A Buchstab–Dickman Note and an Attempted Port to an Extremal Coprime Sieve Problem

1 Setup: the extremal problem

Fix a constant $C > 0$ and let $n \to \infty$. Consider sets

$$A \subseteq \{2, 3, \dots, n\}$$

such that

1. A is pairwise coprime: $(a, b) = 1$ for all distinct $a, b \in A$;
2. the harmonic budget constraint holds:

$$\sum_{a \in A} \frac{1}{a} \leq C.$$

Let

$$U_A(n) := \#\{m \leq n : \ \forall a \in A, \ a \nmid m\}$$

be the number of integers up to n not divisible by any element of A. The extremal quantity of interest is

$$F_n(C) := \min_A \frac{U_A(n)}{n}, \qquad F(C) := \liminf_{n \to \infty} F_n(C).$$

A natural candidate construction is to choose A as a *tail of primes*:

$$A = \{p \in \mathbb{P} : \ y < p \leq n\},$$

for a cutoff $y = y(n)$ chosen so that $\sum_{y<p\leq n} \frac{1}{p} \approx C$.

2 Smooth numbers and the tail-of-primes construction

For $m \geq 2$ let $P^+(m)$ denote the largest prime factor of m (with $P^+(1) = 1$). Define the smooth counting function

$$\Psi(x, y) := \#\{m \leq x : \ P^+(m) \leq y\}.$$

If $A = \{p : \ y < p \leq n\}$ is a set of primes, then an integer $m \leq n$ is not divisible by any element of A if and only if it has no prime factor exceeding y, i.e. $P^+(m) \leq y$. Hence

$$U_A(n) = \Psi(n, y).$$

If y is chosen so that $\sum_{y<p\leq n} \frac{1}{p} \to C$ as $n \to \infty$, then heuristically (and in fact, by Mertens-type estimates) one expects

$$\log y \sim e^{-C} \log n, \qquad \text{i.e.} \qquad y = n^{e^{-C}+o(1)}.$$

Write

$$u = \frac{\log n}{\log y}.$$

Then $u = e^C + o(1)$, and the Dickman–de Bruijn theorem predicts

$$\Psi(n, y) \sim n\rho(u) \sim n\rho(e^C),$$

where ρ is the Dickman function.

This motivates the conjectural value

$$F(C) \stackrel{?}{=} \rho(e^C).$$

3 Buchstab's identity and the Dickman equation

The standard route to Dickman asymptotics begins with a decomposition by the largest prime factor.

3.1 Layer decomposition

For a prime p, integers $m \le x$ with $P^+(m) = p$ are exactly those of the form $m = pt$ with $t \le x/p$ and $P^+(t) \le p$. Hence

$$\#\{m \le x : \ P^+(m) = p\} = \Psi(x/p, p). \tag{1}$$

Equivalently,

$$\Psi(x, p) - \Psi(x, p^-) = \Psi(x/p, p),$$

where p^- is the previous prime.

3.2 Buchstab's identity

Summing (1) over primes in an interval $(y, z]$ yields

$$\Psi(x, z) - \Psi(x, y) = \sum_{y<p\le z} \Psi(x/p, p),$$

so

$$\Psi(x, y) = \Psi(x, z) - \sum_{y<p\le z} \Psi(x/p, p). \tag{2}$$

This is a common form of *Buchstab's identity*.

3.3 From Buchstab to the Dickman differential-delay equation

Fix $u > 1$ and put $y = x^{1/u}$. In (2) take

$$z := x^{1/(u-1)},$$

so that $y < z$ for $u > 1$. Then

$$\Psi(x, x^{1/u}) = \Psi(x, x^{1/(u-1)}) - \sum_{x^{1/u}<p\le x^{1/(u-1)}} \Psi(x/p, p). \tag{3}$$

