Buchstab's Function and Irregularities in the distribution of primes

Recall \( \Phi(x,z) = \# \{ n \leq x : P^{-1}(n) > Z \} \). By Theorem 4.3.1,

\[
\Phi(x,x^{1/2}) = S^+(x) + S^-(x)
\]

\[
\Phi(x,x^{1/2}) = \frac{x\omega(s)}{\log(x^{1/2})} + O \left( \frac{x}{\log^2 x} \right)
\]

Uniformly for \( 1+1/N \leq s \leq N \) (\( N \) is any fixed positive integer), where

\[
\omega(s) = \frac{F(s)+f(s)}{2e^y}
\]

By the differential-delay equations for \( f \) and \( F \), we have

\[
\omega(s) = \frac{1}{s} \quad (1 \leq s \leq 2), \quad (s \omega(s)) = \omega(s-1) \quad (s > 2).
\]

Moreover, \( \omega \) has a continuous derivative for \( s > 2 \).

Lemma \( \omega \):

The function \( \omega(s) - e^{-s} \) changes sign infinitely often.

Proof: Recall that \( \lim_{s \to 0} \omega(s) = e^{-s} \). Suppose that \( \omega(s) \) has finitely many sign changes.

Let \( \Delta(s) = \omega(s) - e^{-s} \). By \( \omega(s) \),

\[
(\Delta) \quad (s \Delta(s))' = \Delta(s-1).
\]

Integrating from \( u \to \infty \) gives \( -u \Delta(u) = \int_{u-1}^{\infty} \Delta(s)ds \). Thus, if \( \Delta(u) > 0 \) for large \( u \) or \( \Delta(u) < 0 \) for large \( u \), then \( \Delta(u) = 0 \) for large \( u \).

Let \( u_0 = \inf \{ u : \Delta(s) = 0 \text{ for } s > u \} \). By \( \omega(s) \), \( u_0 > 2 \). But \( \Delta(u) \) implies that \( \Delta(s-1) = 0 \) for \( s > u_0 \), contradicting the definition of \( u_0 \).

Remarks: In fact, \( \omega(u) - e^{-u} \) changes sign on every unit interval \([u,u+1] \) for \( u > 1 \).
Primes in short intervals

It is generally expected that \( \pi(x) - \pi(x-y) \sim \frac{y}{\log x} \) if \( y \) is "not too small". Unconditionally, it is known that

\[
(\pi) \quad \pi(x) - \pi(x-y) \sim \frac{y}{\log x}
\]

for \( x^{1/3+\varepsilon} \leq y \leq x \) (Huxley, 1972). Assuming the Riemann Hypothesis, \((\pi)\) holds for \( x^{1/2+\varepsilon} \leq y \leq x \), and (Selberg, 1943) holds for "almost all" intervals \([x-y, x]\) if \( y = (\log x)^{1+\varepsilon} \); that is, if \( y = (\log x)^{1+\varepsilon} \), \((\pi)\) holds for all \( x \leq X \) except for a set of \( x \leq X \) of measure \( o(X) \).

By Theorem \( \Delta \), we know that \((\pi)\) is false for some \( x \) and

\[
y = C \log x \frac{\log x}{(\log x)^2}.
\]

Cramér argued heuristically that

\[
0 < \limsup_{n \to \infty} \frac{P_{n+1} - P_n}{\log^2 P_n} < \infty,
\]

so that \((\pi)\) shouldn't hold for \( y = C (\log x)^2 \).

**Theorem M** (H. Maier, 1985)

For every \( \lambda > 1 \), \((\pi)\) is false for \( y = (\log x)^{\lambda} \). More precisely,

\[
\liminf_{x \to \infty} \frac{\pi(x) - \pi(x-y)}{y/\log x} < 1, \quad \limsup_{x \to \infty} \frac{\pi(x) - \pi(x-y)}{y/\log x} > 1.
\]

The proof is based on Lemma 18 and Maier's "matrix method". We also need a result on primes in progressions.

