NOTES ON UPPER AND LOWER BOUNDING $t(N)$

TERENCE TAO

1. Basics

The symbol p will always denote a prime. The primes $2, 3$ will play a special role here and will be referred to as *tiny primes*.

We use $\nu_p(a/b) = \nu_p(a) - \nu_p(b)$ to denote the p-adic valuation of a positive natural number a/b, that is to say the number of times p divides the numerator a, minus the number of times p divides the denominator b. For instance, $\nu_2(32/27) = 5$ and $\nu_3(32/27) = -3$. If one applies a logarithm to the fundamental theorem of arithmetic, one obtains the identity

$$\sum_p \nu_p(r) \log p = \log r \tag{1.1}$$

for any positive rational r.

Asymptotically, it is known that

$$\frac{1}{e} - \frac{O(1)}{\log N} \leq \frac{t(N)}{N} \leq \frac{1}{e} - \frac{c_0 + o(1)}{\log N}$$

where

$$\begin{aligned} c_0 &:= \frac{1}{e} \int_0^1 \left\lfloor \frac{1}{x} \right\rfloor \log\left(ex \left\lceil \frac{1}{ex} \right\rceil \right) \, dx \\ &= \frac{1}{e} \int_1^\infty \lfloor y \rfloor \log \frac{\lceil y/e \rceil}{y/e} \frac{dy}{y^2} \\ &= 0.3044 \dots \end{aligned}$$

We recall Legendre's formula

$$\nu_p(N!) = \sum_{j=1}^\infty \left\lfloor \frac{N}{p^j} \right\rfloor = \frac{N - s_p(N)}{p-1}. \tag{1.2}$$

To bound the factorial, we have the explicit Stirling approximation [4]

$$N \log N - N + \log \sqrt{2\pi N} + \frac{1}{12N+1} \leq \log N! \leq N \log N - N + \log \sqrt{2\pi N} + \frac{1}{12N}, \tag{1.3}$$

valid for all natural numbers N.

We use $O_{\leq}(X)$ to denote any quantity whose magnitude is bounded by at most X (note the absence of an additional constant factor).

To estimate the prime counting function, we have the following good asymptotics up to a large height.

Theorem 1.1 (Buthe's bounds)**.** [1] *For any* $2 \leq x \leq 10^{19}$*, we have*

$$\mathrm{li}(x) - \frac{\sqrt{x}}{\log x}\left(1.95 + \frac{3.9}{\log x} + \frac{19.5}{\log^2 x}\right) \leq \pi(x) < \mathrm{li}(x)$$

and

$$\mathrm{li}(x) - \frac{\sqrt{x}}{\log x} \leq \pi^*(x) < \mathrm{li}(x) + \frac{\sqrt{x}}{\log x}.$$

For $x > 10^{19}$ we have the bounds of Dusart [2]. One such bound is

$$\psi(x) = x + O_{\leq}(59.18\frac{x}{\log^4 x}).$$

2. Criteria for upper bounding $t(N)$

We have the trivial upper bound $t(N) \leq (N!)^{1/N}$. This can be improved to $t(N) \leq N/e$ for $N \neq 1, 2, 4$, answering a conjecture of Guy and Selfridge [3]; see [5]. This was derived from the following slightly stronger criterion, which asymptotically gives $\frac{t(N)}{N} \leq \frac{1}{e} - \frac{c_0+o(1)}{\log N}$:

Lemma 2.1 (Upper bound criterion)**.** [5, Lemma 2.1] *Suppose that* $1 \leq t \leq N$ *are such that*

$$\sum_{p > \frac{t}{\lfloor\sqrt{t}\rfloor}} \left\lfloor \frac{N}{p} \right\rfloor \log\left(\frac{p}{t}\left\lceil \frac{t}{p} \right\rceil\right) > \log N! - N \log t \quad (2.1)$$

Then $t(N) < t$*.*

A surprisingly sharp upper bound comes from linear programming.

Lemma 2.2 (Linear programming bound)**.** *Let* N *be an natural number and* $1 \leq t \leq N/2$*. Suppose for each prime* $p \leq N$*, one has a non-negative real number* w_p *which is weakly non-decreasing in* p *(thus* $w_p \leq w_{p'}$ *when* $p \leq p'$*), and such that*

$$\sum_p w_p \nu_p(j) \geq 1 \quad (2.2)$$

for all $t \leq j \leq N$*, and such that*

$$\sum_p w_p \nu_p(N!) < N. \quad (2.3)$$

Then $t(N) < t$*.*

Proof. We first observe that the bound (2.2) in fact holds for all $j \geq t$, not just for $t \leq j \leq N$. Indeed, if this were not the case, consider the first $j \geq t$ where (2.2) fails. Take a prime p dividing j and replace it by a prime in the interval $[p/2, p)$ which exists by Bertrand's postulate (or remove p entirely, if $p = 2$); this creates a new j' in $[j/2, j)$ which is still at least t. By the weakly decerasing hypothesis on w_p, we have

$$\sum_p w_p \nu_p(j) \geq \sum_p w_p \nu_p(j')$$

and hence by the minimality of j we have

$$\sum_p w_p \nu_p(j) > 1,$$

a contradiction.

Now suppose for contradiction that $t(N) \geq t$, thus we have a factorization $N! = \prod_{j \geq t} j^{m_j}$ for some natural numbers m_j summing to N. Taking p-valuations, we conclude that

$$\sum_{j \geq t} m_j \nu_p(j) \leq \nu_p(N!)$$

for all $p \leq N$. Multiplying by w_p and summing, we conclude from (2.2) that

$$N = \sum_{j \geq t} m_j \leq \sum_p w_p \nu_p(N!),$$

contradicting (2.3). □

This bound is sharp for all $N \leq 600$, with the exception of $N = 155$, where it gives the upper bound $t(155) \leq 46$. A more precise integer program gives $t(155) = 45$.

