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

TERENCE TAO

1. BASICS

Let $\vec{b} = (b_1, \dots, b_{N'})$ be a tuple of natural numbers. The quantity N' will be called the *length* of a tuple and denoted $\ell(\vec{b})$. The *product* $\prod \vec{b}$ of the tuple is defined by $\prod \vec{b} := b_1 \dots b_{N'}$. The tuple $\vec{b}$ is a *factorization* of a natural number M if $\vec{b} = M$, and a *subfactorization* if $\vec{b}|M$.

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. By the fundamental theorem of arithmetic, we see that a tuple $\vec{b}$ is a factorization of M if and only if

$$\nu_p(M/\prod \vec{b}) = 0$$

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

$$\nu_p(M/\prod \vec{b}) \geq 0$$

for all primes p. We refer to $\nu_p(M/\vec{b})$ as the *p-surplus* of $\vec{b}$ (as an attempted factorization) of M at prime p, and $-\nu_p(M/\prod \vec{b}) = \nu_p(\prod \vec{b}/M)$ as the *p-deficit*. Thus a subfactorization (resp. factorization) occurs when all the p-surpluses are non-negative (resp. zero).

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.

A tuple $\vec{b} = (b_1, \dots, b_{N'})$ is said to be *t-admissible* for some $t > 0$ if $b_i \geq t$ for all $i = 1, \dots, N'$. We define $t(N)$ denotes the largest quantity such that there exists a $t(N)$-admissible factorization of $N!$ of length N. Clearly, $t(N)$ is also the largest quantity such that there exists a $t(N)$-admissible subfactorization of $N!$ of length at least N, since when starting from such a subfactorization, we may delete elements and then distribute any p-surpluses arbitrarily to create a factorization of length exactly N.

A good measure of the efficiency of a t-admissible factorization (or subfactorization) $\vec{b}$ is the *t-excess*

$$E_t(\vec{b}) := \sum_{i=1}^{N'} \log \frac{b_i}{t} = \log \prod \vec{b} - \ell(\vec{b}) \log t.$$

This is clearly non-negative when $\vec{b}$ is t-admissible. Combining this with (1.1), we obtain the basic *balance identity*

$$E_t(\vec{b}) + \sum_p \nu_p(N!/\prod \vec{b}) \log p = \log N! - \ell(b) \log t. \tag{1.2}$$

That is to say, the gap between $\log N!$ and $\ell(b)\log t$ must be somehow distributed between the t-excess $E_t(\vec{b})$ and the p-surpluses $\nu_p(N!/\prod \vec{b})$. In particular, we have the following equivalent definition of $t(N)$:

Lemma 1.1 (Equivalent form of $t(N)$)**.** *$t(N)$ is the supremum of all t for which there exists a t-admissible subfactorization $\vec{b}$ of $N!$ with*

$$E_t(\vec{b}) + \sum_p \nu_p(N!/\vec{b}) \log p \leq \log N! - N \log t.$$

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

We recall Legendre's formula

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

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

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

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 1.2 (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. Powers of 2 and 3

For any $t \geq 1$, there exists n such that

$$t \leq 2^n < 2t;$$

indeed one can take $n = \lfloor \log t / \log 2 \rfloor$. If one also admits powers of two, one can do better. For instance:

Lemma 2.1. *For any (real) $t \geq 40.5$, there exist natural numbers n, m such that*

$$t \leq 2^n 3^m \leq \frac{32}{27} t.$$

For comparison, we have $\log \frac{32}{27} = 0.16989\ldots$, representing about a four-fold improvement over $\log 2 = 0.69314\ldots$.

Proof. If $t \leq 48 = 2^4 \times 3 = \frac{32}{27} \times 40.5$ then we can take $n = 4, m = 1$, so assume $t > 2^4 \times 3$. Let $2^n 3^m$ be the smallest number of this form that is at least t, then we must have $n \geq 5$ or $m \geq 2$ (or both). Thus at least one of $\frac{3^3}{2^5} 2^n 3^m$ and $\frac{2^3}{3^2} 2^n 3^m$ is an integer, and is thus at most t by construction. Hence either $2^n 3^m \leq \frac{2^5}{3^3} t$ or $2^n 3^m \leq \frac{3^2}{2^3} t$. Since $\frac{3^2}{2^3} \leq \frac{2^5}{3^3} = \frac{32}{27}$, the claim follows. □

Asymptotically, we can do even better:

Lemma 2.2 (Baker bound)**.** *For $t \geq 2$, we can find natural numbers n, m such that*

$$t \leq 2^n 3^m \leq \exp(O(\log^{-c} t)) t$$

for some absolute constant $c > 0$.

