

On the minimal length of addition chains

Jean-Marie De Koninck, Nicolas Doyon, William Verreault

Abstract

We denote by $\ell(n)$ the minimal length of an addition chain leading to n and we define the counting function

$$F(m, r) := \# \{n \in [2^m, 2^{m+1}) : \ell(n) \leq m + r\},$$

where m is a positive integer and $r \geq 0$ is a real number. We show that for $0 < c < \log 2$ and for any $\varepsilon > 0$, we have as $m \rightarrow \infty$,

$$F\left(m, \frac{cm}{\log m}\right) < \exp\left(cm + \frac{\varepsilon m \log \log m}{\log m}\right)$$

and

$$F\left(m, \frac{cm}{\log m}\right) > \exp\left(cm - \frac{(1 + \varepsilon)cm \log \log m}{\log m}\right).$$

This extends a result of Erdős which says that for almost all n , as $n \rightarrow \infty$,

$$\ell(n) = \frac{\log n}{\log 2} + (1 + o(1)) \frac{\log n}{\log \log n}.$$

AMS subject classification numbers: 11B83, 11Y55

Key words and phrases: Scholz conjecture, shortest addition chain, distribution of arithmetic functions

1 Introduction

We say that $1 = a_0 < a_1 < \dots < a_k = n$ is an *addition chain* of length k if $a_j = a_s + a_i$ with $0 \leq i, s < j$ for all j , $1 \leq j \leq k$. In other words, a term in an addition chain is obtained either by adding two distinct previous terms or by doubling a previous term.

For example, the following is an addition chain:

$$\begin{aligned} a_0 &= 1, \\ a_1 &= a_0 + a_0 = 1 + 1 = 2, \\ a_2 &= a_1 + a_1 = 2 + 2 = 4, \\ a_3 &= a_2 + a_0 = 4 + 1 = 5, \\ a_4 &= a_3 + a_1 = 5 + 2 = 7, \\ a_5 &= a_3 + a_2 = 5 + 4 = 9. \end{aligned}$$

A question of interest is: *Given a positive integer n , what is the shortest addition chain leading to n ?* For example, if $n = 10$, the following addition chain leads to 10: $a_0 = 1, a_1 = 1+1 = 2, a_2 = 2+2 = 4, a_3 = 4+4 = 8, a_4 = 8+2 = 10$. We can show that there exists no shorter addition chain leading to 10.

When n is a power of two, say $n = 2^r$, the shortest addition chain leading to n is of length r and consists only of doubling steps, since we have the addition chain $a_0 = 1, a_1 = 2a_0 = 2, a_2 = 2a_1 = 4, \dots, a_r = 2a_{r-1} = 2^r$.

Observe that the shortest addition chain leading to n is not necessarily unique. For instance, we have $a_0 = 1, a_1 = 2, a_3 = 4, a_4 = 5$ as well as $b_0 = 1, b_1 = 2, b_2 = 3, b_4 = 5$ both being chains of minimal length leading to 5.

Definition 1.1. *We denote by $\ell(n)$ the length of a shortest addition chain leading to n .*

In this note, we investigate a question related to the distribution of $\ell(n)$: *For a given positive integer k , how many integers n can be obtained by an addition chain of length at most k ?* For example, if we take $k = 3$, the following integers can be obtained by addition chains of length at most 3:

$$\begin{aligned} a_0 &= \mathbf{1}, \\ a_0 &= 1, a_1 = \mathbf{2}, \\ a_0 &= 1, a_1 = 2, a_2 = \mathbf{3}, \\ a_0 &= 1, a_1 = 2, a_2 = \mathbf{4}, \\ a_0 &= 1, a_1 = 2, a_2 = 3, a_3 = \mathbf{5}, \\ a_0 &= 1, a_1 = 2, a_2 = 3, a_3 = \mathbf{6}, \\ a_0 &= 1, a_1 = 2, a_2 = 4, a_3 = \mathbf{8}. \end{aligned}$$

It follows that **7** different integers can be obtained by addition chains of length at most 3.

Definition 1.2. *We denote by $H(k)$ the number of distinct integers that can be obtained by an addition chain of length at most k . Explicitly,*

$$H(k) := \#\{n : \ell(n) \leq k\}.$$

Given that the largest integer that can be obtained by an addition chain of length k is 2^k , we trivially have

$$H(k) \leq 2^k.$$

It turns out that, in some sense, this simple upper bound is relatively close to the truth. The sequence $H(n)$ begins with

$$H(1) = 2, H(2) = 4, H(3) = 7, H(4) = 12, \dots$$

Doubling steps, that is steps given by $a_r = 2a_{r-1}$, play an important role in addition chains since for addition chains of minimal length, most of the steps

must be doubling steps. It is therefore natural to work with intervals of the form $[2^m, 2^{m+1})$. Instead of working directly with the function H , we will investigate a closely related function.

Definition 1.3. *Let m be a positive integer and r a positive real number. We define*

$$F(m, r) := \# \{n \in [2^m, 2^{m+1}) : \ell(n) \leq m + r\}.$$

We will investigate the asymptotic behaviour of the function F . A critical region is given by

$$r = \frac{cm}{\log m},$$

where $0 < c < \log 2$ is a fixed constant and $m \rightarrow \infty$. As we will discuss in the next section, previous results imply that $F(m, r) = 2^m$ when $c > \log 2$ and m is large enough.

Relying on this notation, we will show

$$F\left(m, \frac{cm}{\log m}\right) = \exp(cm(1 + o(1))),$$

as well as more precise results.

2 Previous work

Lower and upper bounds for the function $\ell(n)$ have been obtained.

Arnold Scholz [6] seems to have been the first (in 1937) to study the length of an addition chain. In his paper, one can find what is now known as the *Scholz conjecture*, namely that

$$\ell(2^{n+1} - 1) \leq n + \ell(n + 1).$$

This conjecture has been shown to be true for all $n \leq 5\,784\,688$ by Clift [3]. This conjecture has also been shown to be true for infinite families of integers, see [1] and [8].

