q = \tilde{q}^{-1} \prod_{q \in S} q$, set $b = lp^{\lambda k_1 + k} Q$ and set $a = b - (l-1)p^{\lambda k_1 + k} = p^{\lambda k_1 + k}(lQ - l + 1)$. By arguments that are analogous and mostly identical to our arguments in the $r_i \neq 0$ case, we obtain $v_{a,b-1} \geq \frac{p}{\gcd(lQ, X_{a,b-1})} v_{a,b} = \frac{p}{\gcd(l, X_{a,b-1})} v_{a,b}$ while, once more applying PNT, we have that the difference $b - a$ has order $(d(d+1)\varphi(t) + o(1)) \log(a)$, where $\varphi(t)$ is Euler's totient function. So if we may assume $l < p$, then we can conclude that we can generalize Theorem 7 to more sequences of r_i, at the cost of increasing our upper bound on the lower limit. For example, as explained at the start of Section 2.6, we know that $l < p$ always holds in the case that $r_i \neq 0$ for all i with $\gcd(i,t) = 1$, so we can generalize Theorem 7 to that case, if we can find an upper bound on d. As mentioned before, d is in general smaller than or equal to t. But in this case where $r_i \neq 0$ for all i coprime to t, we can do a bit better.

The Jacobstahl function $j(t)$ is defined as the smallest positive integer j such that every sequence of j consecutive integers contains an integer coprime to t. In particular, in the sequence $n - j(t), n - j(t) + 1, \ldots, n - 1$, there is an integer i which is coprime with t. Since i being coprime to t implies $r_i \neq 0$, we conclude $d \leq j(t)$. By a result of Iwaniec ([13]) there exists an absolute constant c such that $j(t) < c \log(t)^2$, which gives the following bound.

Theorem 8. *There exists an absolute constant c such that, if $r_i \neq 0$ for all i with $\gcd(i,t) = 1$, then* $\liminf_{a \to \infty} \frac{b(a) - a}{\log a} \leq c \log(t)^4 \varphi(t)$.

Using similar arguments it seems likely that one could reduce the bound on this lower limit, by letting Q be the product of a larger set of primes. For example, when $\gcd(d,t) = 1$, then one can use the original interval $I = (\frac{1}{2}(l-1)p^{\lambda k_1 + k}, (l-1)p^{\lambda k_1 + k})$ again, and choose $q \in I$ with $q \equiv d \pmod{t}$. This would already reduce the upper bound in Theorem 8 to $2\varphi(t)$. Furthermore, it is conceivable that we can employ techniques introduced in Sections 2.7, 2.8 and 2.9 to ensure $\gcd(l, X_{a,b-1}) < p$ holds, so that we can prove the finiteness of $\liminf_{a \to \infty} \left(\frac{b(a)-a}{\log a} \right)$ in full generality.

3.3 Improvements in the classical case

When $r_1 = t = 1$, we can strengthen Theorems 6 and 7 a bit.

Theorem 9. *If $r_i = 1$ for all i, then* $0.54 < \liminf_{a \to \infty} \frac{b(a) - a}{\log a} < 0.61$.

Proof. In order to show these tighter bounds on the lower limit, divisibility properties of the polynomials $f_d(x) = \sum_{i=0}^{d} \prod_{\substack{j=0 \\ j\neq i}}^{d} (x-j)$ turn out to be important, so we define $\delta(f_d)$ to be the density of primes p such that $f_d(x) \equiv 0 \pmod p$ is solvable. By a (slight extension of a) theorem of Frobenius, which we will meet shortly (see Lemma 34) and which is a consequence of the more well-known Chebotarev's density theorem, this density exists and one can in principle calculate it. With $c = \sum_{d=1}^{\infty} \frac{\delta(f_d)}{d(d+1)}$ we will first prove $\frac{1}{1+c} \leq \liminf_{a\to\infty} \left(\frac{b(a)-a}{\log a}\right) \leq \frac{1}{2c}$, and later on show $0.82 < c < 0.85$, from which $0.54 < \frac{1}{1+c}$ and $\frac{1}{2c} < 0.61$ follow. Let us start with the upper bound and prove that the lower limit is at most $\frac{1}{2c}$.

To prove this upper bound, let $\epsilon > 0$ be small and assume n is a large integer. For $1 \leq d \leq \sqrt{n} - 1$, define S_d to be the set of primes p with $\frac{n}{d+1} < p \leq \frac{n}{d}$ such that $f_d(x) \equiv 0 \pmod p$ is solvable and define $\delta_n(f_d) = |S_d|\pi(n)^{-1}$. If we fix d and let n go to infinity, then we claim that $\delta_n(f_d) \to \frac{\delta(f_d)}{d(d+1)}$. This is seen by combining that, on the one hand, $|S_d| \left(\pi(\frac{n}{d}) - \pi(\frac{n}{d+1})\right)^{-1}$ converges to $\delta(f_d)$ by Frobenius' theorem, while on the other hand $\left(\pi(\frac{n}{d}) - \pi(\frac{n}{d+1})\right)\pi(n)^{-1}$ converges to $\frac{1}{d(d+1)}$ by PNT. Now choose n to be a large enough integer such that $\sum_{d=1}^{\lfloor\sqrt{n}\rfloor-1} \delta_n(f_d) > c - \frac{\epsilon}{2}$.

Lemma 28. *For a prime $p \in S_d$, let x_p be such that $f_d(x_p) \equiv 0 \pmod p$. Then for all i with $0 \leq i \leq d$ we have $x_p \not\equiv i \pmod p$.*

Proof. By contradiction; assume $x_p \equiv i \pmod p$ for some i with $0 \leq i \leq d$. Then $0 \equiv f_d(x_p) \equiv \prod_{\substack{j=0 \\ j\neq i}}^{d} (x_p - j) \pmod p$ and by Euclid's lemma $x_p - j \equiv 0 \pmod p$ for some $j \neq i$, which gives $i \equiv j \pmod p$, which is impossible as $0 < |i-j| \leq d < \frac{n}{d+1} < p$. □

Let q be the largest prime in S_2, so that we have $f_2(x_q) = 3x_q^2 - 6x_q + 2 \equiv 0 \pmod q$. Then $x'_q = -x_q + 2$ is a root of $f_2(x) \pmod q$ as well, since $f_2(x'_q) = 3(-x_q+2)^2 - 6(-x_q+2) + 2 = 3x_q^2 - 6x_q + 2 \equiv 0 \pmod q$. Moreover $x'_q = -x_q + 2 \not\equiv x_q \pmod q$ as otherwise $x_q \equiv 1 \pmod q$, which contradicts Lemma 28. So x_q and x'_q are two distinct roots of $f_2(x) \pmod q$.

We should point out that later on (as part of the proof of Lemma 35) we will prove that $f_d(x-l)$ is an even function whenever $d = 2l$ is even. And from this it follows that whenever p divides $f_d(x_p)$, p will divide $f_d(-x_p+d)$ as well, while, again by Lemma 28, $x_p \not\equiv -x_p + d \pmod{p}$. Apart from $p = q$, in this paper we will not make use of the fact that for even d and all $p \in S_d$, $f_d(x) \equiv 0 \pmod{p}$ has at least two solutions, but it might prove useful if one wants to improve upon our bounds even further.

Anyway, let S be the union of all S_d over all d with $1 \leq d \leq \sqrt{n} - 1$, define $Q = \prod_{p \in S} p$ and define $Q_p = \frac{Q}{p}$ for a prime divisor p of Q. Furthermore, let x_0 and x_1 be the unique positive integers smaller than Q such that the following congruences hold: $x_0 \equiv x_1 \equiv x_p Q_p^{-1} \pmod{p}$ for all $p \in S \setminus \{q\}$, $x_0 \equiv x_q Q_q^{-1} \pmod{q}$ and $x_1 \equiv x'_q Q_q^{-1} \pmod{q}$. Then x_0 and x_1 differ by a multiple of Q_q as they are congruent modulo every prime divisor of Q_q, so at least one of them is bigger than Q_q. Define $x = \max(x_0, x_1) > Q_q$ and redefine $x_q := x'_q$ if $x_1 > x_0$, so that $x \equiv x_p Q_p^{-1} \pmod{p}$ holds for all $p \in S$. Finally, define $b = xQ$ and $a = b - n$ and note that since Q is divisible by a fraction of $\sum_{d=1}^{\lfloor \sqrt{n} \rfloor - 1} \delta_n(f_d) > c - \frac{\epsilon}{2}$ of all primes below n, we obtain $Q > e^{(c - \frac{\epsilon}{2} + o(1))n}$ by applying PNT. Therefore $b = xQ > \frac{Q^2}{q} > e^{(2c - \epsilon + o(1))n}$, from which the desired upper bound on the lower limit then follows if we prove $v_{a,b} < v_{a,b-1}$. To prove this, we need two little lemmas.

