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

TERENCE TAO

1. Basics

The symbol p will always denote a prime.

We use $v_p(a/b) = v_p(a) - v_p(b)$ 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, $v_2(32/27) = 5$ and $v_3(32/27) = -3$. If one applies a logarithm to the fundamental theorem of arithmetic, one obtains the identity

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

for any positive rational r.

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 1.1. 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 $\nu_2(5!/\prod \mathcal{B}) = 1$, is in balance at 3, and has a 5-deficit of $\nu_2(\prod \mathcal{B}/5!) = 1$, 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 $b \in \mathcal{B}$. We define $t(N)$ denotes the largest quantity such that there exists a $t(N)$-admissible factorization of $N!$ of cardinality N. Clearly, $t(N)$ is also the largest quantity such that there exists a $t(N)$-admissible subfactorization of $N!$ of cardinality at least N, since when starting from such a subfactorization, we may delete elements and then distribute any surpluses at any primes arbitrarily to create a factorization of cardinality exactly N.

Example 1.2. The finite multiset $\{2, 2, 3, 3, 4, 4, 5, 7\}$ is a 2-admissible subfactorization of $8! = 2^7 \times 3^2 \times 5 \times 7$, having a 2-surplus of 1. If one deletes a copy of 2 to make the cardinality exactly 8, one now has a surplus of 2 at 2; one can distribute these two powers of 2 to the remaining element of 2 to obtain a factorization $\{3, 3, 4, 4, 5, 7, 8\}$ that is still 2-admissible, and is in fact now 3-admissible.

A useful measure of the efficiency of a t-admissible finite multiset $\mathcal{B}$ is the *t-excess*

$$\mathrm{E}_t(\mathcal{B}) := \sum_{i=1}^{N'} \log \frac{b_i}{t} = \log(\prod \mathcal{B}) - |\mathcal{B}| \log t.$$

Example 1.3. The 3-excess of $\{3, 3, 4, 4, 5, 7, 8\}$ is

$$\mathrm{E}_t(\{3, 3, 4, 4, 5, 7, 8\}) = 2 \log \frac{3}{3} + 2 \log \frac{4}{3} + \log \frac{5}{3} + \log \frac{7}{3} + \log \frac{8}{3} = 2.914 \dots .$$

The t-excess clearly non-negative when $\mathcal{B}$ is t-admissible. Combining this with (1.1), we obtain the basic *accounting identity*

$$\mathrm{E}_t(\mathcal{B}) + \sum_p \nu_p(N!/\prod \mathcal{B}) \log p = \log N! - |\mathcal{B}| \log t. \tag{1.2}$$

In particular, when one has a subfactorization, the gap between $\log N!$ and $|\mathcal{B}| \log t$ must be somehow distributed between the t-excess $\mathrm{E}_t(\mathcal{B})$ and the p-surpluses $\nu_p(N!/\prod \mathcal{B})$.

Example 1.4. The 3-admissible finite multiset $\{3, 3, 4, 4, 5, 7, 8\}$ is a factorization of 8! of cardinality 7, and the gap

$$\log 8! - 7 \log 3 = 2.914 \dots$$

is entirely absorbed by the 3-excess of the multiset. If one replaces the element 8 of this multiset with 4, this reduces the excess to

$$\mathrm{E}_t(\{3, 3, 4, 4, 5, 7, 8\}) = 2 \log \frac{3}{3} + 3 \log \frac{4}{3} + \log \frac{5}{3} + \log \frac{7}{3} = 2.221 \dots ,$$

but creates a 2-surplus of 1 that contributes $\log 2 = 0.693 \dots$ to (1.2), restoring the accounting identity to balance.

From (1.2), we have the following equivalent definition of $t(N)$:

Lemma 1.5 (Equivalent form of $t(N)$)**.** *$t(N)$ is the supremum of all t for which there exists a t-admissible subfactorization $\mathcal{B}$ of $N!$ with*

$$\mathrm{E}_t(\mathcal{B}) + \sum_p \nu_p(N!/\mathcal{B}) \log p \leq \log N! - N \log t.$$

The advantage of this formulation is that one no longer needs to directly track the cardinality $|\mathcal{B}|$ of the t-admissible subfactorization $\mathcal{B}$. The formulation highlights the need to locate subfactorizations in which the t-excess and the p-surpluses are both kept as low as possible.