Observe that it is easy to establish that

$$\ell(n) \leq \log_2 n + \nu(n),$$

where $\log_2 m = \log m / \log 2$ and $\nu(n)$ is the number of digits equal to one in the binary expansion of n . The best result in the direction of the Scholz conjecture is the one obtained in 1939 by Arnold T. Brauer [2], namely the upper bound

$$\ell(2^{n+1} - 1) \leq n + \ell^*(n + 1),$$

where $\ell^*(n)$ is the minimum length of only those addition chains in which every a_p can be written as $a_{p-1} + a_r$, $r \leq p - 1$. In fact, Brauer proved that

$$\ell(n) \leq \min_{1 \leq r \leq m} \left\{ \left(1 + \frac{1}{r}\right) \log_2 n + 2^r - 2 \right\},$$

where $2^m \leq n < 2^{m+1}$. In Brauer's upper bound, choosing $r = (1 - \varepsilon) \frac{\log \log n}{\log 2}$, it follows that for n sufficiently large,

$$\ell(n) \leq \log_2 n + (1 + \varepsilon) \frac{\log n}{\log \log n}. \quad (2.1)$$

In 1960, Erdős [4] showed that (2.1) is best possible in the sense that for almost all n (i.e., for all n except a sequence of density 0),

$$\ell(n) = \log_2 n + (1 + o(1)) \frac{\log n}{\log \log n}.$$

In 1975, Arnold Schönhage [7] proved that

$$\ell(n) \geq \log_2 n + \log_2 \nu(n) - 2.13.$$

To summarize, what has been proved so far unconditionally is that

$$\log_2 n + \log_2 \nu(n) - 2.13 \leq \ell(n) \leq \lceil \log_2 n \rceil + \nu(n) - 1.$$

Providing a more concrete bound, Thurber [10] showed that if $\nu(n) \geq 9$, then

$$\ell(n) \geq \lfloor \log_2 n \rfloor + 4.$$

The problem of numerically computing $\ell(n)$ for a given n is also an important challenge. Methods to effectively generate an addition chain of minimal length leading to a given integer can be found in [9].

A broad view of the results regarding addition chains can be found in the book of Richard Guy [5, p.169–171].

3 Result statement

Our main results consist in an upper bound and a lower bound for the function F .

Theorem 3.1. *For any $\varepsilon > 0$ and for $0 < c < \log 2$, we have for m large enough*

$$F\left(m, \frac{cm}{\log m}\right) < \exp\left(cm + \frac{\varepsilon m \log \log m}{\log m}\right).$$

Theorem 3.2. *For any $\varepsilon > 0$ and for $0 < c < \log 2$, we have for m large enough*

$$F\left(m, \frac{cm}{\log m}\right) > \exp\left(cm - \frac{c(1 + \varepsilon)m \log \log m}{\log m}\right).$$

These results partially answer a comment of Paul Erdős [4] who wrote that it would be interesting to obtain an asymptotic value for the distribution of the function $\ell(n)$.

4 Proof of the upper bound

The proof of the upper bound is the most complicated of the two and accordingly, it is divided into several subsections.

4.1 Notation and definitions

We begin the proof of the upper bound by providing a few more definitions. We denote by γ the golden ratio

$$\gamma := \frac{1 + \sqrt{5}}{2} \approx 1.618.$$

Definition 4.1. For a fixed integer $m \geq 2$, we define

$$\delta = \delta(m) := \frac{1}{\log m}.$$

We expand upon the ideas of Erdős as we will divide the addition steps into several types.

Definition 4.2. Let $m > 4$ be a positive integer and r a positive real number. Consider an addition chain consisting of $1 = a_0 < \dots < a_{\lfloor m+r \rfloor} = n$ with $n \in [2^m, 2^{m+1})$. Let the set $\mathcal{A}$ consist of the doubling steps of this chain. We subdivide the nondoubling steps into three sets $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ according to their associated relative growth rate:

$$\begin{aligned} \mathcal{A} &:= \{1 \leq j \leq m+r : a_j = 2a_{j-1}\}, \text{ the doubling steps,} \\ \mathcal{B} &:= \{1 \leq j \leq m+r : \gamma a_{j-1} \leq a_j < 2a_{j-1}\}, \text{ the large steps,} \\ \mathcal{C} &:= \{1 \leq j \leq m+r : (1+\delta)a_{j-1} \leq a_j < \gamma a_{j-1}\}, \text{ the midsize steps,} \\ \mathcal{D} &:= \{1 \leq j \leq m+r : a_j < (1+\delta)a_{j-1}\}, \text{ the small steps.} \end{aligned}$$

We denote their respective cardinalities by

$$A := \#\mathcal{A}, B := \#\mathcal{B}, C := \#\mathcal{C}, D := \#\mathcal{D}.$$

4.2 Remarks on the size of the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$

We now obtain bounds on A , B , C and D .

Lemma 4.1. Suppose that an integer $n \in [2^m, 2^{m+1})$ is obtained by an addition chain of length $\lfloor m+r \rfloor$ such that the number of steps of each type is given by A , B , C and D , respectively. Then

$$B + C + D \leq r \frac{1}{1 - \log_2 \gamma}.$$

Proof. In [7], it was shown by induction that

$$n \leq 2^A \gamma^{B+C+D}. \quad (4.1)$$

The intuition being that a chain with additive (nondoubling) steps cannot grow faster than the Fibonacci sequence.

Using the fact that $A + B + C + D = \lfloor m + r \rfloor$ and $n \geq 2^m$, we obtain from (4.1) that

$$2^m \leq 2^{m+r-(B+C+D)} \gamma^{B+C+D},$$

which yields

$$1 \leq 2^r \left(\frac{\gamma}{2} \right)^{B+C+D}.$$

From this, we get

$$B + C + D \leq r \frac{1}{1 - \log_2 \gamma},$$

completing the proof of Lemma 4.1. $\square$

We also have the following upper bound on the size of the set $\mathcal{D}$.

Lemma 4.2. *Suppose that an integer $n \in [2^m, 2^{m+1})$ is obtained by an addition chain of length $\lfloor m + r \rfloor$ such that the number of steps of each type is given by A, B, C and D , respectively. We have*

$$D \leq \frac{r - C(1 - \log_2 \gamma)}{1 - \log_2(1 + \delta)}.$$

Proof. Using the definition of the sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$, we have

$$\begin{aligned} 2^m &\leq n \leq 2^{A+B} \gamma^C (1 + \delta)^D \\ &\leq 2^{m+r-C-D} \gamma^C (1 + \delta)^D. \end{aligned}$$

From this, we get

$$1 \leq 2^r \left(\frac{\gamma}{2} \right)^C \left(\frac{1 + \delta}{2} \right)^D.$$

Taking the base 2 logarithm of each side of the inequality yields

$$0 \leq r - C(1 - \log_2 \gamma) - D(1 - \log_2(1 + \delta)),$$

which we can rearrange to conclude

$$D \leq \frac{r - C(1 - \log_2 \gamma)}{1 - \log_2(1 + \delta)},$$

thus proving Lemma 4.2. $\square$

4.3 The number of ways of choosing the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$

The sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ specify the *position* in the addition chain of the steps of a given type.

We define

$$S = S(A, B, C, D, m, r)$$

as the number of ways of choosing these sets for given values of A, B, C, D, m and r .

It is trivial that

$$S \leq \binom{\lfloor m+r \rfloor}{A, B, C, D}$$

where we use the multinomial coefficient notation:

$$\binom{\lfloor m+r \rfloor}{A, B, C, D} = \frac{\lfloor m+r \rfloor!}{A!B!C!D!}$$

with $\lfloor m+r \rfloor = A + B + C + D$.