3. A general factorization algorithm

In this section we present and then analyze an algorithm that, when given parameters $1 \leq t \leq N$, will attempt to construct a factorization $N! = \prod \mathcal{B}$ of $N!$ by a finite multiset $\mathcal{B}$ of N elements that are all at least t. The algorithm will not always succeed, but when it does, it will certify that $t(N) \geq t$.

3.1. **Notational preliminaries.** We begin with some key definitions.

Let $\mathcal{B} = \{b_1, \dots, b_M\}$ be a finite multiset of natural numbers (thus each natural number may appear in $\mathcal{B}$ multiple times); the ordering of elements in the multiset will not be of relevance to us. The *cardinality* $|\mathcal{B}| = M$ of the multiset is the number of elements counting multiplicity; for example,

$$|\{2, 2, 3\}| = 3.$$

The *product* $\prod \mathcal{B}$ of the finite multiset is defined by $\prod \mathcal{B} := \prod_{b \in \mathcal{B}} b$, where we count for multiplicity; for example

$$\prod \{2, 2, 3\} = 12.$$

The tuple $\mathcal{B}$ is a *factorization* of a natural number M if $\mathcal{B} = M$, and a *subfactorization* if $\mathcal{B}|M$. For example, $\{2, 2, 3\}$ is a factorization of 12 and a subfactorization of 24.

By the fundamental theorem of arithmetic (or (1.1)), we see that a finite multiset $\mathcal{B}$ is a factorization of M if and only if

$$v_p(M/\prod \mathcal{B}) = 0$$

for all primes p, and a subfactorization if and only if

$$v_p(M/\prod \mathcal{B}) \geq 0$$

for all primes p. We refer to $v_p(M/\prod \mathcal{B})$ as the *p-surplus* of $\mathcal{B}$ (as an attempted factorization) of M at prime p, and $-v_p(M/\prod \mathcal{B}) = v_p(\prod \mathcal{B}/M)$ as the *p-deficit*, and say that the factorization is *p-balanced* if $v_p(M/\prod \mathcal{B}) = 0$. Thus a subfactorization (resp. factorization) occurs when one has non-negative surpluses (resp. balance) at all primes p.

Example 3.1. Suppose one wishes to factorize $5! = 2^3 \times 3 \times 5$. The attempted factorization $\mathcal{B} := \{3, 4, 5, 5\}$ has a 2-surplus of $v_2(5!/\prod \mathcal{B}) = 1$, is in balance at 3, and has a 5-deficit of $v_2(\prod \mathcal{B}/5!) = 1$, so it is not a factorization or subfactorization of $5!$. However, if one replaces one of the copies of 5 in $\mathcal{B}$ with a 2, this erases both the 2-surplus and the 5-deficit, and creates a factorization $\{2, 3, 4, 5\}$ of $5!$.

A finite multiset $\mathcal{B}$ is said to be *t-admissible* for some $t > 0$ if $b \geq t$ for all $b \in \mathcal{B}$. Then $t(N)$ is largest quantity such that there exists a $t(N)$-admissible factorization of $N!$ of cardinality N.

Call a natural number 3-*smooth* if it is of the form $2^n 3^m$ for some natural numbers n, m. Given a positive real number x, we use $\lceil x \rceil^{\langle 2,3 \rangle}$ to denote the smallest 3-smooth number greater than or equal to x. For instance, $\lceil 5 \rceil^{\langle 2,3 \rangle} = 6$ and $\lceil 10 \rceil^{\langle 2,3 \rangle} = 12$.

3.2. **Description of algorithm.** We now describe an algorithm that, for given $1 \leq t \leq N$, either successfully demonstrates that $t(N) \geq t$, or halts with an error.

(0) Select a natural number A and another parameter $1 \leq K \leq \sqrt{t}$. There is some freedom to select parameters here, but generally speaking one would like to have $\log N \ll A \ll K \ll \sqrt{t}$.
(1) Let I denote the elements of the interval $[t, t(1 + 3/A)]$ that are coprime to 6. Let $\mathcal{B}^{(1)}$ be the elements of I, each occurring with multiplicity A. This multiset is t-admissible, and $\prod \mathcal{B}^{(1)}$ is not divisible by tiny primes 2, 3. (It will however have has approximately the right number of primes for $3 < p \leq t/K$, though it may have quite different prime factorization at primes $p > t/K$.)
(2) Remove any element from $\mathcal{B}^{(1)}$ that contains a prime factor p with $p > t/K$, and call this new multiset $\mathcal{B}^{(2)}$. It remains t-admissible with no tiny prime factors.
(3) For each $p > t/K$, add in $v_p(N!)$ copies of the number $p\lceil t/p \rceil$ to $\mathcal{B}^{(2)}$, and call this new multiset $\mathcal{B}^{(3)}$. (A variant of the method: add in $p\lceil t/p \rceil^{\langle 2,3 \rangle}$ instead. This is slightly less efficient, but slightly easier to analyze.) Now $\mathcal{B}^{(3)}$ is t-admissible and in balance at all primes $p > t/K$, but will typically be in a slight deficit at primes $3 < p \leq t/K$, particularly in the range $3 < p \leq K$. (It will now also contain a few tiny prime factors, but will generally still have a large surplus at those primes.)