Assuming inductively an asymptotic of the form

$$\Psi(x/p,p)\approx \frac{x}{p}\rho\biggl(\frac{\log(x/p)}{\log p}\biggr),$$

note that if one sets $v:=\frac{\log x}{\log p}$ (i.e. $p=x^{1/v}$), then

$$\frac{\log(x/p)}{\log p}=\frac{\log x-\log p}{\log p}=v-1.$$

Dividing (3) by x suggests

$$\frac{\Psi(x,x^{1/u})}{x}\approx\frac{\Psi(x,x^{1/(u-1)})}{x}-\sum_{x^{1/u}<p\le x^{1/(u-1)}}\frac{1}{p}\,\rho(v-1).$$

The prime harmonic measure $\sum_{p\in(\cdot)}\frac{1}{p}$ is asymptotically approximated by an integral with density $\frac{dt}{t\log t}$, so

$$\sum_{x^{1/u}<p\le x^{1/(u-1)}}\frac{1}{p}\,\rho\biggl(\frac{\log x}{\log p}-1\biggr)\approx\int_{x^{1/u}}^{x^{1/(u-1)}}\rho\biggl(\frac{\log x}{\log t}-1\biggr)\frac{dt}{t\log t}.$$

With the substitution $v=\frac{\log x}{\log t}$, one has $\frac{dt}{t\log t}=-\frac{dv}{v}$ and the limits become $t=x^{1/u}\mapsto v=u$ and $t=x^{1/(u-1)}\mapsto v=u-1$, giving

$$\int_{u-1}^{u}\frac{\rho(v-1)}{v}\,dv.$$

Thus one is led to the integral recurrence

$$\rho(u)=\rho(u-1)-\int_{u-1}^{u}\frac{\rho(v-1)}{v}\,dv,\qquad u>1,$$

which is equivalent (upon differentiating) to the classic delay differential equation

$$u\rho'(u)+\rho(u-1)=0,\quad u>1,\qquad\text{with}\qquad\rho(u)=1\ \text{ for }\ 0\le u\le 1.\tag{4}$$

This is the mechanism by which Buchstab's identity and prime harmonic measure yield the Dickman function.

4 Attempted port to the extremal uncovered set and the state obstruction

Fix an admissible set A (pairwise coprime, $\sum_{a\in A}\frac{1}{a}\le C$), and define the uncovered set

$$S_A:=\{m\in\mathbb{N}:\ \forall a\in A,\ a\nmid m\}.$$

Define the *restricted smooth counting function*

$$\Psi_A(x,y):=\#\{m\le x:\ P^+(m)\le y,\ m\in S_A\}.$$

Note that $\Psi_\varnothing(x,y)=\Psi(x,y)$.

4.1 Layer decomposition still holds, but the recursion does not close

Decompose by the largest prime factor:

$$\Psi_A(x,z) - \Psi_A(x,y) = \sum_{y<p\le z} \#\{m \le x :\ P^+(m) = p,\ m \in S_A\}.$$

For each prime p, write $m = pt$ with $t \le x/p$ and $P^+(t) \le p$. The condition $m \in S_A$ depends on whether p appears in some element of A.

Because A is pairwise coprime, *at most one* element of A is divisible by p. Let $a(p) \in A$ denote that element if it exists.

Case 1: p is *good* (no entanglement). If $p \nmid a$ for all $a \in A$, then $\gcd(p,a) = 1$ for every $a \in A$, and for $m = pt$ one has

$$a \mid m \iff a \mid t.$$

Hence $m \in S_A$ if and only if $t \in S_A$, and therefore

$$\#\{m \le x :\ P^+(m) = p,\ m \in S_A\} = \Psi_A(x/p, p).$$

For such primes, the smooth-number Buchstab recursion carries over verbatim.

