The Minimum LCM of a sequence

Daniel Mineev

September 9, 2024

1 Introduction

Here, I would like to discuss the problem of minimizing the LCM of a sequence of k **distinct** integers given their exact sum to be n. This value I will denote as $L_K(N)$. I will establish a lower and an upper bound on this value for any $n, k \in \mathbb{N}$. The lower bound, as you will see, is extremely simple to establish; a far more complex problem is establishing an upper bound.

Theorem 1. $L_k(n) \geq n / \left(\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k} \right)$

Consider the following problem, listed in the shortlist for IMO 2022, as an example of where this function has been used.

Question 1. *A number is called Norwegian if it has three distinct positive divisors whose sum is equal to 2022. Determine the smallest Norwegian number. (Note: The total number of positive divisors of a Norwegian number is allowed to be larger than 3.)*

This problem is equivalent to calculating the exact value of $L_3(2022)$. The official solution is quite bad; it is structured in a form that may be hard to develop. However, through this upper and lower bound, the example in the problem can be found way quicker.

2 Lower Bound

Here, I will prove that the lower bound on the value of $L_k(n)$ is $n / \left(\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k} \right)$, along with the fact that this bound is tight.

Theorem 2. $L_k(n) \geq n / \left(\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k} \right)$

Proof. Let us assume the opposite, that indeed, $L_k(n) < n / \left(\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k} \right)$, then let $L_k(n)/a_1 < L_k(n)/a_2 < \ldots < L_k(n)/a_k$ be the k divisors of $L_k(n)$ (meaning, the k divisors, which sum to n, from the definition), then we conclude,

that,

$$n = \sum_{i=1}^{k} \frac{L_k(n)}{a_i} = L_k(n) \left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_k} \right)$$

$$< \frac{n}{\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k}} \left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_k} \right) \quad (1)$$

However, as one can probably already see, $\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_k} \leq \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k}$, however even if this sum equals its maximum possible value the last expression will only be n, thus it is impossible for there to be a strict less sign in the inequality, contradiction. Thus, the inequality is indeed true. ∎

Notice that the lower bound on $L_k(n)$ gives something more interesting. Obviously one can see that from this, one can conclude that,

$$[a_1, a_2, \ldots, a_k] \geq \frac{1}{\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{k}} \sum_{i=1}^{k} a_i \quad (2)$$

is true for any $a_1, \ldots, a_k \in \mathbb{N}$, assuming that for no two a_i, a_j we have $a_i = a_j$. As a small example of the two equations, we get the following nice simple tricks, (given, that $a \neq b$)

$$[a, b] \geq \frac{2}{3}(a + b) \quad (3)$$

$$(a, b) = \frac{ab}{[a, b]} \leq \frac{2ab}{3(a + b)} \quad (4)$$

This is a direct consequence, of the lower bound of $L_2(N)$, which is $L_2(N) \leq \frac{2}{3}N$. While these results are, indeed, elementary, as I will show further in the document, they are extremely useful if one knows about them.

3 Upper Bound

3.1 $L_2(N)$ Analysis

Let us try to calculate the precise value of $L_2(N)$.

Theorem 3. *Given $n \in \mathbb{N}$, that $n \geq 3$, the value of $L_2(N)$ is $(q-1)r$, where q is the smallest divisor of N greater than 2 (this is achieved by the partitioning, $\frac{n}{q}, n - \frac{n}{q}$).*

Proof. Let us assume this is not the case, then it must hold, that, the partition of n was into two parts xd and yd, where $x, y > 1$, and d is the greatest common denominator of the two numbers in the partition (consequently $(x, y) = 1$). Notice, that $[xd, yd] = d[x, y]$, however, due to $(x, y) = 1$, one can conclude that $[x, y] = xy$, thus $[xd, yd] = xyd$. However, notice, that xy is minimized, when $x = 2$ and $y = \frac{n}{d} - 2$, but notice that even then,

2

$$dxy = 2d\left(\frac{n}{d} - 2\right) = 2(n - d) \qquad (5)$$

However, notice, that, if one considers the partition where $x = 1$ and $y = \frac{n}{d} - 1$, then the least common multiple will be the following,

$$dxy = d\left(\frac{n}{d} - 1\right) = n - d \qquad (6)$$

Which is better, than than $2(n-d)$, consequently, it is always benificial to make the partition of the form d, dx, thus, to minimize the least common denominator, the value of dx, must be minimized, in other words the value of $\frac{n}{d} - 1$ must be minimized, however this problem is equivelent to maximizing d (note, that the maximized value of d is represented as r in the statement). However, to simplify the statement, notice that maximizing d means minimizing $\frac{n}{d}$, however $\frac{n}{d} \geq 3$ must hold, for the partitioning to be legal (e.i. the numbers would be distinct), consequently, one sees, that if q is the smallest divisor, such that $q \geq 3$, the value of $L_2(N)$ is exactly $(q - 1)r$. ∎

Notice, that trivially (through simply considering the partitioning $1, n - 1$), $L_2(N)$ is at least bounded by $N - 1$, however through this theorem, we understand that this bound is tight for prime values of N, thus proving, that indeed, the upper bound $N - 1$ is tight.