(4) For each prime $3 < p \leq t/K$ at which there is a surplus $\nu_p(N!/\prod \mathcal{B}) > 0$, replace $\nu_p(N!/\prod \mathcal{B})$ copies of p in $\mathcal{B}^{(3)}$ with $\lceil p \rceil^{\langle 2,3\rangle}$ instead, and call this new multiset $\mathcal{B}^{(4)}$. Thus $\mathcal{B}^{(4)}$ has no surplus at primes $3 < p \leq t/K$ (and is still t-admissible and in balance for $p > t/K$).

(5) For the primes $3 < p \leq t/K$ at which there is a deficit $\nu_p(\prod \mathcal{B}/N!) > 0$, multiply all these primes together, and use the greedy algorithm to group them into factors $x_1, \dots, x_M$ in the range $(\sqrt{t/K}, t/K]$, together with possibly one exceptional factor x_* in the range $(1, t/K]$. For each of these factors x_i or x_*, add the quantity $x_i \lceil t/x_i \rceil^{\langle 2,3\rangle}$ or $x_* \lceil t/x_* \rceil^{\langle 2,3\rangle}$ to $\mathcal{B}^{(4)}$, and call this new multiset $\mathcal{B}^{(5)}$.

(6) By construction, $\mathcal{B}^{(5)}$ is t-admissible and will be in balance at all primes $p > 3$, and is thus $N!/\prod \mathcal{B}^{(5)}$ is of the form $2^n 3^m$ for some integers n, m. If at least one of n, m is negative, then `HALT` the algorithm with an error. Otherwise, select a 3-smooth number $2^{n_1} 3^{m_1}$ greater than equal to t with $n_1/m_1 \leq n/m$ (which one can interpret as $n_1 m \leq n m_1$ in case some of the denominators here vanish), and similarly select a 3-smooth number $2^{n_2} 3^{m_2}$ greater than or equal to t with $n_2/m_2 \geq n/m$. (It is reasonable to select the smallest such 3-smooth numbers in both cases, although this is not absolutely necessary for the algorithm to be successful.) By construction, we can express (n, m) as a positive linear combination $\alpha_1(n_1, m_1) + \alpha_2(n_2, m_2)$ of (n_1, m_1) and (n_2, m_2). Add $\lfloor \alpha_1 \rfloor$ copies of $2^{n_1} 3^{m_1}$ and $\lfloor \alpha_2 \rfloor$ copies of $2^{n_2} 3^{m_2}$ to $\mathcal{B}^{(5)}$, and call this tuple $\mathcal{B}^{(6)}$. (This will largely eliminate the surplus at 2 and 3.)

(7) If the multiset $\mathcal{B}^{(6)}$ has cardinality less than N, `HALT` the algorithm with an error. Otherwise, delete elements from $\mathcal{B}^{(6)}$ to bring the cardinality to N, and arbitrarily distribute any surplus primes to one of the remaining elements, and call the resulting multiset $\mathcal{B}^{(7)}$. By construction, $\mathcal{B}^{(7)}$ is a t-admissible factorization of $N!$ into N numbers, demonstrating that $t(N) \geq t$.

3.3. **Analysis of Step 7.** We now analyze the above algorithm, starting from the final step (7) and working backwards to (1), to establish sufficient conditions for the algorithm to successfully demonstrate that $t(N) \geq t$.

It will be convenient to introduce the following notation. For $a_+, a_- \in [0, +\infty]$, we define the asymmetric norm $|x|_{a_+,a_-}$ of a real number x by the formula

$$|x|_{a_+,a_-} := \begin{cases} a_+|x| & x \geq 0 \\ a_-|x| & x \leq 0. \end{cases}$$

If a_+, a_- are finite, this function is Lipschitz with constant $\max(a_+, a_-)$. One can think of a_+ as the "cost" of making x positive, and a_- as the "cost" of making x negative.

We now begin the analysis of Step 9. This procedure will terminate successfully as long as the length $|\mathcal{B}^{(6)}|$ of the tuple is at least N. To ensure this, we introduce the *t-excess* of a multiset $\mathcal{B}$ by the formula

$$\mathrm{E}_t(\mathcal{B}) := \prod_{b \in \mathcal{B}} \log \frac{b}{t} = \log \prod \mathcal{B} - |\mathcal{B}| \log t.$$

Thus, to ensure the success of this step, it suffices to establish the inequality

$$\mathrm{E}_t(\mathcal{B}^{(6)}) \leq \log \prod \mathcal{B}^{(6)} - N \log t.$$

From (1.1) we have

$$\log \prod \mathcal{B}^{(6)} = \log N! - \sum_p \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \log p,$$

so we can rewrite the previous condition (using the fact that $\mathcal{B}^{(6)}$ is a subfactorization of $N!$) as

$$\mathrm{E}_t(\mathcal{B}^{(6)}) + \sum_p \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \right|_{\log p, \infty} \leq \log N! - N \log t.$$

If we assume that $t = N/e^{1+\delta}$ for some $\delta > 0$, we can use the Stirling approximation (1.3) to reduce to the sufficient condition

$$\mathrm{E}_t(\mathcal{B}^{(6)}) + \sum_p \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(6)}} \right) \right|_{\log p, \infty} \leq \delta N + \log \sqrt{2\pi N}. \tag{3.1}$$

3.4. **Analysis of Step 6.** Now we analyze Step 6. For any $L \geq 1$, let κ_L be the least quantity such that

$$x \leq \lceil x \rceil^{\langle 2,3 \rangle} \leq \exp(\kappa_L) x \tag{3.2}$$

holds for all $x \geq L$. Just from considering the powers of two, we have the trivial upper bound

$$\kappa_L \leq \log 2. \tag{3.3}$$

We shall obtain better estimates on this quantity in Section ???. For now we use this quantity to help achieve efficient subfactorizations of 3-smooth numbers, as follows.