Lemma 29. *For all $p \in S$, $L_{a,b}$ is not divisible by p^2.*

Proof. The denominators between a and b that are divisible by $p \in S_d$ are b, $b-p, \ldots, b-dp$, as $b - dp \geq b - n > b - (d+1)p$. Since $\frac{b-ip}{p} = xQ_p - i \equiv x_p - i \not\equiv 0 \pmod{p}$ for $0 \leq i \leq d$ by Lemma 28, we see that $b - ip$ is not divisible by p^2 for any $0 \leq i \leq d$, so $L_{a,b}$ is not divisible by p^2 either. □

Lemma 30. *For all $p \in S$, $X_{a,b}$ is divisible by p, while p does not divide $X_{a,b-1}$.*

Proof. This should be reminiscent of Lemma 3. Let $p \in S_d$ be a prime divisor of Q. For $X_{a,b} \pmod{p}$ we calculate:

$$
\begin{aligned}
X_{a,b} &= L_{a,b} \sum_{i=a}^{b} \frac{1}{i} \\
&\equiv L_{a,b} \sum_{i=0}^{d} \frac{1}{b - ip} && \pmod{p} \\
&\equiv \frac{L_{a,b}}{p} \sum_{i=0}^{d} \frac{1}{xQ_p - i} && \pmod{p} \\
&\equiv \frac{L_{a,b}}{p} \sum_{i=0}^{d} \frac{1}{x_p - i} && \pmod{p} \\
&\equiv \frac{L_{a,b}}{p} \frac{f_d(x_p)}{\prod_{i=0}^{d}(x_p - i)} && \pmod{p} \\
&\equiv 0 \pmod{p}
\end{aligned}
$$

As for $X_{a,b-1}$, assume by contradiction that p does divide $X_{a,b-1}$. Then $X_{a,b} = \frac{L_{a,b}}{L_{a,b-1}} X_{a,b-1} + \frac{L_{a,b}}{b} \equiv \frac{L_{a,b}}{b} \pmod{p}$. This latter quantity is non-zero by Lemma 29, contradicting the fact that we just proved that p divides $X_{a,b}$. □

And now we are almost done. For all primes $p \in S$, by Lemma 30 we have $e_p(g_{a,b}) = 1$, $e_p(g_{a,b-1}) = 0$, while $e_p(L_{a,b}) = e_p(L_{a,b-1}) = 1$, by the proof of Lemma 29. On the other hand, for all primes $p \notin S$, let k_p be such that $p^{k_p} \leq n < p^{k_p+1}$. Then for these primes we have $e_p(g_{a,b-1}) \leq e_p(g_{a,b}) + \min(e_p(x), k_p)$[3] and $e_p(L_{a,b}) = e_p(L_{a,b-1}) + \max(0, e_p(x) - k_p)$. Adding these two (in)equalities gives $e_p(L_{a,b}) + e_p(g_{a,b-1}) \leq e_p(L_{a,b-1}) + e_p(g_{a,b}) + e_p(x)$ for all primes $p \notin S$. Combining both the estimates on the primes that do and do not belong to S, and we get:

$$
\begin{aligned}
L_{a,b} g_{a,b-1} &= \prod_{p \text{ prime}} p^{e_p(L_{a,b}) + e_p(g_{a,b-1})} \\
&= \prod_{p \in S} p^{e_p(L_{a,b}) + e_p(g_{a,b-1})} \prod_{p \notin S} p^{e_p(L_{a,b}) + e_p(g_{a,b-1})} \\
&\leq \prod_{p \in S} p^{e_p(L_{a,b-1}) + e_p(g_{a,b}) - 1} \prod_{p \notin S} p^{e_p(L_{a,b-1}) + e_p(g_{a,b}) + e_p(x)} \\
&= \prod_{p \text{ prime}} p^{e_p(L_{a,b-1}) + e_p(g_{a,b})} \prod_{p \in S} p^{-1} \prod_{p \notin S} p^{e_p(x)} \\
&= L_{a,b-1} g_{a,b} \frac{x}{Q} \\
&< L_{a,b-1} g_{a,b}
\end{aligned}
$$

[3] jamaarhoezodan

We therefore have $v_{a,b} = \frac{L_{a,b}}{g_{a,b}} < \frac{L_{a,b-1}}{g_{a,b-1}} = v_{a,b-1}$, and this finishes the proof of the upper bound $\liminf_{a\to\infty} \left(\frac{b(a)-a}{\log a}\right) \leq \frac{1}{2c}$.

For the lower bound, let n be large and assume that $b-a = n$ with $b > e^{(1+c+\epsilon)n}$. We now aim to prove that $v_{a,b} > v_{a,b-1}$, from which $\frac{1}{1+c} \leq \liminf_{a\to\infty} \left(\frac{b(a)-a}{\log a}\right)$ follows. Define for an integer d with $1 \leq d \leq \sqrt{n}-1$ the set T_d of primes p for which $\frac{n}{d+1} < p \leq \frac{n}{d}$ and such that $p \notin S_d$. Let T be the union of all sets T_d, assume p^{k_p} exactly divides $L_{a,b-1}$ and define $P = \prod_{p\in T} p^{k_p}$ and $P' = \prod_{p\in T} p$. To determine the size of P we use PNT once more, and conclude $P \geq P' = e^{(1-c+o(1))n}$. To prove $v_{a,b} > v_{a,b-1}$ we use the identity $v_{a,b} = \frac{L_{a,b}}{g_{a,b}}$ again and we see that what we want to show is equivalent to $\frac{L_{a,b}}{L_{a,b-1}} > \frac{g_{a,b}}{g_{a,b-1}}$.

Write $b = xyz$ such that $y = \gcd(b, P)$ and $z = \gcd(b, \frac{L_{a,b-1}}{P})$. Since $\gcd(P, \frac{L_{a,b-1}}{P}) = 1$, it follows that $yz = \gcd(b, L_{a,b-1})$ and $x = \frac{L_{a,b}}{L_{a,b-1}}$. By reasoning similar to the proof of Lemma 26, we see $yz = \gcd(b, L_{a,b-1}) \leq e^{(1+o(1))n}$ and therefore $x = \frac{L_{a,b}}{L_{a,b-1}} > e^{(c+\epsilon+o(1))n}$, so it suffices to prove $\frac{g_{a,b}}{g_{a,b-1}} < e^{(c+\epsilon+o(1))}$.

If p^k is any prime power larger than $b-a$, then either p^k does not divide $L_{a,b}$, or there is exactly one integer i in the interval $[a, b]$ for which p^k divides i. In that case $X_{a,b} \not\equiv 0 \pmod{p}$ as there is only one non-zero term in its sum. In both cases we conclude that p^k does not divide $g_{a,b}$. So (the numerator of) $\frac{g_{a,b}}{g_{a,b-1}}$ is only divisible by prime powers smaller than or equal to n. But we claim that it is not divisible by any prime dividing P, so that $\frac{g_{a,b}}{g_{a,b-1}} \leq \frac{L_n}{P'} = e^{(c+o(1))n}$, finishing our proof of the lower bound.

Lemma 31. *For all $p \in T$, the largest power of p that divides $g_{a,b-1}$ is at least as large as the largest power of p that divides $g_{a,b}$.*

Proof. Let $p \in T_d$. To prove the lemma, we consider three different cases. The first case is where p^k divides $L_{a,b}$ for some $k \geq 2$. Then, just as we saw before, p^k does not divide any other integer the interval $[a, b]$, since $b-a = n < p^2 \leq p^k$. So we see that $X_{a,b} \not\equiv 0 \pmod{p}$ as we only have one non-zero term modulo p. Therefore we are free to assume that the largest power of p that divides $L_{a,b}$ (and by extension $g_{a,b}$) is p^1. Now, the second case is where p does not divide b. In that case $X_{a,b} = \frac{L_{a,b}}{L_{a,b-1}} X_{a,b-1} + \frac{L_{a,b}}{b} \equiv \frac{L_{a,b}}{L_{a,b-1}} X_{a,b-1} \pmod{p}$ which is equal to zero if and only if $X_{a,b-1} \equiv 0 \pmod{p}$ as well. For the third and final case assume that p does divide b. Then we can follow the analogous calculation of $X_{a,b} \pmod{p}$ in Lemma 30, to see that $X_{a,b} \not\equiv 0 \pmod{p}$, as otherwise $f_d(x) \equiv 0 \pmod{p}$ is solvable, contrary to $p \notin S_d$.

□

To finish the proof of Theorem 9, we have to bound c and the following lemma takes care of that.

Lemma 32. *With c equal to $\sum_{d=1}^{\infty} \frac{\delta(f_d)}{d(d+1)}$ we have $0.82 < c < 0.85$, from which $0.54 < \frac{1}{1+c}$ and $\frac{1}{2c} < 0.61$ follow by computation.*

Proof. In order to calculate c, we need to know the value of $\delta(f_d)$, so we should try to find some properties of the polynomials f_d. Not all properties listed next will be fully exploited in this paper, but they might very well be useful if one wants to pin down the value of c more precisely.

Lemma 33.