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

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

We recall Legendre's formula

$$\nu_p(N!) = \sum_{j=1}^\infty \left\lfloor \frac{N}{p^j} \right\rfloor = \frac{N - s_p(N)}{p-1}. \tag{1.3}$$

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

$$N \log N - N + \log \sqrt{2\pi N} + \frac{1}{12N+1} \leq \log N! \leq N \log N - N + \log \sqrt{2\pi N} + \frac{1}{12N}, \tag{1.4}$$

valid for all natural numbers N.

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

Theorem 1.6 (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

$$|\psi(x) - x| \leq 59.18 \frac{x}{\log^4 x}.$$

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*

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

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*

$$\sum_p w_p \nu_p(j) \geq 1 \tag{2.2}$$

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

$$\sum_p w_p \nu_p(N!) < N. \tag{2.3}$$

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). $\square$

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. Powers of 2 and 3

We now begin the study of constructions that can establish lower bounds of the form $t(N) \geq t$ for some

$$1 \leq t \leq N. \tag{3.1}$$

It will be convenient to parameterize

$$t = \frac{N}{e^{1+\delta}} \tag{3.2}$$

where we shall assume

$$\delta > 0; \tag{3.3}$$

for instance, if $t = N/3$, we will have $\delta = \log \frac{3}{3} \approx 0.098$. We also a parameter $L \geq 1$ for which

$$9L \leq t \tag{3.4}$$

and divide the primes into three categories:

- The *tiny primes* $p = 2, 3$;
- The *small primes* $3 < p \leq \sqrt{t/L}$;
- The *large primes* $p > \sqrt{t/L}$.

For any $B \geq 1$, define a *B-smooth number* to be a number whose prime factors are all at most B. Here we will be primarily interested in the cases $B = 2, 3$.

For any $x \geq 1$, let $\lceil x \rceil^{\langle 2 \rangle}$ denote the least 2-smooth number which is greater than equal to x. Since the 2-smooth numbers are just the powers of two, we have the explicit formula

$$\lceil x \rceil^{\langle 2 \rangle} = 2^{\lceil \log x / \log 2 \rceil}$$

as well as the bounds

$$x \leq \lceil x \rceil^{\langle 2 \rangle} < 2x.$$

Similarly, let $\lceil x \rceil^{\langle 2,3 \rangle}$ denote the least 3-smooth number which is greater than equal to x. We clearly inherit the previous bound,

$$x \leq \lceil x \rceil^{\langle 2,3 \rangle} < 2x,$$

but now expect to do better. To quantify this, define κ_L for each $L > 1$ to be the least quantity such that

$$x \leq \lceil x \rceil^{\langle 2,3 \rangle} \leq \exp(\kappa_L) x \tag{3.5}$$

for all $x \geq L$. Then 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 3.1. *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

$$t \leq 2^{n_2-1}3^{m_1-1} \leq \max\left(\frac{3^{m_1}}{2^{n_1}}, \frac{2^{n_2}}{3^{m_2}}\right) t, \tag{3.6}$$

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 (3.6), and the claim follows. $\square$

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

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 3.1. The parameters which attain the minimum or maximum are indicated in **boldface**.

Remark 3.2. It should be unsurprising that the continued fraction convergents $1/1, 2/1, 3/2, 8/5, 19/12, \dots$ to
$$\frac{\log 3}{\log 2} = 1.5849\cdots = [1; 1, 1, 2, 2, 3, 1, \dots]$$
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.