We can however improve upon this by observing the following result.

Lemma 4.3. *A step in $\mathcal{B}$ must be preceded by a step in either $\mathcal{C}$ or $\mathcal{D}$.*

Proof. Suppose that $j \in \mathcal{B}$. Since a step in $\mathcal{B}$ is not a doubling step, we have

$$a_j \leq a_{j-1} + a_{j-2}.$$

By definition of the set $\mathcal{B}$, we also have

$$a_j \geq \gamma a_{j-1},$$

and since a_j, a_{j-1} are integers while γ is irrational, it must be that $a_j > \gamma a_{j-1}$. Thus

$$\gamma a_{j-1} < a_{j-1} + a_{j-2},$$

from which we conclude

$$a_{j-1} < \frac{1}{\gamma-1} a_{j-2} = \gamma a_{j-2}$$

and

$$j-1 \in \mathcal{C} \cup \mathcal{D}. \quad \square$$

From the above, we can conclude the following bound.

Lemma 4.4. *Let m be a positive integer and r a positive real number such that $r = cm/\log m$ with c a fixed constant, $0 < c < \log 2$. We have as $m \rightarrow \infty$*

$$S \leq \binom{\lfloor m+r \rfloor}{C+D} \exp(O(r)).$$

Proof. The number of ways of choosing the sets $\mathcal{C}$ and $\mathcal{D}$ is given by

$$\binom{\lfloor m+r \rfloor}{C+D} \binom{C+D}{D}.$$

By Lemma 4.3, the number of ways to choose the set $\mathcal{B}$ is less than or equal to

$$\binom{B+C+D-1}{B-1}.$$

We have thus shown

$$S \leq \binom{\lfloor m+r \rfloor}{C+D} \binom{C+D}{D} \binom{B+C+D-1}{B-1}.$$

This implies

$$S \leq \binom{\lfloor m+r \rfloor}{C+D} 2^{C+D} 2^{B+C+D}.$$

Given that from Lemma 4.1, $B+C+D = O(r)$, we conclude

$$S \leq \binom{\lfloor m+r \rfloor}{C+D} \exp(O(r)).$$

□

We will also make use of the following lemma.

Lemma 4.5. *Assume as above that m is a positive integer and that r is a positive real number with $r = cm/\log m$ and c a fixed constant such that $0 < c < \log 2$. Further assume that the addition steps are grouped in k sequences of consecutive addition steps. Then*

$$S \leq \binom{\lfloor m+r \rfloor}{k} \exp(O(r))$$

as $m \rightarrow \infty$.

Proof. Suppose that the steps in $\mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$ are grouped into k blocks and that the length of each block is respectively $b_1, b_2, \dots, b_k$. We then have

$$b_1 + b_2 + \dots + b_k = B + C + D.$$

The number of ways of choosing the integers $b_1, b_2, \dots, b_k$ is less than or equal to

$$2^{B+C+D} = \exp(O(r)).$$

The number of ways of choosing if the addition steps belong to $\mathcal{B}$, $\mathcal{C}$ or $\mathcal{D}$ is less than or equal to

$$3^{B+C+D} = \exp(O(r)).$$

It remains to specify how the A doubling steps will be distributed amongst the k additive blocks. This is equivalent to choosing a sequence of integers $t_1, t_2, \dots, t_{k+1}$ in such a way that

$$t_1 + t_2 + \dots + t_{k+1} = A$$

with $t_1 \geq 0$, $t_{k+1} \geq 0$ and $t_j > 0$, $2 \leq j \leq k$. The number of ways of doing so is less than or equal to

$$\binom{A+k}{k} = \binom{\lfloor m+r \rfloor}{k},$$

thus completing the proof of Lemma 4.5. □

4.4 The number of ways of choosing the added elements at each step

To completely specify an addition chain, we still have to choose which elements will be added at each step.

First observe that for doubling steps in $\mathcal{A}$ there is no choice to make, since if $j \in \mathcal{A}$ then necessarily $a_j = 2a_{j-1}$. We start by proving the following result.

Lemma 4.6. *Let m be a positive integer and r a positive real number such that $r = cm/\log m$ with c a fixed constant, $0 < c < \log 2$. Let $1 = a_0 < \dots < a_{\lfloor m+r \rfloor} = n$ be an addition chain of length $\lfloor m+r \rfloor$ leading to n with $n \in [2^m, 2^{m+1})$. Assume that the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are fixed. Denote by R the number of ways of choosing which integers will be added in steps belonging to $\mathcal{B}$. Then,*

$$R = \exp(O(r))$$

as $m \rightarrow \infty$.

Proof. Suppose that $j \in \mathcal{B}$ and that

$$a_j = a_s + a_k, \quad j > s \geq k, \quad k < j-1.$$

We have

$$\gamma a_{j-1} \leq a_j = a_s + a_k \leq a_{j-1} + a_k,$$

from which we deduce

$$a_k \geq (\gamma - 1)a_{j-1}.$$

It follows that all u such that $k < u < j$ are in $\mathcal{C} \cup \mathcal{D}$ as assuming the contrary would yield $a_{j-1} > a_k$. The number of ways of choosing s and k is thus less than or equal to $h(j)^2$ where for $j \in \mathcal{B}$,

$$h(j) := \max_{t \geq 1} \{j - u \in \mathcal{C} \cup \mathcal{D} \text{ for all } 1 \leq u \leq t\}.$$

In other words, $h(j)$ is the maximal integer t such that all the integers in the interval $[j - t, j - 1]$ are in the set $\mathcal{C} \cup \mathcal{D}$. We thus have

$$R \leq \prod_{j \in \mathcal{B}} h(j)^2. \quad (4.2)$$

From Lemma 4.3, we have

$$\sum_{j \in \mathcal{B}} h(j) \leq C + D. \quad (4.3)$$

From (4.3) and the arithmetic-geometric mean inequality, we have

$$\prod_{j \in \mathcal{B}} h(j)^{1/B} \leq \frac{1}{B} \sum_{j \in \mathcal{B}} h(j) \leq \frac{C + D}{B}.$$

From (4.2), we can conclude

$$R \leq \left(\frac{C + D}{B} \right)^{2B} = \exp(O(r)). \quad \square$$

We will now prove an upper bound for the number of ways of choosing the integers which will be added in steps belonging to $\mathcal{C}$. This turns out to be one of the most technical steps in the proof.