**Lemma Ga** (Gallagher, 1970)

For certain positive constants \( c_1, c_2, c_3 \) the following holds:

If \( q \geq c_1 \), and for all Dirichlet characters \( \chi \) modulo \( q \), \( L(\sigma, \chi) \neq 0 \) for \( \sigma \gtrsim \frac{c_2}{\log(q(q+2))} \), and \( (a,q)=1 \), then

\[
\pi(2x, q, \chi) - \pi(x, q, \chi) = \frac{\xi(\chi, q, \chi)}{\phi(q)} \left( 1 + O\left( e^{-c_3 \frac{\log x}{\log q} + e^{-c_2 \sqrt{\log x}}} \right) \right).
\]

**Lemma**

For infinitely many integers \( k \), \( q = \Pi_{p \leq k} p \) satisfies the hypotheses of Lemma Ga, if \( c_2 \) is small enough.
Proof By classical results, if $c_4>0$ is small enough, then for all $q$,

\[(*) \quad L\left(\frac{r+it}{2}, \chi\right) \neq 0 \quad \text{for} \quad \sigma > 1 - \frac{\varepsilon_4}{\log(qk+3)} \]

with the possible exception of one real zero for one real $\chi$.

Let $k$ be large, $q = \prod p_k$. If the exceptional zero $\alpha$ exists for $\chi \mod q$, then $\alpha > 1 - \frac{c_4}{\log(qk)}$. Since $\log q \sim k$, there is a $k' \geq k$ so that if $q' = \prod_{p \leq k'} p$, then

\[1 - \frac{c_4}{\log(qk)} < \alpha < 1 - \frac{c_4/2}{\log(qk')}.\]

As $\chi$ induces a character $\chi'$ modulo $q'$ (both with the same conductor), $L(\beta, \chi') = 0$ and hence $(\chi', \beta)$ is the unique pair of (character, zero) that violates $(*)$ with $q$ replaced by $q'$. Hence, with $c_2 = \frac{c_4}{2}$, the hypothesis of Lemma 1a holds either with modulus $q$ or with $q'$.

Proof of Theorem M

Fix $\lambda$ and some $\lambda_1 > \lambda$. Let $k$ be sufficiently large such that

\[q = \prod_{p \leq k} p\]

satisfies the hypothesis of Lemma 1a. Put

\[U = \left| k^{\lambda_1} \right| \]
\[D = \frac{\log \log q}{1} + 1 \]
\[R = q^{\frac{1}{2}} \]
\[x = R_2 = q^2.\]

Consider the matrix

\[
M = \begin{pmatrix}
1 + R_2 & 2 + R_2 & \cdots & U + R_2 \\
1 + (2R_1)R & 2 + (2R_1)R & \cdots & U + (2R_1)R \\
\vdots & \vdots & & \vdots \\
1 + (2R-1)R & 2 + (2R-1)R & \cdots & U + (2R-1)R
\end{pmatrix}
\]

If $k$ is large, then $U \ll q$, since $U \sim (\log q)^\lambda$ (here, by $\sim$ and $o(1)$ we mean as $k \to \infty$, $U, D, R$ and $x$ also $\to \infty$ as $k \to \infty$). Hence the entries of $M$ are distinct integers in $(x, 2x]$. 
Let $N$ be the number of primes in $M$. By lemma Ga,

$$N = \sum_{\substack{1 \leq a \leq U \\ (a,q) = 1}} \pi(2x, q, a) - \pi(x, q, a)$$

$$= \sum_{\substack{1 \leq \varepsilon \leq \nu \\ (\varepsilon, q) = 1}} \frac{\text{li}(2x) - \text{li}(x)}{\phi(q)} \left(1 + O\left(e^{\frac{\varepsilon^2}{\log x}} + e^{-c_2 \log \log q}\right)\right)$$

$$\sim \frac{x}{\phi(q) \log x} \cdot \Phi(U, k).$$

By (\Phi),

$$\Phi(U, k) \sim \frac{\omega(x) U}{\log k},$$

and by Mertens estimate,

$$\phi(q) = \sigma(q) \prod_{p \leq k} (1 - \frac{1}{p}) \sim \frac{e^{-\gamma} q}{\log k}.$$ 