Lemma 3.2. *Let $L \geq 1$. Let $t > 3L$ and let $2^n 3^m$ be a 3-smooth number obeying the conditions*

$$\frac{\log(3L) + \kappa}{\log t - \log(3L)} \leq \frac{n \log 2}{m \log 3} \leq \frac{\log t - \log(2L)}{\log(2L) + \kappa}. \tag{3.4}$$

Then one can find a t-admissible subfactorization $\mathcal{B}$ of $2^n 3^m$ such that

$$\mathrm{E}_t(\mathcal{B}) \leq \kappa_L \frac{n \log 2 + m \log 3}{\log t} \tag{3.5}$$

and

$$|\nu_2(2^n 3^m/\mathcal{B})|_{\log 2, \infty} + |\nu_3(2^n 3^m/\mathcal{B})|_{\log 3, \infty} \leq 2(\log t + \kappa_L). \tag{3.6}$$

In practice, $\log t$ will be significantly larger than $\log(2L)$ or $\log(3L)$, and so the hypothesis (3.4) will be quite mild, as long as n and m are both reasonably large.

Proof. Let $2^{n_0}, 3^{m_0}$ be the largest powers of 2 and 3 less than or equal to t/L respectively, thus

$$L \leq \frac{t}{2^{n_0}} \leq 2L \tag{3.7}$$

and

$$L \leq \frac{t}{3^{m_0}} \leq 3L. \tag{3.8}$$

From (3.2), the 3-smooth numbers $\lceil t/2^{n_0} \rceil^{\langle 2,3 \rangle} = 2^{n_1} 3^{m_1}$, $\lceil t/3^{m_0} \rceil^{\langle 2,3 \rangle} = 2^{n_2} 3^{m_2}$ obey the estimates

$$\frac{t}{2^{n_0}} \leq 2^{n_1} 3^{m_1} \leq e^{\kappa} \frac{t}{2^{n_0}} \tag{3.9}$$

and

$$\frac{t}{3^{m_0}} \leq 2^{n_2}3^{m_2} \leq e^{\kappa}\frac{t}{3^{m_0}}, \tag{3.10}$$

or equivalently

$$t \leq 2^{n_0+n_1}3^{m_1}, 2^{n_2}3^{m_0+m_2} \leq e^{\kappa}t. \tag{3.11}$$

We can use (3.7), (3.9) to bound

$$\begin{aligned}\frac{n_0+n_1}{m_1} &\geq \frac{n_0}{\log(e^{\kappa}\frac{t}{2^{n_0}})/\log 3}\\ &\geq \frac{(\log t - \log(2L))/\log 2}{(\log(3L)+\kappa)/\log 3}\end{aligned}$$

(with the convention that this bound is vacuously true for $m_1 = 0$). Similarly, from (3.8), (3.10) we have

$$\begin{aligned}\frac{n_2}{m_0+m_2} &\leq \frac{\log(e^{\kappa}\frac{t}{3^{m_0}})/\log 2}{m_0}\\ &\leq \frac{(\log(2L)+\kappa)/\log 2}{(\log t - \log(3L))/\log 3}\end{aligned}$$

and hence by (3.4)

$$\frac{n_2}{m_0+m_2} \leq \frac{n}{m} \leq \frac{n_0+n_1}{m_1}. \tag{3.12}$$

Thus we can write (n, m) as a non-negative linear combination

$$(n, m) = \alpha_1(n_0+n_1, m_1) + \alpha_2(n_2, m_0+m_2)$$

for some real $\alpha_1, \alpha_2 \geq 0$. We now take our subfactorization $\mathcal{B}$ to consist of $\lfloor \alpha_1 \rfloor$ copies of the 3-smooth number $2^{n_0+n_1}3^{m_1}$ and $\lfloor \alpha_2 \rfloor$ copies of the 3-smooth number $2^{n_2}3^{m_0+m_2}$. By (3.11), each term $2^{n'}3^{m'}$ here is admissible and contributes an excess of at most κ, which is in turn bounded by $\frac{\kappa}{\log t}(n'\log 2 + m'\log 3)$. Adding these bounds together, we obtain (3.5).

The expression $2^n3^m/\prod \mathcal{B}$ contains at most $n_0+n_1+n_2$ factors of 2 and at most $m_0+m_2+m_1$ factors of 3, hence

$$\nu_2(2^n3^m/\prod \mathcal{B})\log 2 + \nu_3(2^n3^m/\prod \mathcal{B})\log 3 \leq \log 2^{n_0+n_1}3^{m_1} + \log 2^{n_2}3^{m_0+m_2},$$

and the bound (3.6) follows. □

We now use this lemma to analyze Step 6 as follows.