Lemma 4.7. *Let m be a positive integer and r a positive real number such that $r = cm / \log m$ with c a real constant such that $0 < c < \log 2$. Let $1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor} = n$ with $n \in [2^m, 2^{m+1})$ be an addition chain of length $\lfloor m + r \rfloor$. Assume that the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are fixed. Denote by T the number of ways of choosing which integers will be added in steps belonging to $\mathcal{C}$. For any $\varepsilon > 0$, we have as $m \rightarrow \infty$*

$$T \leq (\log m)^{(1+\varepsilon)2C} \exp(O(r)).$$

Proof. The intuitive idea behind the proof of Lemma 4.7 is that for steps j in $\mathcal{C}$, the number of choices will be limited since if $a_j = a_k + a_s$, then neither a_k nor a_s can be very small (smaller than δa_j).

We denote by $\beta_1, \dots, \beta_C$ the elements of the set $\mathcal{C}$ in increasing order. For $1 \leq s \leq C$, we write

$$a_{\beta_s} = a_{\beta'_s} + a_{\beta''_s},$$

where β'_s and β''_s are not necessarily in $\mathcal{C}$.

We will now get an upper bound for the number of ways of choosing the values of β'_s and β''_s .

Given that $\beta_s \in \mathcal{C}$, we have $a_{\beta_s} \geq (1 + \delta)a_{\beta_{s-1}}$. It follows that $a_{\beta'_s}, a_{\beta''_s} \leq a_{\beta_{s-1}} \leq (1 + \delta)^{-1}a_{\beta_s}$. From this and from $a_{\beta'_s} + a_{\beta''_s} = a_{\beta_s}$, we get

$$a_{\beta'_s} = a_{\beta_s} - a_{\beta''_s} \geq a_{\beta_s} - \frac{1}{1 + \delta}a_{\beta_s} = a_{\beta_s} - (1 - \delta + \delta^2 - \dots)a_{\beta_s} \geq \delta a_{\beta_s}.$$

Similarly, we have $a_{\beta'_s} \geq \delta a_{\beta_s}$.

For $1 \leq s \leq C$, let

$$u(s) := \#\{j : \delta a_{\beta_s} \leq a_j \leq a_{\beta_s}, j \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}\}.$$

Then

$$(1 + \delta)^{u(s)} \leq \frac{1}{\delta},$$

which yields

$$u(s) \leq \frac{-\log \delta}{\log(1 + \delta)} \leq \frac{-2 \log \delta}{\delta} \quad (4.4)$$

provided that δ is small enough (which is the case if m is large enough). Let $t = t(s)$ be the smallest integer such that

$$\#\{j : t(s) \leq j \leq \beta_s, j \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}\} \leq \frac{-2 \log \delta}{\delta}.$$

The interest of this is that β'_s and β''_s must be chosen in the interval $[t(s), \beta_s]$.

Therefore, the number of integers in this interval is at most $\frac{-2 \log \delta}{\delta} + \eta(s)$, where

$$\eta(s) := \#\{j : t(s) \leq j \leq \beta_s, j \in \mathcal{D}\}.$$

Hence, the number of ways of choosing the integers β'_s and β''_s for $1 \leq s \leq C$ is less than or equal to

$$\prod_{s=1}^C \left(\frac{-2 \log \delta}{\delta} + \eta(s) \right)^2.$$

We have

$$\sum_{s=1}^C \left(\frac{-2 \log \delta}{\delta} + \eta(s) \right) = \frac{-2C \log \delta}{\delta} + \sum_{s=1}^C \eta(s).$$

Because each element in $\mathcal{D}$ is counted in at most $\frac{-2 \log \delta}{\delta}$ sets of the form $\{j : t(s) \leq j \leq \beta_s, j \in \mathcal{D}\}$, we have

$$\sum_{s=1}^C \eta(s) \leq \frac{-2D \log \delta}{\delta}.$$

We can conclude

$$\sum_{s=1}^C \left(\frac{-2 \log \delta}{\delta} + \eta(s) \right) \leq \frac{-2(C + D) \log \delta}{\delta}.$$

By the arithmetic-geometric mean inequality, we get

$$\prod_{s=1}^C \left(\frac{-2 \log \delta}{\delta} + \eta(s) \right)^{1/C} \leq \frac{1}{C} \sum_{s=1}^C \left(\frac{-2 \log \delta}{\delta} + \eta(s) \right)$$

$$\leq \frac{-2(C+D)\log \delta}{C\delta}.$$

Raising both sides to the power $2C$, it follows that

$$\prod_{s=1}^C \left(\frac{-2\log \delta}{\delta} + \eta(s) \right)^2 \leq \left(\frac{-2\log \delta}{\delta} \right)^{2C} \left(\frac{C+D}{C} \right)^{2C}. \quad (4.5)$$

Writing $C = \xi D$, we get

$$\begin{aligned} \left(\frac{C+D}{C} \right)^{2C} &= \left(\frac{\xi+1}{\xi} \right)^{2\xi D} \\ &= \exp(2\xi D \log(1 + 1/\xi)) \\ &\leq \exp(2D). \end{aligned}$$

From Lemma 4.1, it follows that

$$\left(\frac{C+D}{C} \right)^{2C} = \exp(O(r)).$$

From the definition of δ , we have for any $\varepsilon > 0$ and for m large enough,

$$\left(\frac{-2\log \delta}{\delta} \right)^{2C} = (2\log m \log \log m)^{2C} \leq (\log m)^{(1+\varepsilon)2C}.$$

We can thus conclude

$$T \leq (\log m)^{(1+\varepsilon)2C} \exp(O(r)). \quad \square$$

Finally, we provide an upper bound for the number of ways of choosing which integers need to be added in steps belonging to $\mathcal{D}$.

Lemma 4.8. *Let m be a positive integer and let r be a positive real number such that $r = cm/\log m$ with c a fixed constant, $0 < c < \log 2$. Let $1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor} = n$ be an addition chain of length $\lfloor m+r \rfloor$ leading to n with $n \in [2^m, 2^{m+1})$. Assume that the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are fixed. The number of ways of choosing the integers added in steps belonging to $\mathcal{D}$ is less than or equal to*

$$\exp(D \log m + (1 + o(1))D \log \log m)$$

as $m \rightarrow \infty$.

Proof. We follow the approach used by Erdős in [4]. Assume that $j \in \mathcal{D}$ with $a_j = a_s + a_k$. The pair (s, k) must be chosen in the box $[1, m+r] \times [1, m+r]$ which contains $\lfloor m+r \rfloor^2$ elements. Thus, the number of ways of choosing which integers are added in steps of $\mathcal{D}$ is less than or equal to

$$\binom{\lfloor m+r \rfloor^2}{D}.$$

From this, we can prove Lemma 4.8 by observing

$$\binom{\lfloor m+r \rfloor^2}{D} \leq \left(\frac{e(m+r)^2}{D} \right)^D \leq \exp(D \log m + (1 + o(1))D \log \log m). \quad \square$$

4.5 Some intermediate results

We now have the tools to obtain some intermediate results that will take us closer to the proof of the upper bound.