While this does indeed already tell us a lot, the function is bounded by two linear functions, consequently $L_2(N) = O(N)$, however this does not tell the full story, the question that I will raise in this section, will be to find a function $\widehat{L}_2(N)$, such that,

$$\left[\sum_{k=1}^{N} L_2(k)\right]/N = \left[\sum_{k=1}^{N} \widehat{L}_2(k)\right]/N \qquad (7)$$

In simple terms, a nice approximation of $L_2(N)$ in the form of a linear function, $\widehat{L}_2(N)$. To start the analysis, I will perform a common trick, notice, that $L_2(N)$ from the previous theorem is a very hard function to work with, consequently to eliminate any drastic changes in the value of the function, I propose analysing the value of the following,

$$AL_2(N) = \left[\sum_{k=1}^{N} L_2(k)\right]/N \qquad (8)$$

Notice, that the problem of identifing the smallest divisor greater than 2, is not extremely different from identifying the smallest prime divisor, this is due to the fact, that the smallest divisor greater than 2, must be part of the set $\{4, 3, 5, 7, \ldots\}$ (in other words, the union of $\{4\}$ and the set of primes, greater than 2) let π_n represent the probability that the n-th special number is the minimal divisor greater than 2, then, indeed, a recursive relationship holds (where s_n is the n-th special number),

3

$$\pi_n = \frac{1}{s_n}\left(1 - \sum_{i=1}^{n-1}\pi_i\right) = \frac{\pi_{n-1}(s_{n-1}-1)}{s_n} \tag{9}$$

The last transition comes from,

$$1 - \sum_{i=1}^{n-2}\pi_i = \pi_{n-1}s_{n-1}$$

$$\implies 1 - \sum_{i=1}^{n-1}\pi_i = \pi_{n-1}s_{n-1} - \pi_{n-1} = \pi_{n-1}(s_{n-1}-1) \tag{10}$$

Now, all that is left, to reiterate the recursive relationship to obtain,

$$\pi_n = \frac{(s_2-1)\cdot(s_3-1)\dots(s_{n-1}-1)}{s_1\dots s_n} = \frac{1}{s_1 s_n}\prod_{i=2}^{n-1}\frac{s_i-1}{s_i} \tag{11}$$

First of all, quick note, from this, one can conclude, that, $\pi_n \le \frac{1}{s_1 s_n}$, which has as a direct biproduct,

$$\lim_{n\to\infty}\pi_n = 0 \tag{12}$$

Now, I shall return to the definition of $AL_2(N)$, to get, (where, $d(k)$ represents the smallest divisor, which is greater than 2 of k)

$$AL_2(N) = \left[\sum_{k=1}^{N}L_2(k)\right]/N = \left[\sum_{k=1}^{N}k - \frac{k}{d(k)}\right]/N$$

$$= \left[\frac{N(N+1)}{2}\right]/N - \sum_{k=1}^{N}\frac{k}{Nd(k)} = \frac{N+1}{2} - \mathbb{E}\left[\frac{k}{d(k)}\right] \tag{13}$$

The expectation is just a convienient way to rephrase the previous, assuming that the random variable k is uniformly selected from 1 to N. Thus, I propose analysing that value,

$$\mathbb{E}\left[\frac{k}{d(k)}\right] = \sum_{k=1}^{\infty}\frac{\pi_k N}{s_k} \le N\sum_{k=1}^{\infty}\frac{1}{s_1 s_k}\cdot\frac{1}{s_k} = \frac{N}{s_1}\sum_{k=1}^{\infty}\frac{1}{s_k^2} \tag{14}$$

One of the methods here, is to immidietely write the last expression using the prime zeta function,

$$\sum_{k=1}^{\infty}\frac{1}{s_k^2} = \frac{1}{4^2} + \sum_{k=2}^{\infty}\frac{1}{p_i^2} = \frac{1}{16} - \frac{1}{2^2} + P(2) \approx 0.13974 = C \tag{15}$$

Thus, in the end, we have, that,

$$AL_2(N) = \frac{N+1}{2} - \mathbb{E}\left[\frac{k}{d(k)}\right] \approx \frac{N+1}{2} - \frac{N}{3} \cdot C$$
$$= \frac{N+1}{2} - 0.04658N \approx 0.45342N \quad (16)$$