Proposition 3.3. *Let $L \geq 1$. Let $3L < t = N/e^{1+\delta}$ for some $\delta > 0$, and let $1 \leq K \leq t$ and $A \geq 1$. Suppose that the above algorithm with the indicated parameters reaches the end of Step 5 with a multiset $\mathcal{B}^{(5)}$ obeying the following hypotheses:*

(i) *(Small excess and surplus at non-tiny primes)*

$$\mathrm{E}_t(\mathcal{B}^{(5)}) + \sum_{p>3}\left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(5)}}\right)\right|_{\log p,\infty} \leq \delta N + \log\sqrt{2\pi} - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}. \tag{3.13}$$

(ii) *(Large surpluses at tiny primes) The surpluses $\nu_2(N!/\prod \mathcal{B}^{(5)}), \nu_3(N!/\prod \mathcal{B}^{(5)})$ are positive (so in particular Step 7 does not halt with an error) and obey the bounds*

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{\nu_2(N!/\prod \mathcal{B}^{(5)})\log 2}{\nu_3(N!/\prod \mathcal{B}^{(5)})\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. Write $n := \nu_2(N!/\prod \mathcal{B}^{(5)})$ and $m := \nu_3(N!/\prod \mathcal{B}^{(5)})$. From (1.2) we have $n \leq N$ and $m \leq N/2$, hence

$$n \log 2 + m \log 3 \leq N \log \sqrt{12}.$$

Applying Lemma 3.2, we can find a subfactorization $\mathcal{B}'$ of $2^n 3^m$ with an excess of at most $(\kappa_L \log \sqrt{12})\frac{N}{\log t}$, and with

$$|\nu_2(2^n 3^m/\prod \mathcal{B}')|_{\log 2,\infty} + |\nu_3(2^n 3^m/\prod \mathcal{B}')|_{\log 3,\infty} \leq 2(\log t + \kappa_L) \leq 2\log N$$

where we have used (3.3) and the fact that $\log t \leq \log N - 1$. Then $\mathcal{B}^{(6)} = \mathcal{B}^{(5)} \cup \mathcal{B}'$ is another t-admissible multiset, and from (3.13) and the observation that $-2 + 3\kappa_L \leq \log \sqrt{2\pi}$, we obtain the previously obtained sufficient condition (3.1). □

3.5. **Analysis of Step 5.**

Proposition 3.4. *Let $L \geq 1$. Let $9L < t = N/e^{1+\delta}$ for some $\delta > 0$, and let $1 \leq K \leq t$ and $A \geq 1$. Suppose that the above algorithm with the indicated parameters reaches the end of Step 4 to produce a multiset $\mathcal{B}^{(4)}$ obeying the following hypotheses.*

(i) *(Small excess and surplus at non-tiny primes)*

$$\begin{aligned} \mathrm{E}_t(\mathcal{B}^{(4)}) &+ \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\kappa_K \min(\frac{\log p}{\log \sqrt{t/K}},1),\infty} \\ &\leq \delta N - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}. \end{aligned} \tag{3.14}$$

(ii) *(Large surpluses at tiny primes) Whenever n_{**}, m_{**} are natural numbers obeying the bounds*

$$n_{**}\log 2 + m_{**}\log 3 \leq \sum_{3<p\leq t/K} \left|\nu_p\left(\frac{N!}{\prod \mathcal{B}^{(4)}}\right)\right|_{\frac{\log\sqrt{tK}+\kappa_K}{\log\sqrt{t/K}}\log p,\infty} + \log t + \kappa_L,$$

then $\nu_2(N!/\prod \mathcal{B}^{(4)}) > n_{**}$, $\nu_3(N!/\prod \mathcal{B}^{(4)}) > m_{**}$, and furthermore

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{(\nu_2(N!/\prod \mathcal{B}^{(4)}) - n_{**})\log 2}{(\nu_3(N!/\prod \mathcal{B}^{(4)}) - m_{**})\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. By (3.14), $\mathcal{B}^{(4)}$ is a subfactorization of $N!$. Consider all the p-surplus primes in the range $3 < p \leq t/K$, thus each such prime is considered with multiplicity $\nu_p(N!/\prod \mathcal{B})$. Using the greedy algorithm, one can factor the product of all these primes into M factors $c_1, \dots, c_M$ in the interval $[\sqrt{t/K}, t/K]$, times one exceptional factor c_* in $[1, \sqrt{t/K}]$, for some M. If we let M' denote the number of factors in $c_1, \dots, c_M$ that are not divisible by a prime larger than $\sqrt{t/K}$, We have the bound

$$(\sqrt{t/K})^{M'} \leq \prod_{3<p\leq\sqrt{t}} \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(4)}} \right)$$

and hence

$$M' \leq \sum_{3<p\leq\sqrt{t/K}} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(4)}} \right) \right|_{\frac{\log p}{\log \sqrt{t/K}}, \infty}.$$

Restoring the factors divisible by primes $p \geq \sqrt{t/K}$, we conclude that

$$M \leq \sum_{3<p\leq t/K} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(4)}} \right) \right|_{\min(\frac{\log p}{\log \sqrt{t/K}}, 1), \infty}. \tag{3.15}$$

For each of the M factors c_i, we introduce the 3-smooth number $\lceil t/c_i \rceil^{\langle 2,3 \rangle} = 2^{n_i} 3^{m_i}$, which by (3.2) lies in the interval $[t/c_i, e^{\kappa_K} t/c_i]$; similarly, for the exceptional factor c_* we introduce a 3-smooth number $\lceil t/c_* \rceil^{\langle 2,3 \rangle} = 2^{n_*} 3^{m_*}$ in the interval $[t/c_*, e^{\kappa_K} t/c_*]$. If we then adjoin the 3-smooth numbers $\lceil t/c_i \rceil^{\langle 2,3 \rangle} c_i = 2^{n_i} 3^{m_i} c_i$ for $i = 1, \dots, M$ as well as $\lceil t/c_* \rceil^{\langle 2,3 \rangle} c_* = 2^{n_*} 3^{m_*} c_*$ to the tuple $\mathcal{B}^{(4)}$ to create a new tuple $\mathcal{B}^{(5)}$. The quantity $\log \lceil t/c_* \rceil^{\langle 2,3 \rangle} = n_i \log 2 + m_i \log 3$ is bounded by $\log \sqrt{tK} + \kappa_K$, and the quantity $\log \lceil t/c_* \rceil^{\langle 2,3 \rangle} = n_* \log 2 + m_* \log 3$ is similarly bounded by $\log t + \kappa$, hence if we denote $n_{**} := n_1 + \dots + n_M + n_*$ and $m_{**} := m_1 + \dots + m_M + m_*$, we have

$$n_{**} \log 2 + m_{**} \log 3 \leq \frac{\log \sqrt{tK} + \kappa_K}{\log \sqrt{t/K}} \sum_{3<p\leq t/K} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(4)}} \right) \right|_{\log p, \infty} + \log t + \kappa_K.$$

