**NOTES ON THE GUY-SELFRIDGE PROBLEM**

TERENCE TAO

## 1. Basics

We recall Legendre's formula

$$\nu_p(N) = \sum_{j=1}^\infty \lfloor \frac{N}{p^j} \rfloor = \frac{N - s_p(N)}{p-1}. \tag{1.1}$$

## 2. Criteria for factorization

Suppose we are trying to factorize $N!$ into factors of size at least $t$. A candidate tuple $\vec b = (b_1, \dots, b_{N'})$ is said to be *admissible* if $b_j \geq t$ for all $j = 1, \dots, N'$. The *undershoot* $S_p^-(\vec b)$ and *overshoot* of a tuple at a prime $p$ are defined by the formulae

$$S_p^-(\vec b) := (\nu_p(N) - \sum_{i=1}^{N'} \nu_p(b_i))_+$$

$$S_p^+(\vec b) := (\sum_{i=1}^{N'} \nu_p(b_i) - \nu_p(N!))_+.$$

By the fundamental theorem of arithmetic, we obtain a perfect factorization $N! = b_1 \dots b_{N'}$ if the undershoots and overshoots all vanish. The *excess* $E(\vec b)$ is defined to be the quantity

$$E(\vec b) := \sum_{i=1}^{N'} \log \frac{b_i}{t}. \tag{2.1}$$

This quantity is non-negative for admissible tuples. Intuitively, the smaller the excess, the more efficient the candidate factorization.

**Proposition 2.1.** *Let $2 \leq t \leq N$. Suppose one can find an admissible tuple $\vec b$ obeying the inequality*

$$E(\vec b) + \sum_p (S_p^-(\vec b) \log p + S_p^+(\vec b) \log 2) \leq \log N! - N \log t \tag{2.2}$$

*as well as the side condition*

$$\sum_{p>2} S_p^+(\vec b) \left\lceil \frac{\log p}{\log 2} \right\rceil < S_2^-(\vec b). \tag{2.3}$$

*Then $t(N) \geq t$.*

*Proof.* Since there is a non-zero undershoot at $p = 2$, there is no overshoot at this prime. Suppose that we have an overshoot at an odd prime $p > 2$, thus one of the $b_i$ has a factor of $p$. If we replace this factor of $p$ by the larger quantity $2^{\lceil \frac{\log p}{\log 2} \rceil}$, then $S_p^+(\vec{b})$ goes down by one, $E(\vec{b})$ goes up by at most $\log 2$, and both sides of (2.3) decrease by the same amount. Thus neither (2.2) nor (2.3) can be destroyed by this process. Iterating, we may assume without loss of generality that there are no overshoots, thus

$$W(\vec{b}) + \sum_p (v_p(N!) - \sum_{i=1}^{N'} v_p(b_i)) \log p \leq \log N! - N \log t.$$

From the fundamental theorem of arithmetic we have

$$\sum_p v_p(b_i) \log p = \log b_i; \quad \sum_p v_p(N!) \log p = \log N!.$$

Using this and (2.1), we may rearrange the previous inequality as

$$N \log t \leq N' \log t.$$

If we then delete all but $N$ of the terms in the tuple $(b_1, \dots, b_{N'})$, and then distribute all undershoots amongst these surviving terms arbitrarily, we obtain a factorization $N! = a_1 \dots a_N$ with all $a_i \geq t$, so that $t(N) \geq t$ as required. □

In view of this proposition, we no longer need to keep direct track of the number of terms in the factorization; as long as we keep the excess small, and have not too much undershoot or overshoot, taking particular care to avoid undershoot at large primes, and to keep some undershoot at 2 for the side condition, we get a lower bound on $t(N)$. For $t = N/3$, the right-hand side $\log N! - N \log t$ is roughly $N \log \frac{3}{e} \approx 0.098N$ by the Stirling approximation.

## 3. Explicit estimates for primes

**Theorem 3.1** (Buthe's bounds)**.** [1] *For any* $2 \leq x \leq 10^{19}$, *we have*

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

*and*

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

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

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

## 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.