A HEURISTICS ON PRIMES IN AN ADDITION CHAIN

THEOPHILUS AGAMA

Abstract. Let

$$E(n) : 1 = s_0 < s_1 = 2 < \ldots < s_h = n$$

with $h \asymp n^{1-\epsilon}$ be an addition chain leading to n. We develop a heuristic for the least number of primes in an addition chain for all sufficiently large targets n.

1. Introduction

Let

$$E(n) : 1 = s_0 < s_1 = 2 < \ldots < s_h = n$$

be an addition chain leading to $n \geq 3$. Addition chains provide natural, combinatorial recipes for forming integers by repeated summation of previous terms [1]. Their length and internal structure arise in complexity theory and algorithmic number theory, but here we raise a different, purely arithmetic question: Under mild regularity, how many primes must an addition chain contain?

In this paper, we state and motivate Conjecture 2.1 and provide a compact and persuasive heuristic for it. Informally, the conjecture predicts that any addition chain whose upper half follows a reasonably regular linear profile contains at least

$$\frac{h}{2\log(\frac{h}{2})}(1 + o(1))$$

primes, as $n \to \infty$ (with h growing with n). To make this precise, we assume the following transparent regularity conditions on the tail of the addition chain:

- **Small step closure:**

$$|s_i - s_{i-1}| < s_{i-1} \quad \text{for all } 2 \leq i \leq h.$$

Date: September 30, 2025.

2010 *Mathematics Subject Classification.* Primary 11P32, 11A41; Secondary 11B13, 11H99.

Key words and phrases. primes, lower bounds.

- **Interpolation-closeness:**

$$|s_i - i\frac{n}{h}| < s_i \quad \text{for all } \lfloor \tfrac{h}{2} \rfloor + 1 \leq i \leq h.$$

- **Relative shift uniformity:** There exists $0 < \epsilon < 1$ such that for all $\lfloor \frac{h}{2} \rfloor + 1 \leq i \leq h$ we have

$$|s_i - i\frac{n}{h}| < \epsilon s_i.$$

These axioms single out chains whose upper half tracks, to first order, a linear interpolation from $n/2$ to n; they exclude highly irregular constructions (for instance, the standard Fibonacci truncation does not satisfy interpolation-closeness asymptotically). The heuristic argument is based on two complementary principles.

(1) *Equidistribution.* Under interpolation-closeness and relative-shift uniformity, the values

$$s_{\lfloor h/2 \rfloor + 1}, \ldots, s_h = n$$

are roughly equidistributed in the interval $(n/2, n]$. Concretely, if one partitions $(n/2, n]$ into k equal subintervals, there are about $(h/2)/k$ chain values in each subinterval.

(2) *Random-sampling.* Treating those $h/2$ tail values as if they were independent, uniformly chosen integers from $(n/2, n]$, the Prime Number Theorem predicts that the probability a typical such value s_i is prime is about $1/\log s_i$. Since $n/2 < s_i \leq n$, we have $\log s_i \sim \log n$, so the expected number of primes in the upper half is

$$\sum_{i=\lfloor h/2 \rfloor + 1}^{h} \frac{1}{\log s_i} \approx \frac{h}{2 \log n}.$$

In admissible growth regimes (for example $h \asymp n^{1-\epsilon}$) one may rewrite

$$\frac{h}{2\log(n)} = \frac{h}{2\log\big(\frac{h}{2}\cdot\frac{2n}{h}\big)} = \frac{h}{2\log(\frac{h}{2}) + 2\log(\frac{2n}{h})} \sim \frac{h}{2\log(\frac{h}{2})},$$

which yields the scale appearing in the conjecture.

We emphasize that this is a heuristic, not a proof. The main obstacles to a rigorous version are the arithmetic correlations enforced by the additive structure (congruence biases, systematic evenness, and other local divisibility phenomena) and the intrinsic fluctuations in prime distribution on short intervals. However, the two axioms above capture the precise hypotheses under which the random sampling picture becomes plausible, and they provide a natural framework for both theoretical and computational investigation.

2. A conjecture on primes in an addition chains

Although it has been shown that there cannot be *many* primes in an addition chain of *moderate* length, we have barely provided a criterion for counting primes if we allow our construction to include primes. This seems to be a very difficult problem, given the inherent irregular nature of the primes. At the moment, we make the following conjecture, which specifies the best way to include a few primes in an addition chain.

Conjecture 2.1 (Addition chain local prime distribution)**.** *Let $n \geq 3$ and let*

$$1 = s_o < 2 = s_1 < \cdots < s_h = n$$

be an addition chain leading to n with $h \asymp n^{1-\epsilon}$ and satisfying the regularity conditions

- ***Small step closure****:*

  $$|s_i - s_{i-1}| < s_{i-1}$$

  for all $2 \leq i \leq h$.

- ***Interpolation-closeness****:*

  $$|s_i - i\frac{n}{h}| < s_i$$

  for all $\lfloor \frac{h}{2} \rfloor + 1 \leq i \leq h$.

- ***Relative shift uniformity***: *There exists* $0 < \epsilon < 1$ *such that for all* $\lfloor \frac{h}{2} \rfloor + 1 \leq i \leq h$ *we have*
$$|s_i - i\frac{n}{h}| < \epsilon s_i.$$
Denote by $\mathcal{P}_h(n)$ *the number of primes among the chain* $\{s_o, s_1, \ldots, s_h\}$*, then as* $n \longrightarrow \infty$*, we have*
$$\mathcal{P}_h(n) \geq \frac{h}{2\log(\frac{h}{2})}(1 + o(1)).$$

3. Heuristics

Here we provide a heuristic to believe why Conjecture 2.1 could be a reasonable prediction. Let us assume that the chain values

$$s_{\lfloor \frac{h}{2} \rfloor + 1}, \ldots, s_h = n$$

are roughly equidistributed in the interval $(\frac{n}{2}, n]$: there are roughly $\frac{n}{2k}$ chain values in each subinterval in any partition of the interval $(\frac{n}{2}, n]$ into k subintervals of equal lengths. Again, let us pretend that the chain values

$$s_{\lfloor \frac{h}{2} \rfloor + 1}, \ldots, s_h = n$$

are independent random selections from the interval $(\frac{n}{2}, n]$, then the probability that a chain value s_i in this interval is prime is roughly $\frac{1}{\log s_i}$, according to the prime number theorem. Since $\frac{n}{2} < s_i \leq n$, we deduce that $\log n \sim \log s_i$ and get for the expected number of primes

$$\sum_{i=\lfloor \frac{h}{2} \rfloor + 1}^{h} \frac{1}{\log s_i} \approx \frac{h}{2\log n}.$$

Furthermore, if we assume that the chain length is of the order $h \asymp n^{1-\epsilon}$ for a sufficiently small $\epsilon > 0$, then we can write

$$\frac{h}{2\log n} = \frac{h}{2\log(\frac{h}{2}\frac{2n}{h})} = \frac{h}{2\log(\frac{h}{2}) + 2\log(\frac{2n}{h})} \sim \frac{h}{2\log(\frac{h}{2})}.$$

This is the quantity the conjecture predicts as a lower bound, since we have only examined the upper half of the chain. The three regularity conditions suggest treating the upper half as equidistributed and the chain values in this upper half as random selections from the interval $(\frac{n}{2}, n]$. It is important to state that this is far from a proof and should be read as a heuristic grounded on the two axioms of equidistribution and randomness.

References

1. A. Scholz, *253, Jber. Deutsch. Math* Verein. II, vol. 47, 1937, 41–42.