1. *For all $d \in \mathbb{N}$ and all $x \in \mathbb{R}$ we have $f_d(x) = (-1)^d f_d(d-x)$. In other words, $f_d(x+\frac{d}{2})$ is an odd function when d is odd and it is an even function when d is even.* [4]
2. *When d is odd, $\delta(f_d) = 1$.*
3. *All roots of $f_d(x)$ are real and positive. More precisely, if $x_1, x_2, \ldots, x_d$ are the roots of $f_d(x)$ in ascending order, then $i-1 < x_i < i$ for all i with $1 \le i \le d$.*
4. *When d is even, $f_d(x)$ is irreducible if $d \le 500$ or when $d+1$ is prime.*

Proof of 1. By direct calculation:

$$\begin{aligned} f_d(d-x) &= \sum_{i=0}^{d} \prod_{\substack{j=0 \\ j\ne i}}^{d} \big((d-x)-j\big) \\ &= \sum_{i=0}^{d} (-1)^d \prod_{\substack{j=0 \\ j\ne i}}^{d} \big(x-(d-j)\big) \\ &= (-1)^d \sum_{i=0}^{d} \prod_{\substack{j=0 \\ j\ne i}}^{d} (x-j) \\ &= (-1)^d f_d(x) \end{aligned}$$

Plugging in $\frac{d}{2}+x$ gives $f_d(\frac{d}{2}+x) = (-1)^d f_d(\frac{d}{2}-x)$. □

Proof of 2. By the first property we see $f_d(\frac{d}{2}) = 0$ when d is odd, so for odd primes p we have $f_d(x) \equiv 0 \pmod{p}$ for $x \equiv 2^{-1}d \pmod{p}$. □

[4] This was suggested by Will Jagy, see [16].

Proof of 3. When $0 \leq i \leq d$ we claim that $f_d(i) = (-1)^{d-i} i!(d-i)!$, which, one can note, is a slight generalization of Lemma 28. Since this is alternatingly positive and negative, it follows from the intermediate value theorem that for $1 \leq i \leq d$ there must be an $x \in (i-1, i)$ for which $f_d(x) = 0$. To prove our claim, we simply calculate.

$$\begin{aligned} f_d(i) &= \prod_{\substack{j=0 \\ j \neq i}}^{d} (i-j) \\ &= \prod_{j=0}^{i-1} (i-j) \prod_{j=i+1}^{d} (i-j) \\ &= i!(d-i)!(-1)^{d-i} \end{aligned}$$

□

Proof of 4. For $d \leq 500$ we have used the computer program PARI/GP which has the function *polisirreducible* to check whether a polynomial is irreducible or not. Whenever $d+1 = p$ is prime, assume $f_d(x) = a_d x^d + a_{d-1} x^{d-1} + \ldots + a_1 x + a_0$. One can easily check that $a_d = d+1 = p$ and $a_0 = (-1)^d d! = d!$, so that p exactly divides a_d, while it does not divide a_0. We furthermore claim that for all i with $1 \leq i \leq d$, a_i is divisible by p, from which the irreducibility of $f_d(x)$ follows from Eisenstein's criterion after applying the substitution $u = x^{-1}$, which reverses the coefficients of $f_d(x)$.

To prove that p divides all coefficients of $f_d(x)$ except for a_0, we show that $f_d(k) - d! \equiv 0 \pmod{p}$ for all $k \in \mathbb{Z}$, which implies that $f_d(x) - d!$ must be the zero polynomial when reduced modulo p. So let k be an integer and without loss of generality we may assume $0 \leq k \leq p-1$. Then we see $\prod_{\substack{j=0 \\ j \neq i}}^{d} (k-j) \equiv 0 \pmod{p}$, unless $i = k$. So $f_d(k) - d! \equiv \prod_{\substack{j=0 \\ j \neq k}}^{d} (k-j) - d! \equiv d! - d! \equiv 0 \pmod{p}$. □

Note that the conclusion $f_d(k) - d! \equiv 0 \pmod{d+1}$ holds regardless whether $d+1$ is prime or not, but only when $d+1$ is prime does this imply that $f_d(k) - d!$ must be the zero polynomial, when reduced modulo $d+1$. As for non-prime moduli it is possible for a non-zero polynomial to have more roots than its degree; $x^2 - 1 \pmod{8}$ is a typical example.

Since $\delta(f_d) = 1$ when d is odd, for the rest of this section we may assume that $d = 2l$ is even. Denote by G_d the Galois group of $f_d(x)$ and let us view it as a permutation subgroup of the symmetric group on d elements, as elements of G_d permute the roots of $f_d(x)$. From now on, S_d will denote the symmetric group on d elements, as we no longer need our previous definition of S_d as a certain set of primes. Now we have the following special case of a density theorem by Frobenius.

Lemma 34. *If $f_d(x)$ is irreducible, the density $\delta(f_d)$ is equal to the proportion of $\sigma \in G_d$ such that σ is not a derangement. That is, for which a root x of f_d exists such that $\sigma(x) = x$.*

Proof. See [14] for a nice survey with references. They generally work with monic polynomials there, but this assumption can be omitted. $\square$

For example, since it is well-known that S_d contains $d! \sum_{k=0}^{d} \frac{(-1)^k}{k!}$ derangements, if $G_d = S_d$, then $\delta(f_d)$ would be equal to $1 - \sum_{k=0}^{d} \frac{(-1)^k}{k!}$, which for large d approximates $1 - \frac{1}{e} \approx 0.63$. However, we will see that, for $d > 2$, G_d is not isomorphic to S_d, but is instead isomorphic to a subgroup of the so-called signed symmetric group, S_l^+ with $l = \frac{d}{2}$.

The signed symmetric or hyperoctahedral group S_l^+ is the group of permutations σ of $\{-l, -l+1, \ldots, -1, 1, 2, \ldots, l\}$ such that $\sigma(i) = -\sigma(-i)$, for all i. One can also define it as the semidirect product $\mathbb{Z}_2^l \rtimes S_l$, where S_l acts on $\mathbb{Z}_2^l$ in the natural way by permuting coordinates. If one is familiar with wreath products, S_l^+ can furthermore be viewed as the wreath product $\mathbb{Z}_2 \wr S_l$. It is generated by the set of three permutations $\{(-l, -l+1, \ldots, -1)(1, 2, \ldots, l), (-l, -l+1)(1, 2), (-l, 1)\}$.

Lemma 35. *When $d = 2l$ is even, G_d is isomorphic to a subgroup of S_l^+.*

Proof. Define $g_d(x) = f_d(x + \frac{d}{2})$. By Lemma 33, $g_d(x)$ is even and this makes it slightly easier to work with. As $g_d(x)$ and $f_d(x)$ are translates of each other, they have the same Galois group, so it suffices to find the Galois group of $g_d(x)$. Let $\{x_{-l}, x_{-l-1}, \ldots, x_{-1}, x_1, \ldots, x_l\}$ be the roots of $g_d(x)$ with $x_i = -x_{-i}$ and let σ be an element of G_d. If $\sigma(x_i) = x_j$, then $\sigma(-x_i) = -x_j$, since σ is a field automorphism. We can thusly define an injective homomorphism ϕ from G_d to S_l^+ such that for all i, if $\sigma \in G_d$ sends x_i to x_j, then $\phi(\sigma)$ sends i to j.

$\square$

Now that we have narrowed the Galois group of $f_d(x)$ down a little bit, we can try to find to find how many derangements it has. And whenever $G_d \cong S_l^+$ we have an exact formula.

Lemma 36. *The number of derangements in S_l^+ equals $\sum_{m=0}^{l} \frac{2^{l-m} l!}{m!} \sum_{k=0}^{l-m} \frac{(-1)^k}{k!}$.*

Proof. Let $\phi : S_l^+ \to S_l$ be such that $\phi(\sigma) = \tau$ with for all i with $1 \leq i \leq l$, $\tau(i) = |\sigma(i)|$. So τ 'forgets' about all signs of σ. In other words, the pre-image of $\tau \in S_l$ consists of exactly the 2^l elements $\sigma_1, \sigma_2, \ldots, \sigma_{2^l} \in S_l^+$ such that for all j with $1 \leq j \leq 2^l$ and all $i \neq 0$ with $-l \leq i \leq l$, $\sigma_j(i) \in \{-\tau(i), \tau(i)\}$. Let now $\tau \in S_l$ be a permutation that fixes exactly m integers. That is, assume that (possibly after re-indexing) $\tau(i) = i$ for $1 \leq i \leq m$ but $\tau(i) \neq i$ for $m+1 \leq i \leq l$. Then $\sigma \in \phi^{-1}(\tau) \subset S_l^+$ is a derangement if and only if for all i with $1 \leq i \leq m$, $\sigma(i) = -i$. Therefore, for every permutation in S_l that fixes exactly m integers, we have 2^{l-m} derangements in S_l^+.

To find the number of permutations in S_l that fix exactly m integers, we can first choose m out of l integers to fix and then choose a derangement of the other $l-m$ integers. Using the formula for the number of derangements we saw earlier, the number of permutations in S_l that fix exactly m integers is equal to $D_{l,m} = \dfrac{l!}{m!(l-m)!}(l-m)!\displaystyle\sum_{k=0}^{l-m}\frac{(-1)^k}{k!}$. Since we get 2^{l-m} derangements in S_l^+ for every such permutation in S_l, we obtain the result we wanted to prove by multiplying $D_{l,m}$ by 2^{l-m} and summing over all m.