However, I have found this result to not be precise enough, this is due to the rough approximation we make, for the value of π_k, it far from accurate, I am referencing this specific moment in logic,

$$\frac{1}{s_1 s_n} \prod_{i=2}^{n-1} \frac{s_i - 1}{s_i} \leq \frac{1}{s_1 s_n} \quad (17)$$

If one accounts for this error (I have not figured out a way here, to get a nice formula to simplify the last, thus I have simply calculated this with a computer),

$$\sum_{k=1}^{\infty} \frac{\pi_k}{s_k} = \sum_{k=1}^{\infty} \left[\frac{1}{s_1 s_n} \prod_{i=2}^{n-1} \frac{s_i - 1}{s_i} \cdot \frac{1}{s_k} \right] \approx 0.18959 = C \quad (18)$$

From, one concludes, that,

$$AL_2(N) = \frac{N+1}{2} - \mathbb{E}\left[\frac{k}{d(k)}\right] \approx \frac{N+1}{2} - \frac{N}{3} \cdot C$$
$$\approx \left[\frac{1}{2} - 0.06319666\right] \cdot N \approx 0.43680334N \quad (19)$$

Now, notice, that if N is chosen at random from 1 to n then, (where $d(n)$ is the smallest divisor greater, than 2)

$$\mathbb{E}[L_2(N)] \sim \left(\sum_{n=1}^{N} \sum_{i=1}^{k} \pi_i \cdot (n - n/s_i) \right) / N \quad (20)$$

This is due, to the uniformness of the distrubution of π_i along the natural number line (meaning, the distribution of the natural numbers, which are used to calculate the value of π_i), without such uniformness the following statement would not hold. Now, notice, that,

$$\sum_{i=1}^{k} \pi_i(n - n/s_i) \leq \sum_{i=1}^{k} \frac{1}{s_1 s_i}(n - n/s_i) = \sum_{i=1}^{k} \frac{ns_i - n}{s_1 s_i^2} \quad (21)$$

Now, notice, that, by the Prime Number theorem, we get the following, $s_i \sim i \ln(i)$. Consequently,

$$\sum_{i=1}^{k} \frac{ns_i - n}{s_1 s_i^2} \sim n \sum_{i=1}^{k} \frac{i \ln(i) - 1}{i^2 \ln^2(i)} \sim n \sum_{i=1}^{k} \frac{1}{i \ln i} \sim n \ln(\ln(n)) \quad (22)$$

5

From, which we get, the following,

$$\left(\sum_{n=1}^{N}\sum_{i=1}^{k}\pi_i\cdot(n-n/s_i)\right)/N \sim \left(\sum_{n=1}^{N}n\ln(\ln(n))\right)/N \qquad (23)$$

Also, notice, that the last sum can, indeed, be approximated by an integral,

$$\sum_{n=1}^{N}n\ln(\ln(n)) \sim \int_{1}^{N}x\ln(\ln(x)) = \frac{x^2\ln(\ln(x))}{2} \qquad (24)$$

Thus,

$$AL_2(N) = O(N\ln(\ln(N))) \qquad (25)$$

From which we derive the final conclusion, that,

$$AL_2(N) = \left[\sum_{k=1}^{N}L_2(k)\right]/N = O(x\ln(\ln(x))) \qquad (26)$$

$$\sum_{k=1}^{N}L_2(k) = O(N^2\ln(\ln(x))) \qquad (27)$$

$$L_2(N) = O(N^2\ln(\ln(N)) - (N-1)^2\ln(\ln(N-1))) = O(N\ln(\ln(N))) \qquad (28)$$

Thus, we see, the final result on the upper bounding asymptotics of $L_2(N)$,

$$L_2(N) = O(N\ln(\ln(N))) \qquad (29)$$

Through numerical evidence, it seems that the true asymptotics is linear, $O(N)$, I believe the reason this method has not achieved this asymptotics, is due to the very rough upper bounding on the value of π_n, namely in (11) I have completely ignored the additional factor $\frac{s_2-1}{s_2}\cdot\ldots\cdot\frac{s_{n-1}-1}{s_{n-1}}$, which is, indeed, less, than 1 (specifically this allowed me to perform this analysis), however for a better asymptotic bound it would be wise to account for this additional factor.

3.2 $L_3(N)$ Case-wise Analysis

Consider the simplest and most elementary situation where $k = 3$. Now let one consider three situation for different possible congruences mod 3.

1. If $n \equiv 1 \pmod 3$, then an example of a 3-partition could be $1, k, 2k$ (with the sum exactly n) with a LCM of $2k$, where $n = 3k+1$, consequently we achieve an upper bound on the value for $L_3(n)$ to be $\frac{2(n-1)}{3}$.