Each of the new factors in $\mathcal{B}^{(5)}$ contributes an excess of at most κ_K, so the total excess of $\mathcal{B}^{(5)}$ is at most

$$\mathrm{E}_t(\mathcal{B}^{(4)}) + \kappa_K M + \kappa_K$$

which by (3.15) is bounded by

$$\mathrm{E}_t(\mathcal{B}^{(4)}) + \sum_{3<p\leq t/K} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}^{(4)}} \right) \right|_{\kappa_K \min(\frac{\log p}{\log \sqrt{t/K}}, 1), \infty} + \kappa_K.$$

We conclude that $\mathcal{B}^{(5)}$ obeys the hypotheses of Proposition 3.3 (using (3.3) to bound κ_K by $\log \sqrt{2\pi}$), and the claim follows. □

3.6. **Analysis of Step 4.**

Proposition 3.5. *Let $L \geq 1$. Let $9L < t = N/e^{1+\delta}$ for some $\delta > 0$, and suppose that the algorithm reaches the end of Step 3 to produce a multiset $\mathcal{B}^{(3)}$ obeying the following hypotheses:*

(i) *(Small excess and surplus at non-tiny primes) One has*

$$\mathrm{E}_t(\mathcal{B}^{(3)}) + \sum_{3<p\leq t/K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{\kappa_K \min(\frac{\log p}{\log\sqrt{t/K}},1),\kappa_p} \leq \delta N - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}. \quad (3.16)$$

(ii) *(Large surpluses at tiny primes) Whenever n_{**}, m_{**} are natural numbers obeying the bounds*

$$n_{**}\log 2 + m_{**}\log 3 \leq \sum_{3<p\leq t/K} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}^{(3)}}\right)\right|_{\frac{\log\sqrt{tK}+\kappa_K}{\log\sqrt{t/K}}\log p, \log p + \kappa_p} + \log t + \kappa_K,$$

then $\nu_2(N!/\prod \mathcal{B}^{(3)}) > n_{**}$, $\nu_3(N!/\prod \mathcal{B}^{(3)}) > m_{**}$, and furthermore

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{(\nu_2(N!/\prod \mathcal{B}^{(3)}) - n_{**})\log 2}{(\nu_3(N!/\prod \mathcal{B}^{(3)}) - m_{**})\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. Suppose there is a large prime p with a positive surplus $|\nu_p(N!/\prod \mathcal{B})|_{1,0} > 0$. Now we add the element $\lceil t/p\rceil^{\langle 2,3\rangle} p = 2^{n_{t/p}}3^{m_{t/p}}p$ to the multiset, which is at most $\exp(\kappa_{t/p})t$ by (3.2). This procedure reduces the p-deficit by one, adds at most $\kappa_{t/p}$ to the excess, and decrements $\nu_2(N!/\prod \mathcal{B})$, $\nu_3(N!/\prod \mathcal{B})$ by $n_{t/p}, m_{t/p}$ respectively. Since $n_{t/p}\log 2 + m_{t/p}\log 3 \leq \log(t/p) + \kappa_{t/p}$, if we apply this procedure to clear all surpluses at large primes, we have increased the excess by at most

$$\sum_{p>\sqrt{t/L}} \kappa_{t/p}$$

and decreased $\nu_2(N!/\prod \mathcal{B})$, $\nu_3(N!/\prod \mathcal{B})$ by some n', m' with

$$n'\log 2 + m'\log 3 \leq \sum_{p>\sqrt{t/L}} \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\log(t/p)+\kappa_{t/p},0}$$

.

The hypotheses of Proposition 3.4 are now satisfied, and we are done. □

3.7. **Analysis of Steps 1,2,3.** To apply **??**, we now compute the various statistics of $\mathcal{B}^{(3)}$ produced by Steps 1-3.

We begin with the analysis of $\mathcal{B}^{(1)}$, constructed in Step 2 of the algorithm. To count elements coprime to 6, we have the following lemma:

Lemma 3.6. *For any interval $[a, b]$ with $0 \leq a \leq b$, the number of natural numbers in the interval that are coprime to 6 is $\frac{b-a}{3} + O_{\leq}(4/3)$.*

Proof. By the triangle inequality, it suffices to show that the number of natural numbers coprime to 6 in $[0,a]$, minus $a/3$, is $O_{\leq}(2/3)$. The claim is easily verified for $0 \leq a \leq 6$, and the quantity in question is 6-periodic in a, giving the claim. □

The excess of $\mathcal{B}^{(1)}$ is clearly given by

$$\mathrm{E}_t(\mathcal{B}^{(1)}) = A \sum_{n \in I} \log \frac{n}{t}.$$

By the fundamental theorem of calculus, this is

$$A \int_0^{3t/A} |I \cap [t, t+h]| \frac{dh}{t+h}.$$

Bounding $\frac{1}{t+h}$ by $\frac{1}{t}$ and applying Lemma 3.6, we conclude that

$$\mathrm{E}_t(\mathcal{B}^{(1)}) \leq A \int_0^{3t/A} \left(h + \frac{4}{3}\right) \frac{dh}{t} = \frac{9t}{2A} + 4. \tag{3.17}$$