□

Since S_l^+ contains $2^l l!$ integers, the fraction of elements of S_l^+ that is not a derangement equals $1 - \displaystyle\sum_{m=0}^{l}\frac{1}{2^m m!}\sum_{k=0}^{l-m}\frac{(-1)^k}{k!}$ and it might be useful to note that this converges to $1 - \frac{1}{\sqrt{e}} \approx 0.393$ when l goes to infinity. When $G_d \cong S_l^+$, this fraction equals $\delta(f_d)$, by Lemma 34, and using the function *GaloisGroup* from the computer program Magma we have found that $G_d \cong S_l^+$ holds for all even $d \leq 60$, except for $d = 8, 24, 48$. This then finally allows us to find lower and upper bounds on c.

$$
\begin{aligned}
c &= \sum_{d=1}^{\infty}\frac{\delta(f_d)}{d(d+1)} \\
&= \sum_{l=1}^{\infty}\frac{\delta(f_{2l-1})}{2l(2l-1)} + \sum_{\substack{l=1\\ l\neq 4,12,24}}^{30}\frac{\delta(f_{2l})}{2l(2l+1)} + \sum_{l\in\{4,12,24\}}\frac{\delta(f_{2l})}{2l(2l+1)} + \sum_{l=31}^{\infty}\frac{\delta(f_{2l})}{2l(2l+1)} \\
&= \sum_{l=1}^{\infty}\frac{1}{2l(2l-1)} + \sum_{\substack{l=1\\ l\neq 4,12,24}}^{30}\frac{1 - \displaystyle\sum_{m=0}^{l}\frac{1}{2^m m!}\sum_{k=0}^{l-m}\frac{(-1)^k}{k!}}{2l(2l+1)} + \sum_{l\in\{4,12,24\}}\frac{\delta(f_{2l})}{2l(2l+1)} + \sum_{l=31}^{\infty}\frac{\delta(f_{2l})}{2l(2l+1)}
\end{aligned}
$$

The first sum equals $\log(2) \approx 0.6931$ and the second sum is approximately equal to 0.1281, giving $c > 0.82$. For an upper bound we use $\delta(f_{2l}) \leq 1$ which gives 0.016 as upper bound for the third sum and 0.009 as upper bound for the fourth sum, so that, adding it all up, $c < 0.85$.

□

4 Two possible generalizations

4.1 Arbitrary sequences of numerators

We now ask ourselves: can we still prove results similar to those in Section 2 if we no longer assume that the r_i are periodic? First of all, it is clear that we still need at least some condition on the r_i, because, for example, $r_i = (-1)^i i$ would already make for a very uninteresting sequence. But even assuming that the r_i are bounded is not enough to obtain results similar to e.g. Theorem 5.

Define, for example, the sequence with $r_1 = 1$, $r_i = -1$ if $i \geq 2$ is a power of two and $r_i = 0$ otherwise. Then $\displaystyle\sum_{i=1}^{n} \frac{r_i}{i} = \frac{1}{2^k}$, where 2^k is the smallest power of two smaller than or equal to n. So then $v_{1,b}$ is a monotonically increasing function of b, providing a counterexample to a possible strengthening of Theorem 5.

However, this example seems a bit like cheating as well, because r_i equals 0 for almost all i. What if we instead insisted that the r_i were bounded and non-zero? Can we then conclude that for all a there exists a b for which $v_{a,b} < v_{a,b-1}$? As it turns out, the answer is still no. Maybe somewhat surprisingly, given almost any set A, if all we assume is that $r_i \in A$ for all i, then we cannot even exclude the possibility that $v_{1,n} = L_n$ holds for all $n \in \mathbb{N}$, unless A is of a special form. More precisely:

Theorem 10. *If A is a set of non-zero integers containing at least one odd integer, and, for every odd prime p, there exist $a_1, a_2 \in A$ such that $a_1 \not\equiv a_2 \pmod{p}$, then it is possible to assign the r_i values in A, such that the denominator of $\displaystyle\sum_{i=1}^{n} \frac{r_i}{i}$ equals L_n for all $n \in \mathbb{N}$. Conversely, if for some set of non-zero integers A we have $r_i \in A$ for all i, and we know that v_n, defined as the denominator of $\displaystyle\sum_{i=1}^{n} \frac{r_i}{i}$, is only finitely often smaller than L_n, then A must contain at least one odd integer, and, for every odd prime p, there must exist $a_1, a_2 \in A$, such that $a_1 \not\equiv a_2 \pmod{p}$.*

Proof. First we will show that the converse statement is true. Now, if A only contains even integers, it's obvious that $v_{1,n} \leq \frac{L_n}{2}$ for all $n \geq 2$. On the other hand, if there exists an odd prime p such that for all $a_1, a_2 \in A$ we have $a_1 \equiv a_2 \pmod{p}$, then let k be any positive integer and set $n = (p-1)p^k$. We get:

$$
\begin{aligned}
L_n \sum_{i=1}^{n} \frac{r_i}{i} &\equiv L_n r_1 \sum_{i=1}^{p-1} \frac{1}{ip^k} \pmod p \\
&\equiv \frac{L_n r_1}{p^k} \sum_{i=1}^{p-1} \frac{1}{i} \pmod p \\
&\equiv \frac{L_n r_1}{p^k} \sum_{i=1}^{p-1} i \pmod p \\
&\equiv \frac{L_n r_1}{p^k} \frac{p(p-1)}{2} \pmod p \\
&\equiv 0 \pmod p
\end{aligned}
$$

Which implies that $v_{1,n} \leq \frac{L_n}{p}$.

Now we will prove the other direction via induction. For a start, it does not matter what we let r_1 be. Assume now that we have chosen $r_1, r_2, \ldots, r_{n-1} \in A$ so that $\dfrac{X_{n-1}}{L_{n-1}} = \displaystyle\sum_{i=1}^{n-1} \frac{r_i}{i}$ with $\gcd(X_{n-1}, L_{n-1}) = 1$. Then we will show that we can choose $r_n \in A$ so that $\gcd(X_n, L_n) = 1$ holds as well. Note that, in general, $\gcd(X_n, L_n) = 1$ is equivalent to saying that the smallest prime divisor of X_n is bigger than n. In particular, with the induction hypothesis we assume that $X_{n-1} \not\equiv 0 \pmod q$ for all primes $q \leq n-1$.

We have to distinguish between three different cases; either n is a prime power, or n is not a prime power but still divisible by a certain large prime power, or n is not divisible by a large power of a prime.

Case I. n is a prime power.
Assume $n = p^k$, let $q \neq p$ be any other prime smaller than n and choose an arbitrary $r_n \in A$ that is not divisible by p. We claim that both $X_n \not\equiv 0 \pmod p$ and $X_n \not\equiv 0 \pmod q$. First off, note that this case is the only one where $L_n \neq L_{n-1}$ and, more precisely, $L_n = pL_{n-1}$. Now, on the one hand, $X_n = pX_{n-1} + \frac{L_n r_n}{n} \equiv \frac{L_n r_n}{n} \not\equiv 0 \pmod p$. While on the other hand $X_n = pX_{n-1} + \frac{L_n r_n}{n} \equiv pX_{n-1} \not\equiv 0 \pmod q$, by the induction hypothesis.

Case II. $n = lp^k$ for some $1 < l < p$ and $k \geq 1$.
In this case we claim that this prime p is unique. Consider the possibility that n can also be written as $n = \tilde{l} q^{\tilde{k}}$ for some prime $q \neq p$ with $\tilde{l} < q$ and $\tilde{k} \geq 1$. By unique factorization we see $q^{\tilde{k}} | l$ and $p^k | \tilde{l}$, which would imply $l \geq q^{\tilde{k}} > \tilde{l} \geq p^k > l$; contradiction. In other words, if $n = \tilde{l} q^{\tilde{k}}$, then $\tilde{l} > q$, so that, in particular, $q^{\tilde{k}+1}$ must divide L_n. Let now $a_1, a_2 \in A$ be such that $a_1 \not\equiv a_2 \pmod p$. Then, regardless of whether we choose $r_n = a_1$ or $r_n = a_2$, for any $q < n$ different

from p we have $X_n = X_{n-1} + \frac{L_n r_n}{n} \equiv X_{n-1} \pmod{q}$, which we assumed to be non-zero for all $q < n$. On the other hand, $X_{n-1} + \frac{L_n a_1}{n} \not\equiv X_{n-1} + \frac{L_n a_2}{n} \pmod{p}$, so that at least one of those is non-zero modulo p. Set r_n to an a_i for which this holds, and $X_n = X_{n-1} + \frac{L_n r_n}{n} \not\equiv 0 \pmod{p}$.

Case III. $n = lp^k$ implies $l > p$.
As we noted in the previous case, this implies that L_n is divisible by a power of p that exceeds p^k. And so regardless of the value of r_n we get $X_n = X_{n-1} + \frac{L_n r_n}{n} \equiv X_{n-1} \pmod{p}$ which we assumed was non-zero for all $p < n$. And we conclude that, for this case, we may choose r_n arbitrarily.

In all cases it was possible for us to choose $r_n \in A$ in such a way that $X_n \not\equiv 0 \pmod{p}$ holds for all $p \leq n$, and the theorem is proved.