Asymptotically, we have logarithmic-type decay:

Lemma 3.3 (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
$$\frac{p_{2j}}{q_{2j}} \leq \frac{\log 3}{\log 2} \leq \frac{p_{2j+1}}{q_{2j+1}} \tag{3.7}$$
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
$$\frac{\log 3}{\log 2} - \frac{p_{2j}}{q_{2j}} = \frac{1}{q_{2j}q_{2j+1}}. \tag{3.8}$$
By Baker's theorem, $\frac{\log 3}{\log 2}$ is a Diophantine number, giving a bound of the form
$$q_{2j+1} \ll q_{2j}^{O(1)} \tag{3.9}$$
and a similar argument (using $p_{2j+2}q_{2j+1} - p_{2j+1}q_{2j+2} = -1$) gives
$$q_{2j+2} \ll q_{2j+1}^{O(1)}. \tag{3.10}$$
We can rewrite (3.7) as
$$1 \leq \frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}}$$
and routine Taylor expansion using (3.8) 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 3.1 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 (3.9), (3.10) 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.

We can now obtain efficient t-admissible subfactorizations of 3-smooth numbers $2^n 3^m$ when n, m are somewhat comparable.

Lemma 3.4. *Let $L \geq 1$. Let $t > 3L$ and let $2^n 3^m$ be a 3-smooth number obeying the conditions*

$$\frac{\log(3L) + \kappa}{\log t - \log(3L)} \leq \frac{n \log 2}{m \log 3} \leq \frac{\log t - \log(2L)}{\log(2L) + \kappa}. \tag{3.11}$$

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

$$\mathrm{E}_t(\mathcal{B}) \leq \kappa_L \frac{n \log 2 + m \log 3}{\log t} \tag{3.12}$$

and

$$|\nu_2(2^n 3^m/\mathcal{B})|_{\log 2,\infty} + |\nu_3(2^n 3^m/\mathcal{B})|_{\log 3,\infty} \leq 2(\log t + \kappa_L). \tag{3.13}$$

In practice, $\log t$ will be significantly larger than $\log(2L)$ or $\log(3L)$, and so the hypothesis (3.11) 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

$$L \leq \frac{t}{2^{n_0}} \leq 2L \tag{3.14}$$

and

$$L \leq \frac{t}{3^{m_0}} \leq 3L. \tag{3.15}$$

From (3.5), 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

$$\frac{t}{2^{n_0}} \leq 2^{n_1} 3^{m_1} \leq e^{\kappa} \frac{t}{2^{n_0}} \tag{3.16}$$

and

$$\frac{t}{3^{m_0}} \leq 2^{n_2} 3^{m_2} \leq e^{\kappa} \frac{t}{3^{m_0}}, \tag{3.17}$$

or equivalently

$$t \leq 2^{n_0+n_1} 3^{m_1}, 2^{n_2} 3^{m_0+m_2} \leq e^{\kappa} t. \tag{3.18}$$

We can use (3.14), (3.16) 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 (3.15), (3.17) we have

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

and hence by (3.11)

$$\frac{n_2}{m_0 + m_2} \leq \frac{n}{m} \leq \frac{n_0 + n_1}{m_1}. \tag{3.19}$$

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 (3.18), 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 (3.12).

The expression $2^n 3^m / \prod \mathcal{B}$ 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

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

and the bound (3.13) follows. □

4. Criteria for lower bounding $t(N)$

Lemma 1.5 gives an initial criterion for lower bounding $t(N)$. We now perform various manipulations on tuples to replace this criterion with a more tractable one. 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. One can then reformulate Lemma 1.5 as follows.

Proposition 4.1 (Reformulated criterion). *Let $1 \leq t \leq N$, and suppose that one has a t-admissible finite multiset $\mathcal{B}$ obeying the following hypothesis:*

(i) *(Small excess and surplus at all primes)*

$$\mathrm{E}_t(\mathcal{B}) + \sum_p \left| \nu_p\left(\frac{N!}{\prod \mathcal{B}}\right) \right|_{\log p, \infty} \leq \log N! - N \log t. \tag{4.1}$$

Then $t(N) \geq t$.

Indeed, the infinite penalty for making $v_p(N!/\mathcal{B})$ in (4.1) ensures that $\mathcal{B}$ is a subfactorization of $N!$.