Next, we compute p-valuations $v_p(\mathcal{B}^{(1)})$. By construction, this quantity vanishes at tiny primes $p = 2, 3$. For $p > 3$, we can use Lemma 3.6 again to conclude

$$\begin{aligned} v_p(\mathcal{B}^{(1)}) &= A \sum_{1 \leq j \leq \frac{\log N}{\log p}} |I \cap p^j \mathbb{Z}| \\ &= A \sum_{1 \leq j \leq \frac{\log N}{\log p}} \left(\frac{t}{p^j A} + O_{\leq}(4/3) \right) \\ &= \frac{t}{p-1} + O_{\leq}\left(\frac{3t}{N(p-1)} \right) + O_{\leq}\left(\frac{4A}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right) \\ &= \frac{t}{p-1} + +O_{\leq}\left(\frac{4A+1}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right) \end{aligned}$$

since $\frac{3t}{N(p-1)} \leq \frac{3}{4e} \leq \frac{1}{3}$. Meanwhile, from (1.2) one has

$$v_p(N!) = \frac{t}{p-1} + O_{\leq}\left(\left\lceil \frac{\log N}{\log p} \right\rceil \right)$$

and thus

$$v_p(N!/\mathcal{B}^{(1)}) = O_{\leq}\left(\frac{4A+4}{3} \left\lceil \frac{\log N}{\log p} \right\rceil \right). \tag{3.18}$$

Now we pass to $\mathcal{B}^{(2)}$ by performing Step 3 of the algorithm. Removing elements from a t-admissible multiset cannot increase the t-excess, so from (3.17) we have

$$\mathrm{E}_t(\mathcal{B}^{(2)}) \leq \frac{9t}{2A} + 4. \tag{3.19}$$

The elements removed are of the form pm with $m \leq K(1 + \frac{3}{A})$ coprime to 6, and p in the interval $[\frac{t}{m}, \frac{t}{m}(1 + \frac{3}{A})]$ and greater than t/K. We conclude that

$$v_p(\mathcal{B}^{(2)}) = v_p(\mathcal{B}^{(1)})$$

for $K(1+\frac{3}{A}) < p \leq t/K$. For $3 < p < K(1+\frac{3}{K})$ one has

$$\nu_p(\mathcal{B}^{(2)}) = \nu_p(\mathcal{B}^{(1)}) - A \sum_{m \leq K(1+\frac{3}{A})} \nu_p(m)...$$

4. Powers of 2 and 3

We now obtain good bounds on the quantity κ_L. Clearly κ_L is a non-increasing function of L with $\kappa_1 = \log 2$. The following lemma gives improved control on κ_L for large L:

Lemma 4.1. *If n_1, n_2, m_1, m_2 are natural numbers such that $n_1 + n_2, m_1 + m_2 \geq 1$ and*

$$1 \leq \frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}$$

then

$$\kappa_{\min(2^{n_1+n_2}, 3^{m_1+m_2})/6} \leq \log \max\left(\frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}\right).$$

Thus, for instance, setting $n_1 = 3$, $m_1 = 2$, $n_2 = 2$, $m_2 = 1$, we have

$$\kappa_{4.5} \leq \log \frac{2^2}{3} = 0.28768\dots,$$

setting $n_1 = 3$, $m_1 = 2$, $n_2 = 5$, $m_2 = 3$, we have

$$\kappa_{40.5} \leq \log \frac{2^5}{3^3} = 0.16989\dots$$

and setting $n_1 = 11$, $m_1 = 7$, $n_2 = 8$, $m_2 = 5$, we have

$$\kappa_{2^{18}/3} \leq \log \frac{3^7}{2^{11}} = 0.06566\dots$$

($2^{18}/3 = 87381.33\dots$).

Proof. If $\min(2^{n_1+n_2}, 3^{m_1+m_2})/6 \leq t \leq 2^{n_2-1}3^{m_1-1}$, then we have

$$t \leq 2^{n_2-1}3^{m_1-1} \leq \max\left(\frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}\right) t, \tag{4.1}$$

so we are done in this case. Now suppose that $t > 2^{n_2-1}3^{m_1-1}$. If we write $\lceil t \rceil^{\langle 2,3 \rangle} = 2^n 3^m$ be the smallest 3-smooth number that is at least t, then we must have $n \geq n_2$ or $m \geq m_1$ (or both). Thus at least one of $\frac{2^{n_1}}{3^{m_1}} 2^n 3^m$ and $\frac{3^{m_2}}{3^{n_2}} 2^n 3^m$ is an integer, and is thus at most t by construction. This gives (4.1), and the claim follows. □

Some efficient choices of parameters for this lemma are given in Table 1. For instance, $\kappa_{4.5} \leq 0.28768\dots$ and $\kappa_{40.5} \leq 0.16989\dots$.