Definition 4.3. Let m be a positive integer and r a positive real number. We denote by

$$N = N(A, B, C, D, m, r)$$

the number of distinct addition chains of length $\lfloor m + r \rfloor$ leading to an integer $n \in [2^m, 2^{m+1})$ with A, B, C, D fixed.

The results of the previous subsections will yield an upper bound on N .

Lemma 4.9. Let m a positive integer and let r be a positive real number with $r = cm/\log m$ where c is a fixed constant with $0 < c < \log 2$. We have

$$N(A, B, C, D, m, r) \leq \exp(D \log m + (1 + o(1))(2D + 3C) \log \log m)$$

as $m \rightarrow \infty$.

Proof. The proof of Lemma 4.9 follows directly from Lemmas 4.4, 4.6, 4.7 and 4.8. $\square$

From Lemmas 4.2 and 4.9, we have

$$\begin{aligned} N &\leq \exp\left(\frac{r - C(1 - \log_2 \gamma)}{1 - \log_2(1 + \delta)} \log m + (1 + o(1)) \left(2 \frac{r - C(1 - \log_2 \gamma)}{1 - \log_2(1 + \delta)} + 3C\right) \log \log m\right) \\ &\leq \exp\left(\frac{r - C(1 - \log_2 \gamma)}{1 - \log_2(1 + \delta)} \log m + (1 + o(1)) (2r + 3C) \log \log m\right). \end{aligned} \quad (4.6)$$

From (4.6), we can obtain the following upper bound which, though weaker than Theorem 3.1, will be useful.

Lemma 4.10. Let m be a positive integer and $0 < c < \log 2$ a real number. Then, for any $\varepsilon > 0$ and for m large enough,

$$F\left(m, \frac{cm}{\log m}\right) \leq \exp\left(cm + \frac{(2 + \varepsilon)cm \log \log m}{\log m}\right).$$

Proof. Assume that $r = cm/\log m$ with $0 < c < \log 2$. From inequality (4.6), we have as $m \rightarrow \infty$,

$$N(A, B, C, D, m, r) \leq \exp(r \log m + (1 + o(1))2r \log \log m).$$

From this, we conclude that

$$\begin{aligned} F(m, r) &\leq \sum_{A, B, C, D \leq m+r} N(A, B, C, D, m, r) \\ &\leq (m + r)^4 \max_{A, B, C, D \leq m+r} N(A, B, C, D, m, r) \\ &\leq \exp(r \log m + (1 + o(1))2r \log \log m). \end{aligned}$$

Using $r = cm/\log m$ completes the proof of Lemma 4.10. $\square$

It is interesting to note that from Lemma 4.10 and inequality (4.6), we can conclude that addition chains with a large value of C , for instance chains such that

$$C < \frac{7r \log \log m}{(1 - \log_2 \gamma) \log m}, \quad (4.7)$$

have a negligible contribution to $F(m, r)$.

4.6 Addition blocks

We divide the addition steps (that is, steps in $\mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$) in blocks of consecutive addition steps. Say we have K blocks of lengths

$$L_1, L_2, \dots, L_K$$

with

$$L_1 + L_2 + \dots + L_K = B + C + D.$$

Suppose that

$$a_j, \dots, a_{j+L-1}$$

is a sequence of $L > 1$ consecutive addition steps. That is, suppose that $s \in \mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$ for $j \leq s \leq j + L - 1$ and that $a_{j-1} \in \mathcal{A}$, $a_{j+L} \in \mathcal{A}$.

Then, after $a_{j-1} = 2a_{j-2}$, the sequence begins with

$$a_j = a_{j-1} + a_t, \quad t < j - 1.$$

Indeed, if $a_j = a_s + a_t$ with $s, t < j - 1$, we would have $a_j \leq a_{j-2} + a_{j-2} = a_{j-1}$, a contradiction. For s such that $j \leq s \leq j + L - 1$, if we write

$$a_s = a_v + a_t, \quad v \geq t,$$

then $v \geq j - 1$ or else $a_s \leq a_{j-1}$. Indeed, assuming the contrary yields

$$a_s = a_v + a_t \leq a_{j-2} + a_{j-2} = a_{j-1},$$

a contradiction.

We will use this observation and the block structure to refine the upper bound on the number of ways of choosing the integers that are in steps belonging to $\mathcal{D}$.

Lemma 4.11. *Consider a block of L consecutive additive steps with L' steps in $\mathcal{D}$ and L'' steps in $\mathcal{B} \cup \mathcal{C}$ in an addition chain of length $\lfloor m + r \rfloor$. Then, the number of ways of choosing which integers are added in steps of type $\mathcal{D}$ of this block is less than or equal to*

$$\binom{(L' + L'') \lfloor m + r \rfloor}{L'}.$$

Proof. Suppose that the addition block ranges from a_j to a_{j+L-1} , that $j \leq s \leq j+L-1$ and that $a_s = a_v + a_t$ with $t \leq v$. The pair (t, v) must be chosen in the set $[1, m+r] \times [j-1, j+L-2]$. Notice that we use the fact that $v \in [j-1, j+L-2]$ but don't use the condition $t \leq v$ as it would not lead to significant improvements. The cardinality of the set $[1, m+r] \times [j-1, j+L-2]$ is $L[m+r] = (L' + L'')[m+r]$. Since we must choose L' distinct pairs (t, v) , the number of ways of making this choice is less than or equal to

$$\binom{(L' + L'')[m+r]}{L'}. \quad \square$$

Using Lemma 4.11, we can improve the bound obtained in Lemma 4.8.

Lemma 4.12. *Let m be a positive integer and r a positive real number such that $r = cm/\log m$ with c a fixed constant, $0 < c < \log 2$. Consider $1 = a_0 < a_1 < \dots < a_{[m+r]}$ an addition chain of length $[m+r]$ leading to n with $n \in [2^m, 2^{m+1})$. Assume that the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ are fixed. The number of ways of choosing which integers are added in steps belonging to $\mathcal{D}$ is less than or equal to*

$$\exp(D \log m + O(m/\log m)).$$

Proof. Assume that the addition steps are divided in k blocks of lengths $L_1, \dots, L_k$. Assume that the number of steps in $\mathcal{D}$ in the block s , with $1 \leq s \leq k$, is L'_s and that the number of steps in $\mathcal{B} \cup \mathcal{C}$ in this block is L''_s . From Lemma 4.11, the number of ways of choosing the integers added in steps belonging to $\mathcal{D}$ is less than or equal to

$$\begin{aligned} \prod_{1 \leq s \leq k} \binom{(L'_s + L''_s)[m+r]}{L'_s} &\leq \prod_{1 \leq s \leq k} \left(\frac{e(L'_s + L''_s)(m+r)}{L'_s} \right)^{L'_s} \\ &= (e(m+r))^D \prod_{1 \leq s \leq k} \left(1 + \frac{L''_s}{L'_s} \right)^{L'_s} \\ &\leq (e(m+r))^D \exp \left(\sum_{s=1}^k L''_s \right) \\ &= (e(m+r))^D e^{B+C} \\ &= \exp(D \log m + O(m/\log m)). \end{aligned}$$

For the last equality, we used the fact that $B+C = O(r)$ and $r = O(m/\log m)$. $\square$

Using Lemma 4.12, we can improve Lemma 4.10 and obtain the following result.