□

Even though we have shown that we cannot hope to strengthen Theorem 5 by allowing the r_i to be non-periodic, we can however generalize Theorem 3 a bit. That is, let $\{r_i\}_{i \in \mathbb{N}}$ now be a bounded sequence of non-zero integers with $r = \max_i |r_i|$ and assume $m > r$ is just any given integer. Let $2 = p_1 < p_2 < \ldots < p_z < m$ be the sequence of primes smaller than m and let $\tilde{m}$ be any integer larger than m^{12z}. We can then find n such that X_n has a large prime divisor.

Theorem 11. *There exists an integer $n \in [\tilde{m}, \tilde{m} + m^{2z+2})$ for which X_n is divisible by a prime larger than or equal to m.*

The proof of this theorem goes through the exact same steps as the proof of Theorem 3, since in that proof we never really used the fact that our original sequence was periodic, apart from ensuring the existence of i for which $r_i \neq 0$. However, we can see that in this case the constants are much nicer, which might seem very surprising at first. But this is all because of the extra assumption that $r_i \neq 0$ for all i.

For example, we assumed that w from Lemmas 12 and 10 was exponentially large in the length of the interval I in order to prove that $\sum_{i=w+1}^{w+k} \frac{r_i}{i} \neq 0$ [5]. However, when the r_i are guaranteed to be non-zero and $w \geq k \geq m$, we can use Sylvester's Theorem which states that there exists an $i \in [w+1, w+k]$ such that i is divisible by a prime larger than k, which is the length of the interval. This will then imply that $\sum_{i=w+1}^{w+k} \frac{r_i}{i}$ is non-zero, and at no point do we need w to be very large at all.

Similarly, if the r_i are non-zero, we do not have to deal with the possible existence of primes in Σ_3. So terms like $t^{4\mu_2}$ that keep popping up in, for example,

[5] We will come back to this point later on when we discuss ways to optimize our arguments.

Section 2.5 will simply not appear in the proof of Theorem 11.

In conclusion we can say that very few results that we obtained thus far could possibly carry over to the general non-periodic case, even with the obvious extra assumptions that the r_i are non-zero and remain bounded. But the one result that does generalize, actually becomes quite nice.

A natural follow-up question is now: how many of the r_i have to be 0 in order for Theorem 11 to become false? Or, moving even further astray, fix m and let $|r_i| \leq m$ for all i. Furthermore assume that $\left(\sum_{i=1}^{k} \frac{r_i}{i}\right)^{-1}$ is an integer for all k with $1 \leq k \leq n$. What is the largest possible subset A of $\{1, 2, \ldots, n\}$ such that for all $i \in A$ we have $r_i \neq 0$?

The example given at the start of this section provides an example with $|A| > c\log(n)$, but it seems likely that better constructions are possible. However, these questions, interesting and tempting as they may be, do lead us away from the original subject of this paper. So for now we gladly pass these questions on to the next brave soul.

4.2 When the denominators are powers of consecutive integers

Let d be a positive integer. It seems natural to look at sums of the form $\sum_{i=a}^{b} \frac{r_i}{i^d}$ and see whether one can prove something similar to Theorem 2 in this more general case. For a start, what we can say is that it is possible to generalize Theorem 1 and almost the exact same proof can be used. We will use analogous definitions ($L_{a,b}$ should now be the least common multiple of all integers $i^d \in \{a^d, (a+1)^d, .., b^d\}$ for which $r_i \neq 0$) and to specify the dependence on d, $b_d(a)$ will denote the smallest b such that $v_{a,b} < v_{a,b-1}$. Just like the start of Section 2.2, let $p > \max(r, t)$ be a prime number such that $p | X_n$ with $n = lp^k \geq c_1$, where n is the smallest such integer. Let k_1 be the smallest integer for which $p^{\lambda k_1 + k} \geq \max(a, 2t)$ and choose $b = np^{\lambda k_1} = lp^{\lambda k_1 + k}$. We then obtain the following generalization of Theorem 1.

Theorem 12 (Theorem 1, generalized version)**.** *If* $\gcd(l^d, X_{a,b-1}) < p$*, then* $v_{a,b} < v_{a,b-1}$ *and* $b_d(a) \leq b \leq \max(a-1, 2t-1)lp^{\lambda}$*.*

The only difference here is l^d in place of l in the assumption $\gcd(l^d, X_{a,b-1}) < p$, which is of course harder to satisfy when d is large. On the other hand, when d is large, then for certain j it is easier to find a large prime divisor p of X_j, which leads to the following corollary.

Corollary 3. *Let i and $j > i$ be the smallest two (positive) indices such that r_i and r_j are non-zero. If $j = q^k$ is a prime power, then, for all but finitely many d, $b_d(a)$ is finite for all a.*

Note that, in particular, if at least two out of r_1, r_2, r_3, r_4, r_5 are non-zero, then Corollary 3 applies.

Proof (sketch). We define the following constants: $g = \gcd(i, j)$, $h = \gcd(r_i j^d, r_j i^d)$, $A = h^{-1} r_i j^d$ and $B = h^{-1} r_j i^d$. The idea of introducing A and B is that $X_j = g^{-d}(r_i j^d + r_j i^d)$, so that $A + B$ divides X_j. In other words, if we can prove that $A + B$ has a large prime divisor, then X_j must have a large prime divisor as well, and we do this by applying known bounds on the abc conjecture. As usual, define $m = 1 + \max(|r_1|, |r_2|, \ldots, |r_t|, t)$ and let $rad(x)$ be the radical of x; the largest squarefree divisor of x. We then have the following inequalities.

Lemma 37.

1. *If $d > 4m^2$, then $|A + B| > e^{\frac{d}{2}}$.*
2. *$rad(AB) < 2m^4$.*
3. *$rad(A + B) > K^{-1} \frac{\log(A+B)}{rad(AB)}$, for some absolute constant $K \geq 1$.*

Proof of 1. We have $g^d \leq h \leq |r_i r_j| g^d$, so $|A| = \frac{|r_i| j^d}{h} > \frac{|r_i| j^d}{|r_i r_j| g^d} > \frac{1}{m} (\frac{j}{g})^d \geq \frac{2^d}{m}$, since $\frac{j}{g} \in \mathbb{N}$ and $j > i \geq g$. On the other hand, $i < j < 2m$, and this gives us the possibility of finding a lower bound on $|A/B|$, using the inequalities $\log(1 + x) > \frac{x}{2}$ (which is valid for $0 < x \leq 1$) and $e^x > 2x$ (which is valid for all $x \in \mathbb{R}$).

$$
\begin{aligned}
|A/B| &> \frac{1}{m} \left(\frac{j}{i} \right)^d \\
&> \frac{1}{m} \left(\frac{2m+1}{2m} \right)^d \\
&= \frac{1}{m} e^{d \log(1 + \frac{1}{2m})} \\
&> \frac{1}{m} e^{\frac{d}{4m}} \\
&> \frac{1}{m} e^m \\
&> 2
\end{aligned}
$$

Combining our lower bounds on $|A|$ and $|A/B|$, gives us the desired lower bound on $|A + B|$;

$$
\begin{aligned}
|A+B| &\geq |A| - |B| \\
&> |A| - \frac{1}{2}|A| \\
&> \frac{2^d}{2m} \\
&> \frac{e^{\frac{2d}{3}}}{e^m} \\
&= e^{\frac{2d}{3}-m} \\
&> e^{\frac{2d}{3}-\frac{d}{6}} \\
&= e^{\frac{d}{2}}
\end{aligned}
$$

□

Proof of 2. By definition,

$$
\begin{aligned}
rad(AB) &\leq rad(r_i j^d r_j i^d) \\
&\leq rad(r_i) rad(j^d) rad(r_j) rad(i^d) \\
&= rad(r_i) rad(j) rad(r_j) rad(i) \\
&< 2m^4
\end{aligned}
$$

where the final inequality follows from $\max(|r_i|, |r_j|, i) < m$ and $j < 2m$. □

Proof of 3. Since A and B are coprime by construction, we may apply Theorem 1 from [15] and take K large enough such that, with the notation of [15], $KG^{\frac{2}{3}} > c\log^3(G)$. In particular, since $\max_{x \geq 1} \frac{\log^3(x)}{x^{\frac{2}{3}}} < 5$, $K = \max(1, 5c)$ suffices. Note that G in the context of [15] corresponds to $rad(AB(A+B)) = rad(AB)rad(A+B)$ in our notation. The statement $A + B < e^{KG}$ is therefore, in our notation, equivalent to $K^{-1}\log(A + B) < rad(AB)rad(A + B)$, which is the desired inequality.

□

To sketch how to finish the proof of Corollary 3, assume $d > 4Km^4e^{2m^2}$. If we then apply all inequalities from Lemma 37, we obtain $rad(A + B) > \frac{d}{4Km^4} > e^{2m^2}$. Since the product of all primes smaller than m^2 is smaller than e^{2m^2} by Lemma 8, we conclude that X_j, which is divisible by $A + B$, must be divisible by a prime larger than m^2. To satisfy the condition $\gcd(j^d, X_{a,b-1}) = \gcd(q^{kd}, X_{a,b-1}) < p$ we now apply a suitable generalization of Lemma 19 to obtain intervals I such that the largest power of q that divides X_n is smaller than $m^2 < p$, for all $n \in I$. The arguments from Section 2.8 now provide infinitely b for which $v_{a,b} < v_{a,b-1}$.