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

$$\frac{p_{2j}}{q_{2j}} \leq \frac{\log 3}{\log 2} \leq \frac{p_{2j+1}}{q_{2j+1}} \tag{2.1}$$

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 $A_{-1} = 1, B_{-1} = 0, A_0 = b_0, B_0 = 1$, and $[b_0; b_1, b_2, \dots] = [1; 1, 1, 2, \dots]$ 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

$$\frac{\log 3}{\log 2} - \frac{p_{2j}}{q_{2j}} = \frac{1}{q_{2j}q_{2j+1}}. \tag{2.2}$$

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

$$q_{2j+1} \ll q_{2j}^{O(1)} \tag{2.3}$$

and a similar argument gives

$$q_{2j+2} \ll q_{2j+1}^{O(1)}. \tag{2.4}$$

We can rewrite (2.1) as

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

By repeating the proof of Lemma 2.1, we see that if

$$t > 2^{p_{2j+1}-1}3^{q_{2j}-1}, \tag{2.5}$$

then the first expression of the form 2^n3^m that is greater than or equal to t obeys the bound

$$t \leq 2^n3^m \leq t \max\left(\frac{3^{q_{2j}}}{2^{p_{2j}}}, \frac{2^{p_{2j+1}}}{3^{q_{2j+1}}}\right).$$

From (2.1), (2.2) one can bound

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

If one then sets j to be the largest natural number for which (2.5) holds, the claim then follows from (2.3), (2.4). □

We can now obtain efficient t-admissible subfactorizations of 2^n3^m when n, m are somewhat comparable.

Lemma 2.3. *Set $L := 40.5$ and $\kappa := \log \frac{32}{27}$, or else $L \geq 2$ and $\kappa = c \log^{-c} L$ for a sufficiently small constant $c > 0$. Let $t > 3L$ and n, m be positive integers obeying the conditions*

$$\frac{\log(3L) + \kappa}{\log t - \log(3L)} \leq \frac{n \log 2}{m \log 3} \leq \frac{\log t - \log(2L)}{\log(2L) + \kappa}. \tag{2.6}$$

Then one can find a t-admissible subfactorization $\vec{b}$ of 2^n3^m such that

$$E_t(\vec{b}) \leq \frac{\kappa}{\log t}(n \log 2 + m \log 3) \tag{2.7}$$

and

$$|\nu_2(2^n3^m/\vec{b})|_{\log 2,\infty} + |\nu_3(2^n3^m/\vec{b})|_{\log 3,\infty} \leq 2(\log t + \kappa). \tag{2.8}$$

Proof. Let $2^{n_0}, 3^{m_0}$ be the largest powers of 2 and 3 less than t/L respectively. By Lemma 2.1 or Lemma 2.2, we can find natural numbers n_1, m_1, n_2, m_2 such that

$$\frac{t}{2^{n_0}} \leq 2^{n_1}3^{m_1} \leq e^{\kappa}\frac{t}{2^{n_0}} \tag{2.9}$$

and

$$\frac{t}{3^{m_0}} \leq 2^{n_2}3^{m_2} \leq e^{\kappa}\frac{t}{3^{m_0}}, \tag{2.10}$$

or equivalently

$$t \leq 2^{n_0+n_1}3^{m_1}, 2^{n_2}3^{m_0+m_2} \leq e^{\kappa}t. \tag{2.11}$$

We can 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 $m_1 = 0$) and similarly

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

and hence by (2.6)

$$\frac{n_2}{m_0+m_2} \leq \frac{n}{m} \leq \frac{n_0+n_1}{m_1}. \tag{2.12}$$

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 $\vec{b}$ to consist of $\lfloor\alpha_1$ copies of $2^{n_0+n_1}3^{m_1}$ and $\lfloor\alpha_2\rfloor$ copies of $2^{n_2}3^{m_0+m_2}$. By (2.11), 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 (2.7).

The expression $2^n3^m/\prod\vec{b}$ 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

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

and the bound (2.8) follows. □

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

Lemma 1.1 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

$$|x|_{a_+,a_-} := \max(a_+x, -a_-x),$$

thus this is $a_+|x|$ when x is positive and $a_-|x|$ when x is negative. 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 1.1 as follows.

Proposition 3.1 (Reformulated balance criterion)**.** *Let $1 \leq t \leq N$, and suppose that one has a t-admissible tuple $\vec{b}$ such that*

$$E_t(\vec{b}) + \sum_p |v_p(N!/\prod \vec{b})|_{\log p,\infty} \leq \log N! - N \log t. \tag{3.1}$$

Then $t(N) \geq t$.

Indeed, the infinite penalty for making $v_p(N!/\vec{b})$ in (3.1) ensures that $\vec{b}$ is a subfactorization of $N!$.

We will reduce this infinite penalty term later, but let us work on other aspects of the criterion (3.1) 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

4. 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 4.1 (Upper bound criterion)**.** [5, Lemma 2.1] *Suppose that $1 \leq t \leq N$ are such that*

$$\sum_{p > \frac{t}{\lfloor\sqrt{t}\rfloor}} \left\lfloor \frac{N}{p} \right\rfloor \log\left(\frac{p}{t}\left\lceil\frac{t}{p}\right\rceil\right) > \log N! - N \log t \tag{4.1}$$

Then $t(N) < t$.

A surprisingly sharp upper bound comes from linear programming.

Lemma 4.2 (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*

$$\sum_p w_p v_p(j) \geq 1 \tag{4.2}$$

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

$$\sum_p w_p v_p(N!) < N. \tag{4.3}$$

Then $t(N) < t$.

Proof. We first observe that the bound (4.2) 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 (4.2) fails. Take a prime p dividing j and replace it by a prime inthe 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 (4.2) that

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

contradicting (4.3). □

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`