We will reduce this infinite penalty term later, but let us work on other aspects of the criterion (4.1) first. In practice we will apply this criterion with $t := N/e^{1+\delta}$ for some $\delta > 0$; for instance, if we wish to set $t = N/3$, then $\delta = \log \frac{e}{3} \approx 0.098$. From (1.4) we may then replace $\log N! - N \log t = \log N! - N \log N + N + \delta N$ by the slightly smaller quantity

$$\delta N + \log \sqrt{2\pi N}.$$

The $\log \sqrt{2\pi N}$ is a lower order term, and we shall use it only to clean up some other lower order terms.

Using Lemma 3.4, we can leave a large surplus at tiny primes and still obtain a usable criterion:

Proposition 4.2 (Criterion with tiny-prime surplus)**.** *Let $L \geq 1$. Let $3L < t = N/e^{1+\delta}$ for some $\delta > 0$, and suppose that one has a t-admissible finite multiset $\mathcal{B}$ obeying the following hypotheses:*

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

$$\mathrm{E}_t(\mathcal{B}) + \sum_{p>3} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\log p,\infty} \leq \delta N + \kappa_L - \frac{3}{2}\log N - \kappa_L(\log \sqrt{12})\frac{N}{\log t}. \tag{4.2}$$

(ii) *(Large surpluses at tiny primes) The surpluses $v_2(N!/\prod \mathcal{B}), v_3(N!/\prod \mathcal{B})$ are positive and obey the bounds*

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{v_2(N!/\prod \mathcal{B})\log 2}{v_3(N!/\prod \mathcal{B})\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. Write $n := v_2(N!/\prod \mathcal{B})$ and $m := v_3(N!/\prod \mathcal{B})$. From (1.3) we have $n \leq N$ and $m \leq N/2$, hence

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

Applying Lemma 3.4, 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

$$|v_2(2^n3^m/\prod \mathcal{B}')|_{\log 2,\infty} + |v_3(2^n3^m/\prod \mathcal{B}')|_{\log 3,\infty} \leq 2(\log t + \kappa_L) \leq 2\log N - 2 + 2\kappa_L.$$

The union $\mathcal{B} \cup \mathcal{B}'$ of $\mathcal{B}$ and $\mathcal{B}'$ (counting multiplicity) is another t-admissible multiset, and from (4.2) and the observation that $-2 + 3\kappa_L \leq \log \sqrt{2\pi}$, we see that

$$\mathrm{E}_t(\mathcal{B} \cup \mathcal{B}') + \sum_p \left| v_p\left(\frac{N!}{\prod(\mathcal{B} \cup \mathcal{B}')}\right)\right|_{\log p,\infty} \leq \delta N + \log \sqrt{2\pi N},$$

and the claim now follows from Proposition 4.1. $\square$

The criterion (4.2) will still be somewhat expensive at small primes $3 < p \leq \sqrt{t/L}$. We can improve the situation as follows.

Proposition 4.3 (Improved criterion with tiny-prime surplus)**.** *Let $L \geq 1$. Let $9L < t = N/e^{1+\delta}$ for some $\delta > 0$, and suppose that one has a t-admissible finite multiset $\mathcal{B}$ obeying the following hypotheses.*

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

$$\begin{aligned} &\mathrm{E}_t(\mathcal{B}) + \sum_{3<p\leq\sqrt{t/L}} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\frac{\kappa_L \log p}{\log\sqrt{t/L}},\infty} + \sum_{p>\sqrt{t/L}} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\log p,\infty} \\ &\leq \delta N - \frac{3}{2}\log N - \kappa_L(\log\sqrt{12})\frac{N}{\log t}. \end{aligned} \tag{4.3}$$

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

$$n_{**}\log 2 + m_{**}\log 3 \leq \sum_{3<p\leq\sqrt{t/L}} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\frac{\log\sqrt{tL}+\kappa}{\log\sqrt{t/L}}\log p,\infty} + \log t + \kappa,$$

then $v_2(N!/\prod \mathcal{B}) > n_{**}$, $v_3(N!/\prod \mathcal{B}) > m_{**}$, and furthermore

$$\frac{\log(3L)+\kappa_L}{\log t - \log(3L)} \leq \frac{(v_2(N!/\prod \mathcal{B}) - n_{**})\log 2}{(v_3(N!/\prod \mathcal{B}) - m_{**})\log 3} \leq \frac{\log t - \log(2L)}{\log(2L)+\kappa_L}.$$

Then $t(N) \geq t$.