□

It should be noted that the proof of Lemma 19 and the ideas present in Section 2.8 are all mostly independent of the value of d, so they quite easily generalize to work for general d. Moreover, Baker's method again allows us to make everything explicit (see Section 2.9). Lastly, it might not be too difficult to prove the finiteness of $b_d(a)$ for all d and a, but we will not pursue this.
We do however want to point out it would follow from Schanuel's Conjecture that $b_d(a)$ is always finite. This can be shown along the same lines as the proposed proof of Theorem 4 of [2], combining the intervals that we constructed in Lemmas 18 and 19 for certain primes $p_1, \ldots, p_k$. As we mentioned in Section 1 this depends on the linear independence of $\theta_i = \frac{\log(p_1)}{\log(p_i)}$, and this can be proven under the assumption of Schanuel's Conjecture. But from now on, we will restrict to the case $r_i = 1$ and for that, let us introduce some notation.

Let p_d be the smallest prime p for which $p-1$ does not divide d, define q_i to be the smallest prime divisor of X_i and let c_d be the smallest constant such that $b_d(a) \leq c_d \max(1, a-1)$ for all $a \in \mathbb{N}$. Recall that Corollary 1 gave us $c_1 = 6$, since $b_1(1) = b_1(2) = 6$, and it possible to generalize this a little and calculate c_d for all d.

Theorem 13. *If d is odd, then $c_d = 6$. For even d we have the (in)equalities $c_d = b_d(1) = \min_i(iq_i) \leq \frac{1}{2}p_d(p_d-1)$, where the minimum runs over all i with $2 \leq i \leq \frac{1}{2}(p_d-1)$.*

Proof. Since $v_{1,b_d(1)} < v_{1,b_d(1)-1}$, we see $\gcd(X_{b_d(1)}, L_{b_d(1)}) > 1$, so let p be a prime divisor of $\gcd(X_{b_d(1)}, L_{b_d(1)})$. Then in particular p divides $b_d(1)$ and with $b_d(1) = lp$, we see that p must divide X_l. And this implies $c_d \geq b_d(1) \geq \min_{i\geq 2}(iq_i)$.
On the other hand, if for some i and all $n \geq a$, $\gcd(i^d, X_{a,n}) < q_i$ then with $i = l$ and $q_i = p$, the upper bound on $b_d(a)$ in Theorem 12 would simplify and can be rewritten as $c_d \leq iq_i$. We claim that, for any a and $n > a$, the smallest prime divisor of $X_{a,n}$ is at least p_d, which in particular implies $q_i \geq p_d$ for all i. It also implies that $\gcd(i^d, X_{a,b-1}) = 1 < q_i$ whenever $i < p_d$ and from this it follows that $c_d = 6$ when d odd, since for odd d it is easily seen that $q_2 = 3$. We furthermore claim that, when d is even, $q_{\frac{1}{2}(p_d-1)} = p_d$, so that $\min_i(iq_i) \leq \frac{1}{2}p_d(p_d-1)$ and the minimum can be taken over $2 \leq i \leq \frac{1}{2}(p_d-1)$, as $q_i \geq p_d$. To finish the proof of Theorem 13 let us state these claims again and prove them.

Lemma 38. *Let p be a prime such that $p-1$ divides d. Then for all positive integers a and $n \geq a$ we have that p does not divide $X_{a,n}$.*

Lemma 39. *Let p be a prime such that $p-1$ does not divide d. Then, first of all, $X_{p-1} \equiv 0 \pmod{p}$. When d is even we furthermore have $X_{\frac{1}{2}(p-1)} \equiv 0 \pmod{p}$ as well.*

Proof of Lemma 38. The idea is that d is a multiple of $\varphi(p) = p-1$ so that $i^d \equiv 1 \pmod{p}$ for $1 \leq i \leq p-1$. Assume that p^{dk} exactly divides $L_{a,n}$ and let j_1 and j_2 be such that $(j_1 - 1)p^k < a \leq j_1 p^k \leq j_2 p^k \leq n < (j_2 + 1)p^k$ with $1 \leq j_1 \leq j_2 \leq p - 1$. Then let us look at $X_n \pmod{p}$.

$$\begin{aligned} X_{a,n} &\equiv \frac{L_{a,n}}{p^{dk}} \sum_{i=j_1}^{j_2} \frac{1}{i^d} && \pmod{p} \\ &\equiv \frac{L_{a,n}}{p^{dk}} (j_2 + 1 - j_1) && \pmod{p} \end{aligned}$$

And this is non-zero since $1 \leq j_2 + 1 - j_1 \leq p - 1$.

□

Proof of Lemma 39. Let g be a primitive root of p and recall that $\{g, 2g, \ldots, (p-1)g\}$ and $\{\frac{1}{1}, \frac{1}{2}, \ldots, \frac{1}{p-1}\}$ are both complete sets of non-zero residues modulo p. In particular we see that $\sum_{i=1}^{p-1} (ig)^d \equiv \sum_{i=1}^{p-1} i^d \equiv \sum_{i=1}^{p-1} \frac{1}{i^d} \pmod{p}$ and we use this to prove that p divides X_{p-1}.

$$\begin{aligned} 0 &\equiv L_{p-1} \sum_{i=1}^{p-1} \left((ig)^d - i^d \right) && \pmod{p} \\ &\equiv (g^d - 1) L_{p-1} \sum_{i=1}^{p-1} i^d && \pmod{p} \\ &\equiv (g^d - 1) X_{p-1} && \pmod{p} \end{aligned}$$

Since $g^d - 1 \not\equiv 0 \pmod{p}$ as g is a primitive root of p and $p - 1$ does not divide d, we see that X_{p-1} must be divisible by p. Moreover, when d is even, we have $i^d = (-i)^d$, so that the first and second half of the sum are equal to one another;

$$\begin{aligned} 0 &\equiv L_{\frac{1}{2}(p-1)} \sum_{i=1}^{p-1} \frac{1}{i^d} && \pmod{p} \\ &\equiv L_{\frac{1}{2}(p-1)} \left(\sum_{i=1}^{\frac{1}{2}(p-1)} \frac{1}{i^d} + \sum_{i=1}^{\frac{1}{2}(p-1)} \frac{1}{(-i)^d} \right) && \pmod{p} \\ &\equiv 2X_{\frac{1}{2}(p-1)} && \pmod{p} \end{aligned}$$

Since $p - 1$ does not divide d, we know that $p \neq 2$ and we therefore conclude $p | X_{\frac{1}{2}(p-1)}$. □

Corollary 4. *For all d, $c_d = O(\log^{10}(d))$. On the other hand, $c_d > 3\log(d)$ infinitely often.*

Proof. Let c be a small enough constant and q be a prime smaller than $cp_d^{\frac{1}{5}}$. Then in [11] it is proven that there exists a prime $p < p_d$ such that $p \equiv 1 \pmod{q}$. Since d is divisible by $p-1$ for all $p < p_d$, q divides d as well. Therefore $d \geq \prod_{q < cp_d^{\frac{1}{5}}} q = e^{(1+o(1))cp_d^{\frac{1}{5}}}$, implying $p_d = O(\log^5(d))$. Since $c_d < p_d^2$, the upper bound follows. For the lower bound, choose $d = \text{lcm}(1, 2, 4, 6, 10, \dots, p_d - 1)$ with $p_d > 10^6$ and note that d is not divisible by any prime larger than $\frac{1}{2}p_d$. We furthermore have $\frac{\log(p_d)}{\log(\frac{1}{2}p_d)} < 1.053$, since $p_d > 10^6$. Then $c_d = \min_i(iq_i) \geq 2p_d > 3\log(d)$, where the final inequality follows from Lemma 8; $d < p_d^{\pi(\frac{1}{2}p_d)} < e^{1.053 \cdot 0.63 p_d} < e^{\frac{2}{3}p_d}$. ☐

Corollary 5. $c_d = \begin{cases} 6 & \textit{if } d \equiv 1 \pmod{2} \\ 10 & \textit{if } d \equiv 2 \pmod{4} \\ 21 & \textit{if } d \equiv 4, 8 \pmod{12} \\ 34 & \textit{if } d \equiv 12 \pmod{24} \\ 55 & \textit{if } d \equiv 24, 48, 72, 96 \pmod{120} \end{cases}$

Proof. All of these can be relatively quickly checked by calculating p_d (which increases in every case), finding the possible values of q_i for the first few i and applying $c_d = \min_i(iq_i) \leq \frac{1}{2}p_d(p_d - 1)$, when d is even. Let us do this for the final (and hardest) case of $d \equiv 24, 48, 72, 96 \pmod{120}$, and leave the rest for the interested reader. So we will assume that 24 divides d but 5 does not divide d. Since 24 is divisible by $1, 2, 4$ and 6, but not by 10, we see $p_d = 11$ and, using Theorem 13, we obtain $c_d \leq 55$ right away. Furthermore, we claim that X_i is not divisible by 13 for any i, not divisible by 17 for $i \leq 3$ and not divisible by 19 or 23 for $i = 2$, so that iq_i is minimized for $i = 5$, $q_i = 11$. To prove that X_i is not divisible by $13, 17, 19$ or 23 for the relevant values of i, let us deal with them one prime at a time.