Lemma 4.13. *Let m be a positive integer and $c, 0 < c < \log 2$, be a real constant. For any $\varepsilon > 0$, we have for m large enough*

$$F\left(m, \frac{cm}{\log m}\right) \leq \exp\left(cm + \frac{(1+\varepsilon)cm \log \log m}{\log m}\right).$$

While this result is still weaker than Theorem 3.1, we are getting closer.

Proof. The proof of Lemma 4.13 follows rather directly from our previous results. The number of ways of choosing A , B , C and D is $O(m^4)$. By Lemma 4.4, the number of ways of choosing the sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ is less than or equal to

$$\binom{\lfloor m+r \rfloor}{C+D} \exp(O(r)).$$

By Lemmas 4.6 and 4.7, for any $\varepsilon > 0$ and for m large enough, the number of ways of choosing the integers added in steps belonging to $\mathcal{B} \cup \mathcal{C}$ is less than or equal to

$$(\log m)^{(1+\varepsilon)2^C} \exp(O(r)).$$

Using Lemma 4.12 and assuming $r = \frac{cm}{\log m}$, we obtain

$$\begin{aligned} F\left(m, \frac{cm}{\log m}\right) &\leq \binom{\lfloor m+r \rfloor}{C+D} \exp(D \log m) \exp(O(r)) \\ &= \binom{\lfloor m+r \rfloor}{C+D} \exp(D \log m) \exp\left(O\left(\frac{m}{\log m}\right)\right). \end{aligned} \quad (4.8)$$

Finally, Lemma 4.13 follows from (4.8) and Lemma 4.2. $\square$

We end this subsection by observing that from the above results, the number of addition chains of length $\lfloor m+r \rfloor$ leading to n with $n \in [2^m, 2^{m+1})$ and with k addition blocks is less than or equal to

$$\binom{\lfloor m+r \rfloor}{k} \exp\left(D \log m + o\left(\frac{m \log \log m}{\log m}\right)\right).$$

This implies that if $k = o(m/\log m)$, the proof of Theorem 3.1 is complete. In the rest of the demonstration, we will thus assume that there is a positive constant c_1 such that $k > c_1 m/\log m$.

4.7 Minimal addition chains

As discussed in the Introduction, several addition chains of the same length can lead to the same integer. Thus far, we did not distinguish between such chains. To make further progress, we will need a few additional definitions.

Definition 4.4. Let m, n be positive integers and let r be a positive real number. Let

$$1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor} = n$$

and

$$1 = b_0 < b_1 < \dots < b_{\lfloor m+r \rfloor} = n$$

be two addition chains of the same length leading to the same integer.

Assume that $b_j \leq a_j$ for all j . Assume further that there exists an s , $1 \leq s \leq m+r$, such that $b_s < a_s$.

In this case, we say that the chain $a_0, \dots, a_{\lfloor m+r \rfloor}$ **dominates** the addition chain $b_0, \dots, b_{\lfloor m+r \rfloor}$.

If an addition chain dominates another addition chain of the same length and leads to the same integer, we say that the first chain is **not valid**. Our goal is to count only valid addition chains.

For example, consider the two addition chains $a_0 = 1, a_1 = 2, a_3 = 4, a_4 = 5$ and $b_0 = 1, b_1 = 2, b_3 = 3, b_4 = 5$. The chain a_j , $0 \leq j \leq 4$, dominates the chain b_j , $0 \leq j \leq 4$, and thus is not valid.

We also define the notion of a *marked* element of an addition chain.

Definition 4.5. Let $1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor} = n$ be an addition chain. We say that the step s is **marked** if there exists a $j > s+1$ such that $a_j = a_s + a_t$ for some t or else the addition chain cannot be a valid chain of minimal length.

We provide a few examples to illustrate this notion. Suppose that an addition chain begins with

$$a_0 = 1, a_1 = 2, a_2 = 4, a_3 = 5, a_4 = 8, \dots$$

The step 3 ($a_3 = 5$) is **marked** since if it is not used in a later step of the chain, it can simply be removed from the addition chain which would yield a chain of smaller length leading to the same integer.

We now give another example that will be more relevant in the rest of the proof. Suppose that an addition chain begins with

$$a_0 = 1, a_1 = 2, a_2 = 4, a_3 = 8, a_4 = 10, a_5 = 20, \dots$$

then the step 3 ($a_3 = 8$) is marked since if it is not used in a future step of the chain, the chain could be replaced by

$$b_0 = 1, b_1 = 2, b_2 = 4, b_3 = 5, b_4 = 10, b_5 = 20, b_j = a_j, j > 5.$$

Since the chain a_j dominates the chain b_j , it is not valid. The previous example can be generalized. Assume that an addition chain contains the sequence

$$a_0 = 1, \dots, a_s = 2a_{s-1}, \dots, a_j = 2a_{j-1}, a_{j+1} = a_j + a_s, \dots,$$

then the step a_j is marked. Indeed if a_j is not used in a step following a_{j+1} , we can replace the addition chain $a_0, a_1, \dots, a_{\lfloor m+r \rfloor}$, by the chain $b_0, b_1, \dots, b_{\lfloor m+r \rfloor}$, defined as $b_t = a_t$, $0 \leq t \leq j-1$, $b_j = b_{j-1} + b_{s-1}$, $b_{j+1} = 2b_j$ as well as $b_t = a_t$, $t > j+1$. The important point is to observe that

$$b_{j+1} = 2b_j = 2b_{j-1} + 2b_{s-1} = 2a_{j-1} + 2a_{s-1} = a_{j+1}.$$

We extend the definition of a marked step of a chain to a marked addition block.

Definition 4.6. Let $1 = a_0 < a_1 < \dots < a_{m+r}$ be an addition chain. Assume that this chain contains the sequence

$$a_j = 2a_{j-1} < a_{j+1} < \dots < a_{j+L} < a_{j+L+1} = 2a_{j+L},$$

where $a_{j+1}, \dots, a_{j+L} \in \mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$. We say that the addition block $a_{j+1}, \dots, a_{j+L}$ is **marked** if at least one element preceding an element of this block must be used in a subsequent block for the chain to be valid.