Case 2: p is *entangled* (some modulus contains p). Suppose there exists $a(p) \in A$ divisible by p. Write

$$a(p) = p^e b, \qquad e \ge 1, \quad \gcd(p,b) = 1.$$

Then for $m = pt$,

$$a(p) \mid m \iff p^e b \mid pt \iff p^{e-1} b \mid t.$$

Thus the “forbidden divisor” $a(p)$ for m transforms into a *different* forbidden divisor $a(p)/p$ for t. This motivates defining an *updated forbidden set*

$$A_p := (A \setminus \{a(p)\}) \cup \{a(p)/p\}.$$

With this definition, one checks:

$$m \in S_A \iff t \in S_{A_p}.$$

Hence the correct layer identity becomes

$$\#\{m \le x :\ P^+(m) = p,\ m \in S_A\} = \Psi_{A_p}(x/p, p),$$

not $\Psi_A(x/p, p)$.

4.2 Multi-state Buchstab recursion

Plugging the two cases into the layer sum yields a *multi-state* recursion:

$$\Psi_A(x,y) = \Psi_A(x,z) - \sum_{\substack{y<p\le z \\ p \nmid \prod_{a\in A} a}} \Psi_A(x/p, p) - \sum_{\substack{y<p\le z \\ p \mid \prod_{a\in A} a}} \Psi_{A_p}(x/p, p). \tag{5}$$

This is the precise obstruction to porting the standard Dickman derivation: the recursion does not close on a single function $\Psi_A(\cdot,\cdot)$, but instead propagates across a family of evolving states

$$A \to A_p \to (A_p)_q \to \cdots$$

created by repeatedly dividing elements of A by primes encountered as potential largest prime factors.

4.3 What a proof of optimality would need

The tail-of-primes conjecture predicts that the extremal uncovered density is $\rho(e^C)$, achieved by choosing A to be a prime tail. The Buchstab–Dickman mechanism aligns perfectly with this prediction *once A is prime*, since then there are no entangled primes and the recursion closes:

$$\Psi_A(x, y) = \Psi_A(x, z) - \sum_{y<p\leq z,\ p\notin A} \Psi_A(x/p, p).$$

Thus, a plausible proof strategy is:

1. **Structural reduction:** show that allowing composite moduli does not improve the extremal value, i.e. an asymptotically optimal family may be taken to consist essentially of primes (or at least that entangled primes are never helpful).

2. **Continuous limit / control problem:** for prime A, the constraint $\sum_{p\in A} \frac{1}{p} \leq C$ corresponds to removing a set of total "length" C in the prime-harmonic measure $d\mu(p) \approx \frac{dp}{p \log p}$, and one must show that removing the top interval (the tail of primes) minimizes the resulting solution of the Buchstab recursion, yielding $\rho(e^C)$.

The hard step is (1): composite moduli create the state-changes $A \mapsto A_p$ in (5), and controlling (or ruling out) any advantage from this additional flexibility appears to require genuinely robust extremal sieve inequalities.

5 A clean regime: $C \leq \log 2$

When $C \leq \log 2$, the tail-prime cutoff satisfies $y \geq \sqrt{n}$ asymptotically, so any two distinct primes $p, q > y$ satisfy $pq > n$. Thus the sets of multiples $\{m \leq n : p \mid m\}$ are disjoint, and the union bound becomes sharp:

$$\# \bigcup_{p\in A} \{m \leq n : p \mid m\} = \sum_{p\in A} \left\lfloor \frac{n}{p} \right\rfloor = n \sum_{p\in A} \frac{1}{p} + o(n).$$

Hence the uncovered density is asymptotically $1 - \sum_{p\in A} \frac{1}{p}$, and the minimum is $1 - C$. Since $\rho(u) = 1 - \log u$ for $1 \leq u \leq 2$, this matches $\rho(e^C)$ in the range $C \leq \log 2$.

Conclusion. The Buchstab–Dickman machinery gives a transparent derivation of the tail-prime performance and predicts the optimal constant $\rho(e^C)$. Attempting to extend the derivation to arbitrary admissible coprime A reveals a concrete obstruction: composite moduli induce state changes $A \mapsto A_p$ which prevent the recursion from closing on a single function, turning the problem into a multi-state extremal sieve.