By Lemma 38 we have that 13 does not divide X_i for any i, as $12|24$. For 17 we have $\frac{1}{i^d} \equiv \pm 1 \pmod{17}$, as $8|d$. But $\frac{1}{2^8} \equiv 1 \pmod{17}$, so that $X_2 \equiv 2 \pmod{17}$ and $X_3 \pmod{17}$ is either 1 or 3, so definitely non-zero. Finally, the only way either 19 or 23 divides X_2 is if $\frac{1}{2^d}$ is congruent to -1 modulo 19 or 23. But for 23 this congruence is not solvable, while for 19 we have that 2 is a primitive root, so $\frac{1}{2^d} \equiv -1 \pmod{19}$ precisely when $d \equiv 9 \pmod{18}$. But this is impossible as d is even. ☐

With the help of a computer it is not hard to extend Corollary 5. For example, $17|X_6$ when $d \equiv 120 \pmod{240}$, $37|X_3$ when $d \equiv \pm 240 \pmod{720}$, $p_d = 23$ when $\gcd(d, 11 \cdot 720) = 720$, $p_d = 29$ when $\gcd(d, 7 \cdot 7920) = 7920$ and 193 divides X_2 when $d = 7 \cdot 7920$. Working this all out gives $c_d \leq 406$ for $d < 110880$.

Theorem 13 shows that c_d is always equal to $b_d(1)$, which may suggest there exists a constant $c'_d < c_d$ such that $b_d(a) \leq c'_d(a-1)$, as long as a is large enough. This is indeed often the case, at least when d is even. Let c'_d be the smallest constant such that $b_d(a) \le c'_d(a-1)$ holds for all $a \ge 4$. [6] Then for all even $d < 120$ we can improve upon Corollary 5;

Lemma 40. $$c'_d \le \begin{cases} \frac{25}{3} = 8.\overline{3} & \textit{if } d \equiv 2 \pmod 4 \\ \frac{147}{8} = 18.375 & \textit{if } d \equiv 4, 8 \pmod{12} \\ \frac{34}{3} = 11.\overline{3} & \textit{if } d \equiv 12 \pmod{24} \\ \frac{55}{3} = 18.\overline{3} & \textit{if } d = 24 \\ \frac{111}{5} = 22.2 & \textit{if } d = 48 \\ \frac{69}{2} = 34.5 & \textit{if } d = 72 \\ \frac{605}{23} \approx 26.3 & \textit{if } d = 96 \end{cases}$$

Proof (sketch). We will not give all the details, but instead construct functions $f_d(a)$ such that the motivated reader can check themselves that $v_{a,f_d(a)} < v_{a,f_d(a)-1}$ and $f_d(a) \le c(a-1)$ hold (for the constants c appearing in the statement of Lemma 40) whenever $f_d(a)$ is defined, using the ideas that were already present in Section 2.3. Moreover, in every case we make sure that if $f_d(a) = lp^k$ (where the meaning of p in the different cases should be clear), then every prime divisor q of l will be such that $q-1$ divides d, so that $\gcd(l^d, X_{a,f_d(a)-1}) = 1 < p$ follows immediately from Lemma 38 and does not have to be checked separately. Finally, there is no doubt that these values can be extended and improved upon even further, but this paper is long enough as it is.

$$\text{If } d \equiv 2 \pmod 4 : f_d(a) = \begin{cases} 10 & \text{if } 3 \le a \le 5 \\ 21 & \text{if } a = 6 \text{ and } d \equiv 2, 10 \pmod{12} \\ 26 & \text{if } a = 6 \text{ and } d \equiv 6 \pmod{12} \\ 9 \cdot 5^{k-1} & \text{if } 5^k < a \le 6 \cdot 5^{k-1} \text{ for some } k \ge 2 \\ 2 \cdot 5^{k+1} & \text{if } 6 \cdot 5^{k-1} < a \le 5^{k+1} \text{ for some } k \ge 1 \end{cases}$$

$$\text{If } d \equiv 4, 8 \pmod{12} : f_d(a) = \begin{cases} 21 & \text{if } 3 \le a \le 7 \\ 34 & \text{if } a = 8 \text{ and } d \equiv 4 \pmod 8 \\ 78 & \text{if } a = 8 \text{ and } d \equiv 0 \pmod 8 \\ 10 \cdot 7^{k-1} & \text{if } 7^k < a \le 8 \cdot 7^{k-1} \text{ for some } k \ge 2 \\ 3 \cdot 7^{k+1} & \text{if } 8 \cdot 7^{k-2} < a \le 7^{k+1} \text{ for some } k \ge 1 \end{cases}$$

$$\text{If } d \equiv 12 \pmod{24} : f_d(a) = \begin{cases} 34 & \text{if } 4 \le a \le 17 \\ 7 \cdot 17^k & \text{if } 17^k < a \le 2 \cdot 17^k \text{ for some } k \ge 1 \\ 14 \cdot 17^k & \text{if } 2 \cdot 17^k < a \le 3 \cdot 17^k \text{ for some } k \ge 1 \\ 2 \cdot 17^{k+1} & \text{if } 3 \cdot 17^k < a \le 17^{k+1} \text{ for some } k \ge 1 \end{cases}$$

[6] We choose $a \ge 4$ just because it happens to work in all cases we will consider. We conjecturally have $b_d(a) < (1+\epsilon)a$ for large enough a.

$$\text{If } d \equiv 24 \pmod{120} : f_d(a) = \begin{cases} 55 & \text{if } 4 \leq a \leq 11 \\ 8 \cdot 11^k & \text{if } 11^k < a \leq 2 \cdot 11^k \text{ for some } k \geq 1 \\ 9 \cdot 11^k & \text{if } 2 \cdot 11^k < a \leq 3 \cdot 11^k \text{ for some } k \geq 1 \\ 5 \cdot 11^{k+1} & \text{if } 3 \cdot 11^k < a \leq 11^{k+1} \text{ for some } k \geq 1 \end{cases}$$

$$\text{If } d = 48 : f_d(a) = \begin{cases} 55 & \text{if } 4 \leq a \leq 5 \\ 16 \cdot 37^k & \text{if } 37^k < a \leq 2 \cdot 37^k \text{ for some } k \geq 1 \\ 17 \cdot 37^k & \text{if } 2 \cdot 37^k < a \leq 3 \cdot 37^k \text{ for some } k \geq 1 \\ 18 \cdot 37^k & \text{if } 3 \cdot 37^k < a \leq 4 \cdot 37^k \text{ for some } k \geq 1 \\ 34 \cdot 37^k & \text{if } 4 \cdot 37^k < a \leq 5 \cdot 37^k \text{ for some } k \geq 1 \\ 3 \cdot 37^{k+1} & \text{if } 5 \cdot 37^k < a \leq 37^{k+1} \text{ for some } k \geq 0 \end{cases}$$

$$\text{If } d = 72 : f_d(a) = \begin{cases} 55 & \text{if } 3 \leq a \leq 11 \\ 111 & \text{if } 12 \leq a \leq 23 \\ 9 \cdot 23^k & \text{if } 23^k < a \leq 2 \cdot 23^k \text{ for some } k \geq 1 \\ 3 \cdot 23^{k+1} & \text{if } 2 \cdot 23^k < a \leq 23^{k+1} \text{ for some } k \geq 1 \end{cases}$$

$$\text{If } d \equiv 96 \pmod{120} : f_d(a) = \begin{cases} 55 & \text{if } 4 \leq a \leq 11 \\ 111 & \text{if } a = 23 \\ 7 \cdot 11^k & \text{if } 11^k < a \leq 2 \cdot 11^k \text{ for some } k \geq 1 \\ 27 \cdot 11^{k-1} & \text{if } 2 \cdot 11^k < a \leq 23 \cdot 11^{k-1} \text{ for some } k \geq 2 \\ 5 \cdot 11^{k+1} & \text{if } 23 \cdot 11^{k-1} < a \leq 11^{k+1} \text{ for some } k \geq 1 \end{cases}$$

□

5 Final thoughts and remarks

It is not hard to show that for every $\epsilon > 0$ we can improve the inequality $v_{a,b} < v_{a,b-1}$ in Theorem 2 to the slightly stronger $v_{a,b} < \epsilon v_{a,b-1}$. To prove this, first recall that we chose $M = \lfloor e^{3.8m} \rfloor$ in Section 2.8 to make sure that $l > M$ was either divisible by a prime $q \geq m$, or a prime $q < m$ such that q^y divides l with $q^y > m^3$. Assume for simplicity that $\epsilon \leq \frac{1}{11}$, so that $1.26(3 + \frac{\log(\epsilon^{-1})}{\log(m)} + \epsilon^{-1}) < 1.9\epsilon^{-1}$ and now choose $M = \left\lfloor e^{1.9\epsilon^{-1}m} \right\rfloor$. Then we claim that $l > M$ is either divisible by a prime $q > \epsilon^{-1}m$, or a prime $q < m$ such that q^y divides l with $q^y > \epsilon^{-1}m^3$. We prove this by showing that the product of all prime powers smaller than or equal to $\epsilon^{-1}m$ times the product of all prime powers $q^y \leq \epsilon^{-1}m^3$ with $q < m$, is smaller than l.