We illustrate this definition with an example. Consider an addition chain that begins with the sequence

$$\begin{aligned} a_0 &= 1, a_1 = 2a_0 = 2, a_2 = 2a_1 = 4, a_3 = 2a_2 = 8, a_4 = a_3 + a_1 = 10, \\ a_5 &= a_4 + a_2 = 14, a_6 = 2a_5 = 28, \dots \end{aligned}$$

In this example, the addition block a_4, a_5 is marked. Indeed, assume that neither a_3 nor a_4 is used in a subsequent addition block, then the beginning of the chain can be replaced by

$$\begin{aligned} b_0 &= 1, b_1 = 2b_0 = 2, b_2 = 2b_1 = 4, b_3 = b_2 + b_0 = 5, b_4 = b_3 + b_1 = 7, \\ b_5 &= 2b_4 = 14, b_6 = 2b_5 = 28, \dots \end{aligned}$$

This example can be generalized into the following result.

Lemma 4.14. Let $1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor}$ be an addition chain. Let $a_{j+1}, \dots, a_{j+L}$ be an addition block of this chain. For $j+1 \leq s \leq j+L$, write

$$a_s = a_{t(s)} + a_{r(s)}, \quad t(s) \geq r(s).$$

Assume that for all s with $j+1 \leq s \leq j+L$,

$$t(s), r(s) < j+1 \Rightarrow t(s), r(s) \in \mathcal{A},$$

then the addition block $a_{j+1}, \dots, a_{j+L}$ is **marked**.

Proof. Let $1 = a_0 < a_1 < \dots < a_{\lfloor m+r \rfloor}$ be an addition chain. Let $a_{j+1}, \dots, a_{j+L}$ be an addition block of this chain. Assume that none of the steps $a_j, \dots, a_{j+L-1}$ are used in subsequent addition blocks. Then we can show that this addition chain is not valid by comparing it to the addition chain $b_0, \dots, b_{\lfloor m+r \rfloor}$ defined by

$$b_s = a_s, \quad s < j, \quad b_s = a_{s+1}/2, \quad j \leq s \leq j+L-1, \quad b_s = a_s, \quad s \geq j+L. \quad \square$$

For any addition chain, we divide the addition blocks into two types.

Definition 4.7. We say that an addition block is of **type 1** if either

- $a_0 = 1$ is used in an addition of the block.
- An element of a previous addition block of type 1 is used in an addition of the block.

We say that an addition block is of **type 2** otherwise.

Remark 4.1. Observe that it follows directly from this definition and from Lemmas 4.3 and 4.14 that the addition blocks of type 2 are marked.

We are now ready to complete the proof of Theorem 3.1.

4.8 Completing the proof of the upper bound

Let K_1 and K_2 be the number of blocks of type 1 and type 2, respectively. Let D_1 and D_2 stand for the total number of steps in $\mathcal{D}$ in blocks of type 1 and 2, respectively. Finally, let T_1 and T_2 be the number of elements of blocks of type 2 that are used in subsequent blocks of type 1 and type 2, respectively. Since blocks of type 2 are marked, we have

$$T_1 + T_2 \geq K_2.$$

The number of ways to choose the elements added in steps of blocks of type 1 will be less than or equal to

$$\exp(D_1 \log m - (K_1 + T_1) \log \log m(1 + o(1))), \quad m \rightarrow \infty.$$

The number of ways to choose the elements added in steps of blocks of type 2 will be at most

$$\exp(D_2 \log m - T_2 \log \log m(1 + o(1))), \quad m \rightarrow \infty.$$

We get that the total number of ways to choose which integers are added in addition steps will be smaller than or equal to

$$\exp(r \log m - K \log \log m(1 + o(1))), \quad m \rightarrow \infty,$$

where K is the total number of addition blocks in the addition chain. Multiplying this by the number of ways of choosing the elements in each set and using Lemma 4.5, we get

$$F(m, r) \leq \exp\left(r \log m + o\left(\frac{m \log \log m}{\log m}\right)\right), \quad m \rightarrow \infty.$$

This completes the proof of Theorem 3.1.

4.9 Proof of the lower bound

For the lower bound, we will provide a constructive proof. In this proof, we will make use of binary expansion.

Explicitly, the constructive process goes as follows.

- **Step 1.** Choose positive integers k and u . We will later describe how to specify these integers. For now, assume that $ku \leq m$ and that $2^u + k - u \leq r$.
- **Step 2.** Choose nonnegative integers $s_1, \dots, s_{k-1}$ such that

$$\sum_{j=1}^{k-1} s_j + ku = m.$$

- **Step 3.** Define $\mathcal{U}$ as the set of odd positive integers in the interval $[2^{u-1}, 2^u]$. Choose k integers $U_1, \dots, U_k$ in the set $\mathcal{U}$, possibly with repetitions.
- **Step 4.** Let the first steps of the chain be

$$a_0 = 1, a_1 = 2, a_2 = 3, a_3 = 4, \dots, a_{2^u-2} = 2^u - 1.$$

- **Step 5.** Construct the rest of the chain as follows (using binary expansion):

$$\begin{aligned}
a_{2^u-1} &= U_1, \\
&\vdots \\
a_{2^u-1+s_1} &= U_1 \underbrace{0 \dots 0}_{s_1 \text{ zeros}}, \\
&\vdots \\
a_{2^u-1+s_1+u} &= U_1 \underbrace{0 \dots 0}_{s_1 \text{ zeros}} \underbrace{0 \dots 0}_u, \\
a_{2^u+s_1+u} &= U_1 \underbrace{0 \dots 0}_{s_1 \text{ zeros}} U_2, \\
&\vdots \\
N = a_t &= U_1 \underbrace{0 \dots 0}_{s_1 \text{ zeros}} U_2 \underbrace{0 \dots 0}_{s_2 \text{ zeros}} U_3 \dots U_{k-1} \underbrace{0 \dots 0}_{s_{k-1} \text{ zeros}} U_k,
\end{aligned}$$

with

$$t = 2^u + \sum_{j=1}^{k-1} s_j + (k-1)u + k - 2.$$

We can make the following remarks about the chains generated in such a manner.

- **Remark 1.** The number of binary digits of N is

$$\sum_{j=1}^{k-1} s_j + ku.$$

- **Remark 2.** In light of Step 2, the length of such an addition chain is

$$2^u + \sum_{j=1}^{k-1} s_j + (k-1)u + k - 2 = 2^u + m - u + k - 2 < m + r.$$

Here, the inequality comes from our assumption that $2^u + k - u \leq r$.