2. If $n \equiv 2 \pmod 3$, analogously consider the example $2, k, 2k$, again where $n = 3k+2$. As one can see, indeed the LCM is again $2k$, thus again an upper bound is achieved of $\frac{2(n-2)}{3}$.
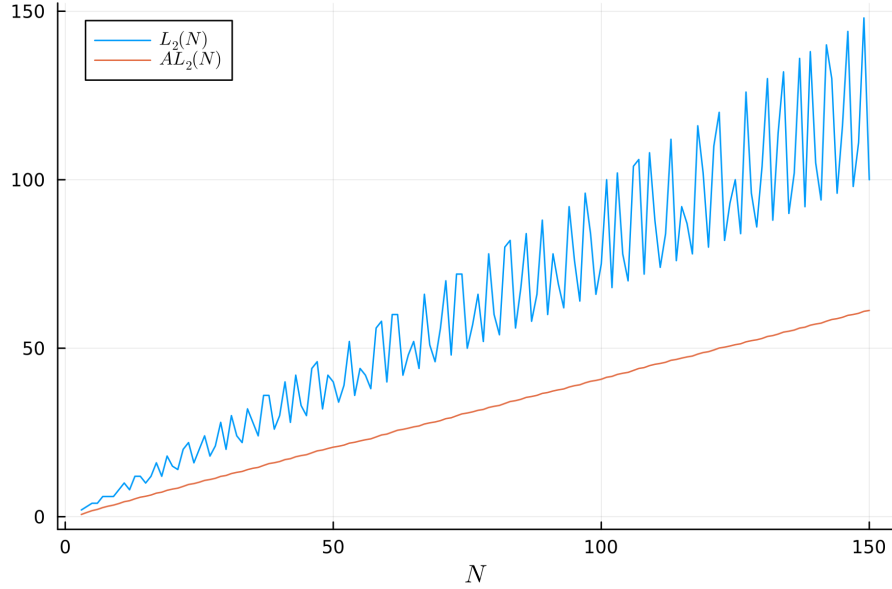
3. Finally, if $n \equiv 0 \pmod 3$, then $n = 3k$, let us again consider the possible congruences modulo 3 of k. (As you can see the work for now is quite technical)

  - If $k \equiv 0 \pmod 3$, then $n = 9t$ and an example of $t, 2t, 6t$ sets the upper bound of $\frac{2n}{3}$
  - If $k \equiv 1 \pmod 3$, then $3, k - 1, 2k - 2$ works, due to $2k - 2 \equiv 0 \pmod 3$
  - If $k \equiv 2 \pmod 3$, then $6, k - 2, 2k - 4$ works, due to $2k - 4 \equiv 0 \pmod 6$ if $k \equiv 2 \pmod 3$

Thus, indeed, one has established an upper bound on the value of $L_3(N)$ to be $\frac{2N}{3}$. It is important to note this bound is strict only for N large enough so that the above examples work (the exact number after which the bound starts working will be discussed a bit later). One can see the established lower and upper bound for the values of $L_3(N)$,

However, this somewhat brute-force approach does not work for larger values of N, in fact already just to find an upper bound for $L_4(N)$ posses a harder problem than before. Thus, calculating an upper bound for higher values of K requires a different perspective. In fact it seems the linear upper bound for $L_4(N)$ while close, is not as tight as it was for $L_3(N)$.
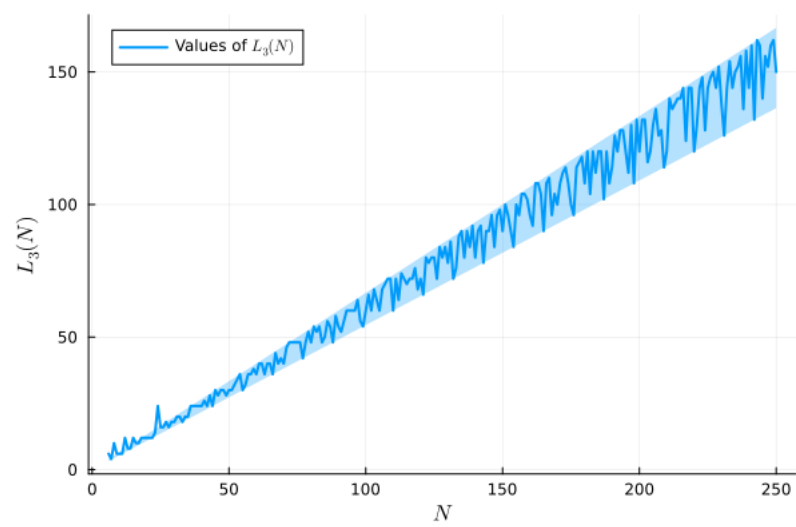
Figure 1