Remark 4.2. It should be unsurprising that the continued fraction convergents $1/1, 2/1, 3/2, 8/5, 19/12, \dots$ to

$$\frac{\log 3}{\log 2} = 1.5849\dots = [1; 1, 1, 2, 2, 3, 1, \dots]$$

are often excellent choices for n_1/m_1 or n_2/m_2, although occasionally other approximants such as $11/7$ are also usable.

n_1	m_1	n_2	m_2	$\min(2^{n_1+n_2}, 3^{m_1+m_2})/6$	$\log\max(3^{m_1}/2^{n_1}, 2^{n_2}/3^{m_2})$
1	1	**1**	**0**	$1/2 = 0.5$	$\log 2 = 0.69314\dots$
1	**1**	2	1	$2^2/3 = 1.33\dots$	$\log(3/2) = 0.40546\dots$
3	2	**2**	**1**	$3^2/2 = 4.5$	$\log(2^2/3) = 0.28768\dots$
3	2	**5**	**3**	$3^4/2 = 40.5$	$\log(2^5/3^3) = 0.16989\dots$
3	**2**	8	5	$2^{10}/3 = 341.33\dots$	$\log(3^2/2^3) = 0.11778\dots$
11	**7**	8	5	$2^{18}/3 = 87381.33\dots$	$\log(3^7/2^{11}) = 0.06566\dots$

TABLE 1. Efficient parameter choices for Lemma 4.1. The parameters which attain the minimum or maximum are indicated in **boldface**.

Asymptotically, we have logarithmic-type decay:

Lemma 4.3 (Baker bound)**.** *We have*

$$\kappa_L \ll \log^{-c} L$$

for all $L \geq 2$ and some absolute constant $c > 0$.

Proof. From the classical theory of continued fractions, we can find rational approximants

$$\frac{p_{2j}}{q_{2j}} \leq \frac{\log 3}{\log 2} \leq \frac{p_{2j+1}}{q_{2j+1}} \tag{4.2}$$

to the irrational number $\log 3/\log 2$, where the convergents p_j/q_j obey the recursions

$$p_j = b_j p_{j-1} + p_{j-2}; \quad q_j = b_j q_{j-1} + q_{j-2}$$

with $p_{-1} = 1, q = -1 = 0, p_0 = b_0, q_0 = 1$, and

$$[b_0; b_1, b_2, \dots] = [1; 1, 1, 2, 2, 3, 1 \dots]$$

is the continued fraction expansion of $\frac{\log 3}{\log 2}$. Furthermore, $p_{2j+1}q_{2j} - p_{2j}q_{2j+1} = 1$, and hence

$$\frac{\log 3}{\log 2} - \frac{p_{2j}}{q_{2j}} = \frac{1}{q_{2j}q_{2j+1}}. \tag{4.3}$$

By Baker's theorem, $\frac{\log 3}{\log 2}$ is a Diophantine number, giving a bound of the form

$$q_{2j+1} \ll q_{2j}^{O(1)} \tag{4.4}$$

and a similar argument (using $p_{2j+2}q_{2j+1} - p_{2j+1}q_{2j+2} = -1$) gives

$$q_{2j+2} \ll q_{2j+1}^{O(1)}. \tag{4.5}$$

We can rewrite (4.2) as

$$1 \leq \frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}}$$

and routine Taylor expansion using (4.3) gives the upper bounds

$$\frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}} \leq \exp\left(O\left(\frac{1}{q_{2j}}\right)\right).$$

From Lemma 4.1 we obtain

$$\kappa_{\min(2^{p_{2j}+p_{2j+1}}, 3^{q_{2j}+q_{2j+1}})/6} \ll \frac{1}{q_{2j}}.$$

The claim then follows from (4.4), (4.5) after optimizing in j.

□

It seems reasonable to conjecture that c can be taken to be arbitrarily close to 1, but this is essentially equivalent to the open problem of determining that irrationality measure of $\log 3/\log 2$ is equal to 2.

5. Asymptotic evaluation of $t(N)$

In this section we establish the asymptotic

$$\frac{t}{N} = \frac{1}{e} - \frac{c_0}{\log N} + O(\log^{1-c} N)$$

for some absolute constant $c > 0$.

We begin with the upper bound. ...

Now we establish the lower bound. Let N be sufficiently large. We introduce parameters

$$A := \lfloor \log^2 N \rfloor$$

and

$$K := \log^3 N.$$

Let I denote the integers in the interval $[t, t+3t/A]$ that are coprime to 6, and let $\mathcal{B}$ be the tuple consisting of these integers, each appearing with multiplicity A. This tuple is t-admissible, and the t-excess can be estimated as

$$\mathrm{E}_t(\mathcal{B}) \leq |\mathcal{B}| \log(1+3/A) \ll A\frac{t}{A}\frac{1}{A} \ll \frac{N}{\log^2 N}$$

by choice of A. As none of the elements of $\mathcal{B}$ are divisible by tiny primes, we have a considerable surplus at those primes. Indeed, from (1.2) we have

$$\nu_p(N!/\prod \mathcal{B}) = \nu_p(N!) = \frac{N}{p-1} - O(\log N)$$

for the tiny primes $p = 2, 3$.

6. Guy–Selfridge conjecture for $N > 10^{19}$

7. Guy–Selfridge conjecture for medium values of N

References

[1] J. Büthe, *Estimating $\pi(x)$ and related functions under partial RH assumptions*, Math. Comp., 85(301), 2483–2498, Jan. 2016.

[2] P. Dusart, *Explicit estimates of some functions over primes*, Ramanujan J. **45** (2018) 227–251.

[3] R. K. Guy, J. L. Selfridge, *Factoring factorial n*, Amer. Math. Monthly **105** (1998) 766–767.
[4] H. Robbins, *A Remark on Stirling's Formula*, Amer. Math. Monthly **62** (1955) 26–29.
[5] T. Tao, *Decomposing factorials into bounded factors*, preprint, 2025. https://arxiv.org/abs/2503.20170