- **Remark 3.** Every distinct chain in this family of addition chains yields a different integer.
- **Remark 4.** The number of ways of choosing the integers $U_1, \dots, U_k$ is

$$\left(\frac{2^u}{4}\right)^k.$$

- **Remark 5.** The number of ways of choosing nonnegative integers $s_1, \dots, s_{k-1}$ such that

$$s_1 + \dots + s_{k-1} = m - ku$$

is equal to

$$\binom{m - ku + k - 2}{k - 1}.$$

From Remarks 1, 2, and 3, every chain constructed is valid and leads to a distinct integer. From Remarks 4 and 5, we can conclude

$$F(m, r) \geq \left(\frac{2^u}{4}\right)^k \binom{m - ku + k - 2}{k - 1} \geq \left(\frac{2^u(m - ku + k - 2)}{4(k - 1)}\right)^k. \quad (4.9)$$

To continue, we will assume $r = cm/\log m$. We will also assume $2^u + k = r$, that is, $u = \log(r - k)/\log 2$.

Observe that since $k \leq r$ and $u \leq \log r/\log 2$, we have for m large enough

$$ku \leq \frac{r \log r}{\log 2} \leq \frac{cm}{\log 2}. \quad (4.10)$$

Using (4.10) in (4.9), we get

$$F\left(m, \frac{cm}{\log m}\right) \geq \left(\frac{2^u m}{k}\right)^k \left(\frac{1}{4} \left(1 - \frac{c}{\log 2}\right)\right)^k.$$

Under the assumptions that $c > 0$ and $k = O(m/\log m)$, we deduce from this

$$F\left(m, \frac{cm}{\log m}\right) \geq 2^{uk} \left(\frac{m}{k}\right)^k \exp\left(O\left(\frac{m}{\log m}\right)\right).$$

Given that $k \leq r = cm/\log m$, we have

$$F\left(m, \frac{cm}{\log m}\right) \geq 2^{uk} \exp(k \log \log m) \exp\left(O\left(\frac{m}{\log m}\right)\right).$$

Using $u = \log(r - k)/\log 2$, it follows that

$$F\left(m, \frac{cm}{\log m}\right) \geq \exp\left(k \log(r - k) + k \log \log m + O\left(\frac{m}{\log m}\right)\right).$$

It remains to choose the value of k to make the right-hand side as large as possible.

We set $r - k = y$ and obtain

$$\begin{aligned} F\left(m, \frac{cm}{\log m}\right) &\geq \exp\left((r - y) \log y + (r - y) \log \log m + O\left(\frac{m}{\log m}\right)\right) \\ &= \exp\left((r - y) \log y + r \log \log m + O(y \log \log m) + O\left(\frac{m}{\log m}\right)\right). \end{aligned}$$

The value of y maximizing $(r - y) \log y$ satisfies

$$\frac{r}{y} = \log y + 1,$$

This is a transcendental equation and a good approximation of its solution is $y = r / \log r$. For the sake of simplicity, we use the following approximation of $r / \log r$:

$$y = \frac{cm}{(\log m)^2}.$$

We can conclude

$$\begin{aligned} &F\left(m, \frac{cm}{\log m}\right) \\ &\geq \exp\left(\left(\frac{cm}{\log m} - \frac{cm}{(\log m)^2}\right)(\log m - 2 \log \log m) + \frac{cm \log \log m}{\log m} + O\left(\frac{m}{\log m}\right)\right) \\ &= \exp\left(cm - \frac{cm \log \log m}{\log m} + O\left(\frac{m}{\log m}\right)\right). \end{aligned}$$

Observe that for any fixed $\varepsilon > 0$ and for m large enough,

$$\exp\left(cm - \frac{cm \log \log m}{\log m} + O\left(\frac{m}{\log m}\right)\right) > \exp\left(cm - \frac{cm \log \log m}{\log m} - \varepsilon \frac{m \log \log m}{\log m}\right).$$

This completes the proof of the lower bound in Theorem 3.1.

5 Conclusion

It would be nice to find an argument that would allow one to close the gap between our lower and upper bounds. However, the approach used in the proof of the lower bound (say with other values for k and u for instance) cannot yield a much better result. This is due to the limited number of choices we have in this approach for each additive step. Indeed, in each additive step, the number of ways of choosing the added integer is $2^u = cm / (\log m)^2$.

One could also consider the following approach to obtain a greater number of chains:

1. Construct an addition chain with r additive steps of the form $a_j = a_{j-1} + a_t$ with $t < j - 1$ and m doubling steps.

2. Choose the position of the additive steps. The number of ways to make this choice is

$$\binom{r+m}{r} = \exp\left((1+o(1))\frac{cm \log \log m}{\log m}\right).$$

3. At each addition step, add any previous odd integer. The number of ways to make this choice is $r!$.

The number of such chains is

$$r! \binom{r+m}{r} = \exp\left(cm + o\left(\frac{m \log \log m}{\log m}\right)\right).$$

The main difficulty with this approach is that some of these chains may lead to the same value. It is not clear how many of these chains lead to different integers. Answering this question would allow one to close the gap between the lower and upper bounds.

Acknowledgments

The research work of the first two authors was supported in part by their respective grants from The Natural Sciences and Engineering Research Council of Canada. The authors would also like to thank the referee for a careful read of the original manuscript.

References

- [1] H. Bahig, K. Nakamura, *Some properties of nonstar steps in addition chains and new cases where the Scholz conjecture is true*, Journal of Algorithms **42**(2) (2002), 304–316.
- [2] A. Brauer, *On addition chains*, Bull. Amer. Math. Soc. **45** (1939), 736–739.
- [3] N.M. Clift, *Calculating optimal addition chains*, Computing **91** (2011), 265–284.
- [4] P. Erdős, *Remarks on number theory III. On addition chains*, Acta Arith. **6** (1960), 77–81.
- [5] R.K. Guy, *Unsolved problems in number theory*, Third Edition, 2004, Springer-Verlag.
- [6] A. Scholz, *Aufgaben und Lösungen*, 253, Jahresbericht der Deutschen Mathematiker-Vereinigung **47** (1937), 41–42.
- [7] A. Schönhage, *A lower bound for the length of addition chains*, Theoret. Comput. Sci. **1** (1975), 1–12.

- [8] A. Tall, *The Scholz conjecture on addition chains is true for infinitely many integers with $\ell(2n) = \ell(n)$* , *Integers* **24A** (2024), A17, 1–12.
- [9] E.G. Thurber, *Efficient generation of minimal length addition chains*, *SIAM Journal on Computing* **28** (4) (1999), 1247–1263.
- [10] E.G. Thurber, *The Scholz–Brauer problem on addition chains*, *Pacific J. Math.* **49** (1), (1973), 229–242.

Département de mathématiques et de statistique, Université Laval, 1045 Avenue de la médecine, Québec, Québec G1V0A6, Canada
 Email address: jmdk@mat.ulaval.ca

Département de mathématiques et de statistique, Université Laval, 1045 Avenue de la médecine, Québec, Québec G1V0A6, Canada
 Email address: nicolas.doyon@mat.ulaval.ca

Department of mathematics, University of Toronto, 40 St George Street, Toronto, Ontario M5S2E4, Canada
 Email address: william.verreault@utoronto.ca