Proof. By (4.3), $\mathcal{B}$ is a subfactorization of $N!$. Consider all the p-surplus primes in the range $3 < p \leq \sqrt{t/L}$, thus each such prime is considered with multiplicity $v_p(N!/\prod \mathcal{B})$. Denoting their product as B, we have

$$\log B = \sum_{3<p\leq\sqrt{t/L}} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\log p,\infty}.$$

Using the greedy algorithm, one can factor B into M factors $c_1, \dots, c_M$ in the interval $[\sqrt{t/L}, \frac{t/L}{]}$, times one exceptional factor c_* in $[1, \sqrt{t/L}]$, for some M. We have the bound

$$(\sqrt{t/L})^M \leq B$$

and hence

$$M \leq \sum_{3<p\leq\sqrt{t/L}} \left| v_p\left(\frac{N!}{\prod \mathcal{B}}\right)\right|_{\frac{\log p}{\log\sqrt{t/L}},\infty}. \tag{4.4}$$

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 (3.5) lies in the interval $[t/c_i, e^{\kappa_L}t/c_i]$; 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 $[t/c_*, e^{\kappa_L}t/c_*]$. 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 $\mathcal{B}$ to create a new tuple $\mathcal{B}'$. This tuple is still t-admissible, but now has no p-surplus (or p-deficit) at any prime $3 < p \leq \sqrt{t/L}$. The quantity $\log\lceil t/c_*\rceil^{\langle 2,3\rangle} = n_i \log 2 + m_i \log 3$

is bounded by $\log \sqrt{tL} + \kappa_L$, 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 $n_{**} := n_1 + \cdots + n_M + n_*$ and $m_{**} := m_1 + \cdots + m_M + m_*$, we have

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

By hypothesis, we now see that $\mathcal{B}'$ has no p-deficit at 2 or 3 either, so $\mathcal{B}'$ is still a subfactorization of $N!$. Each of the new factors in $\mathcal{B}'$ contributes an excess of at most κ_L, so the total excess of $\mathcal{B}'$ is at most

$$\mathrm{E}_t(\mathcal{B}) + \kappa_L M + \kappa_L$$

which by (4.4) is bounded by

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

We conclude that $\mathcal{B}'$ obeys the hypotheses of Equation (4.2), and the claim follows. $\square$

Finally, we relax the subfactorization condition by permitting some p-deficit at various non-tiny primes $p > 3$.

Proposition 4.4 (Improved criterion with tiny-prime surplus, and some deficit)**.** *Let $L \geq 1$. Let $9L < t = N/e^{1+\delta}$ for some $\delta > 0$, and suppose that one has a t-admissible finite multiset $\mathcal{B}$ obeying the following hypotheses:*

(i) *(Small excess and surplus at small and large primes) One has*

$$\begin{aligned} \mathrm{E}_t(\mathcal{B}) + \sum_{3 < p \leq \sqrt{t/L}} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}} \right) \right|_{\frac{\kappa_L \log p}{\log \sqrt{t/L}}, \kappa_p} + \sum_{p > \sqrt{t/L}} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}} \right) \right|_{\kappa_{t/p}, \kappa_p} \\ \leq \delta N - \frac{3}{2} \log N - \kappa_L (\log \sqrt{12}) \frac{N}{\log t}. \end{aligned} \tag{4.5}$$

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

$$\begin{aligned} n_{**} \log 2 + m_{**} \log 3 \leq & \sum_{3 < p \leq \sqrt{t/L}} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}} \right) \right|_{\frac{\log \sqrt{tL} + \kappa_L}{\log \sqrt{t/L}} \log p, \log p + \kappa_p} \\ & + \sum_{p > \sqrt{t/L}} \left| \nu_p \left(\frac{N!}{\prod \mathcal{B}} \right) \right|_{\log(t/p) + \kappa_{t/p}, \log p + \kappa_p} + \log t + \kappa, \end{aligned}$$

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

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

Then $t(N) \geq t$.