$$\begin{aligned}
l &> e^{1.26m(3+\frac{\log(\epsilon^{-1})}{\log(m)}+\epsilon^{-1})} \\
&= (m^3)^{\frac{1.26m}{\log(m)}} \cdot (\epsilon^{-1})^{\frac{1.26m}{\log(m)}} \cdot (\epsilon^{-1}m)^{\frac{1.26\epsilon^{-1}m}{\log(\epsilon^{-1}m)}} \\
&> (\epsilon^{-1}m^3)^{\pi(m)} \cdot (\epsilon^{-1}m)^{\pi(\epsilon^{-1}m)} \\
&\geq \prod_{q<m} q^{\left\lfloor \frac{\log(\epsilon^{-1}m^3)}{\log(q)} \right\rfloor} \prod_{q \leq \epsilon^{-1}m} q^{\left\lfloor \frac{\log(\epsilon^{-1}m)}{\log(q)} \right\rfloor}
\end{aligned}$$

To find an explicit bound on the smallest b such that $v_{a,b} < \epsilon v_{a,b-1}$, one can go through the calculations from Section 2.9 again, which then results in the constant c from Theorem 5 increasing to $c = e^{e^{e^{e^{2\epsilon^{-1}m}}}}$.

In fact, in the case $r_i = t = 1$, we can use Linnik's Theorem to provide us with a prime p that we can apply in Theorem 1 to effectively get $\liminf_{b \to \infty} \frac{v_{a,b}}{v_{a,b-1}} = 0$. Let $k_0 \in \mathbb{N}$ be arbitrary and let p be the smallest prime which is $1 \pmod{2^{k_0}}$. By the current best known bound on Linnik's Theorem (see [11]), $p < c_1 2^{5k_0}$ for some constant c_1. Moreover, by Wolstenholme's Theorem (or common sense), $n = l = p - 1$ is such that p divides X_n, while $\gcd(l, X_{a,b-1}) < p2^{-k_0}$, as $X_{a,b-1}$ is always odd. Again by the proof of Theorem 1 we obtain $v_{a,b} < 2^{-k_0} v_{a,b-1}$ with $b < c_1^2 2^{10k_0} a$. For the sake of clarity and completeness, let us cleanly state the above results.

Corollary 6. *For all $\epsilon \in (0, \frac{1}{11}]$ there exists a constant $c_\epsilon := e^{e^{e^{e^{2\epsilon^{-1}m}}}}$ such that for all $a \in \mathbb{N}$ there exists a $b < c_\epsilon a$ for which $v_{a,b} < \epsilon v_{a,b-1}$.*

Corollary 7. *If $r_i = 1$ for all i, then there is an absolute constant K such that for all $\epsilon \in (0, 1]$ and all $a \in \mathbb{N}$ there exists a $b < K\epsilon^{-5}a$ for which $v_{a,b} < \epsilon v_{a,b-1}$.*

Speaking of effective results, even though we found an upper bound for the constants c in Theorems 4 and 5, these bounds are of course outrageous. In many steps along the way we were perhaps a bit wasteful and there are surely many improvements possible. For example, the reason we needed w from Lemmas 12 and 10 to be exponentially large in the length of the interval I, was so that we could prove that the sum $\sum_{i=w+1}^{w+k} \frac{r_i}{i}$ is non-zero. We proved this by using an integer in the interval $[w, w + c\log(w)]$ which has a prime power divisor larger than the length of the interval, completely analogous to the standard proof that $\sum_{i=1}^{n} \frac{1}{i}$ cannot be an integer for $n > 1$.

In [10] however, it is proven that in every large enough interval $[w, w+\sqrt{w}]$ there will be some integer which is divisible by a prime larger than $w^{1/2+1/15}$, and note that this is larger than the length of the interval. So the conclusion here is that we only need w to be roughly the square of the length of I, instead of the exponential of it. Of course there are some details to be worked out here (most importantly the fact that we need the integer which has a large prime divisor to be inside a residue class $i \pmod{t}$ for which $r_i \neq 0$) and it is not trivial, but it can be done. If we then work this all out, we can bring the constants c in Theorems 4 and 5 one exponent down. Actually, when $r_i \neq 0$ for all i we do not even need the rather difficult result on large prime divisors in short intervals. We can just make do with Sylvester's Theorem, as explained below Theorem 11.

On another note, it can be conjectured that $b(a-1) > b(a)$ happens infinitely often, which might not be too hard to prove when $r_1 = t = 1$, or perhaps even in general. Other questions also remain in the classical special case of $r_1 = t = 1$. For example, it is still open if $\gcd(X_n, L_n) = 1$ holds for infinitely many n or not. This is equivalent to asking whether there are infinitely many n such that, if $l = l(p)$ is the first digit of n in base p, it holds that $\sum_{i=1}^{l} \frac{1}{i} \not\equiv 0 \pmod{p}$ for all $p < n$. So it is a natural question to ask how many l there are for a given p for which that sum can be equal to $0 \pmod{p}$. Even though this does not seem to immediately settle the original question, we do have the following (admittedly rather weak) result.

Lemma 41. *For any prime p there are at most $cp^{2/3}$ integers $l < p$ for which $f(l) = \sum_{i=1}^{l} \frac{1}{i} \equiv 0 \pmod{p}$, where $c = \frac{3^{2/3}}{2}$.*

We will just give the idea of the proof, as for a more complete write-up one can check Lemmas 2.3 and 2.4 in [12], where the same result with the same proof was found, independently.

Proof. For a fixed k, $f(l+k) - f(l)$ is a rational function in l where the numerator is a polynomial of degree at most $k-1$. So it has at most $k-1$ roots modulo p. Therefore, there can be at most $k-1$ integers l for which $f(l+k) \equiv f(l) \equiv 0$

(mod p). So if we have more than $1 + 2 + \ldots + k$ integers l for which $f(l) \equiv 0$ (mod p), then these integers cannot stay bounded inside an interval of length less than $1 \cdot 2 + 2 \cdot 3 + \ldots + k(k+1) > \frac{1}{3}(k + \frac{1}{2})^3$, which needs to be less than p. Therefore, $k + \frac{1}{2} < (3p)^{1/3}$ and $1 + 2 + \ldots + k < \frac{1}{2}(k + \frac{1}{2})^2 < \frac{1}{2}(3p)^{2/3} = cp^{2/3}$ with $c = \frac{3^{2/3}}{2}$. $\square$

References

[1] P. Erdős, R.L. Graham, *Old and New Problems and Results in Combinatorial Number Theory.* Enseign. Math. (2), vol. 28, Enseignement Math., Geneva, 1980. Also available here.

[2] P. Shiu, *The denominators of harmonic numbers.* Available here.

[3] A. Eswarathasan, E. Levine, *p-Integral harmonic sums.* Discr. Math. 91, 249–257, 1991. Also available here.

[4] J. Wolstenholme, *On certain properties of prime numbers.* The Quarterly Journal of Pure and Applied Mathematics 5, 35–39, 1862. Also available here.

[5] R. Meštrović, *Wolstenholme's Theorem: its generalizations and extensions in the last hundred and fifty years (1862 - 2012).* Available here.

[6] R. Wu, Q. Tan, and S. Hong, *New lower bounds for the least common multiple of arithmetic progressions.* Chinese Annals of Mathematics, 2012. Also available here.

[7] J. B. Rosser, L. Schoenfeld, *Approximate formulas for some functions of prime numbers.* Illinois J. Math. 6, no. 1, 64–94, 1962. Also available here.

[8] M. Laurent, M. Mignotte, Y. Nesterenko, *Formes linéaires en deux logarithmes et déterminants d'interpolation.* Journal of Number Theory, vol. 55, no. 2, 285–321, 1995. Also available here.

[9] S. Laishram, T. N. Shorey, *The greatest prime divisor of a product of terms in an arithmetic progression* Acta Arithmetica, 120(3), 299–306, 2005. Also available here.

[10] K. Ramachandra, *A note on numbers with a large prime factor.* J. London Math. Soc. 1, 303–306, 1969. Also available here.

[11] T. Xylouris, *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression.* Bonner Mathematische Schriften, 404, 2011. Entire thesis available here.

[12] B. Wu, Y. Chen, *On certain properties of harmonic numbers.* Journal of Number Theory, vol 175, 66–86, 2016. Also available here.

[13] H. Iwaniec, *On the problem of Jacobstahl.* Demonstratio Mathematica, Volume 11, Issue 1, 225–232, 1978. Also available here.

[14] P. Stevenhagen, H. W. Lenstra, *Chebotarëv and his density theorem.* Mathematical Intelligencer, Volume 18, Issue 2, 26–37, 1996. Also available here.

[15] C. L. Stewart, K. Yu *On the abc conjecture, II.* Duke Mathematical Journal, Volume 108, Number 1, 169–181, 2001. Also available here.

[16] W. van Doorn, W. Jagy, *What is the proportion of primes for which a polynomial has a root*. Question on math.stackexchange. See here.

[17] Also available here