Proof. Consider all the primes with a positive deficit, that is to say the primes p with $|\nu_p(N!/\prod \mathcal{B})|_{0,1} > 0$. As there is a surplus at tiny primes, such primes p must be small or large. If p is one of these primes, we select an element of the tuple that contains p as a factor, and replace it with $\lceil p \rceil^{\langle 2,3 \rangle} = 2^{n_p} 3^{n_p}$, thus increasing this element by a factor of at most $\exp(\kappa_p)$ thanks to (3.5); meanwhile, $\nu_2(N!/\prod \mathcal{B}) \log 2 + \nu_3(N!/\prod \mathcal{B}) \log 3$ goes down by at most $\log p + \kappa p$. Performing this for all the primes in deficit, we can clear this deficit at the cost of raising the excess of $\mathcal{B}$ by at most $\sum_{p>3} \kappa_p$, and decreasing $\nu_2(N!/\prod \mathcal{B})$, $\nu_3(N!/\prod \mathcal{B})$ by some n, m with $n \log 2 + m \log 3 \leq \sum_{p>3} |\nu_p(N!/\prod \mathcal{B})|_{0, \log p + \kappa_p}$.

Now 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 (3.5). 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 4.3 are now satisfied, and we are done. □

5. A GENERAL FACTORIZATION ALGORITHM

Given parameters $1 \leq t \leq N$, one can construct a t-admissible subfactorization $\mathcal{B}$ of $N!$ as follows.

(1) Select a natural number A and another parameter K. There is some freedom to select parameters here, but generally speaking one would like to have $\log N \ll A \ll K \ll \sqrt{N}$.
(2) Let I denote the elements of the interval $[t, t(1+3/A)]$ that are coprime to 6. Initialize $\mathcal{B}$ to be the elements of I, each occurring with multiplicity A. This tuple has no factors at tiny primes 2, 3, but has approximately the right number of primes for $3 < p \leq t/K$.
(3) Remove any element from $\mathcal{B}$ that contains a prime factor p with $p > t/K$.
(4) For each $p > t/K$, add in $\nu_p(N!)$ copies of the number $p\lceil t/p \rceil$. (A variant of the method: add in $t\lceil t/p \rceil^{\langle 2,3 \rangle}$ instead. This is slightly less efficient, but slightly easier to analyze.) Now $\mathcal{B}$ is 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 $3 < p \leq K$.
(5) 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 $\mathcal{B}$ with $\lceil p \rceil^{\langle 2,3 \rangle}$ instead, to eliminate this surplus.

(6) For the primes $3 < p \leq t/K$ at which there is a deficit $v_p(\prod \mathcal{B}/N!) > 0$, 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 $x_* \lceil t/x_* \rceil^{\langle 2,3\rangle}$. This will eliminate the deficit at these primes.
(7) After this process, the ratio $N!/\prod \mathcal{B}$ should now be a 3-smooth number $2^n 3^m$ with $n, m \geq 0$. If this is not the case (because one of the exponents has become negative), halt with an error. Otherwise, continue.
(8) Let $2^{n_1} 3^{m_1}$ be the least 3-smooth number 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 simliarly let $2^{n_2} 3^{m_2}$ be the least 3-smooth number greater than or equal to t with $n_2/m_2 \geq n/m$. 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 $\mathcal{B}$. This will largely eliminate the surplus at 2 and 3.

In practice, one expects this to be a rather efficient factorization that should extract close to the optimal number of factors of $N!$ greater than or equal to t.

6. 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

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

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 (1.3) we have

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

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

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

8. 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`