Therefore,

$$(N) \quad N \sim \frac{ux e^{\omega(\lambda_1)}}{x \log x}.$$ 

Part I. Take $\lambda_1$ so that $e^{\omega(\lambda_1)} < 1$. Such $\lambda_1$ exists by lemma $\omega$. There are $R$ rows of $M$, each consisting of an interval of $U$ integers.

By (N), some row has

$$\leq (1 + o(1)) \frac{U e^{\omega(\lambda_1)}}{\log x}$$

primes. Note that

$$U \sim (\log q)^{\lambda_1} = (\log x)^{\lambda_1 + o(1)}.$$ 

Break this row into $(\log q)^{\lambda_1 + o(1)}$ intervals of length $y = (\log 2x)^{\lambda}$. Then some subinterval contains at most

$$(1 + o(1)) \frac{U}{(\log x)^{\lambda_1 + o(1)}} \frac{e^{\omega(\lambda_1)}}{\log x} \sim \frac{y}{\log x} e^{\omega(\lambda_1)},$$

primes. That is, for some $x' \in (x, 2x],$

$$\frac{\pi(x') - \pi(x' - (\log x')^{\lambda_1})}{(\log x')^{\lambda_1 + 1}} \leq \frac{\pi(x') - \pi(x' - y)}{(\log x')^{\lambda_1 + 1}} \leq (1 + o(1)) e^{\omega(\lambda_1)}.$$

For large enough $k$ the right side is <1.
Part II. Take $\lambda$, so that $e^{\pi \omega(\lambda)} > 1$. Such $\lambda$, exists by Lemma $\omega$.
By (M), some row of $M$ has
$$\geq (1+o(1)) \frac{\log x}{\log x} e^\pi \omega(\lambda),$$
primes. Let $y = (\log x)^{\lambda}$. By an argument similar to that in Part I, some interval of length $y$ in some row of $M$ has
$$\geq (1+o(1)) \frac{y}{\log x} e^\pi \omega(\lambda),$$
primes. Then, for some $x' < (x, 2x]$,\[
\frac{\pi(x') - \pi(x' - (\log x)^{\lambda})}{(\log x)^{\lambda-1}} \geq \frac{\pi(x) - \pi(x' - y)}{(\log x)^{\lambda-1}} \geq (1+o(1)) e^\pi \omega(\lambda).
\]
For large $k$, the right side is $> 1$.

Primes in progressions to large moduli

Let $\Delta(x; q) = \sum_{q \equiv a \pmod{q}} \max_{g \leq x} \left| \psi(g; q, a) - \frac{g}{q} \right|$, $\psi(g; q, a) = \sum_{\substack{n \leq g \equiv a \pmod{q}}} \Lambda(n)$.

Theorem BV is: $\forall A > 0 \exists \delta$ such that
$$\Delta(x; \sqrt{\log x}^{-\delta}) \ll \frac{x}{(\log x)^{2A}}.$$

Elliott-Halberstam conjecture: $\forall A > 0, \forall \epsilon > 0$, $\Delta(x; x^{-1-\epsilon}) \ll \frac{x}{(\log x)^A}$.

E.H. also conjecture: $\forall A > 0 \exists B > 0$ so that $\Delta(x; x; (\log x)^{-B}) \ll \frac{x}{(\log x)^{B}}$.
False!

Theorem. For all $B > 0$,
$$\Delta(x; x; (\log x)^{-B}) \gg \frac{x}{(\log x)^{B}}.$$ (Friedlander, Granville, 1989)

Proof. Fix $\lambda > B$ with $e^{\pi \omega(\lambda)} \neq 1$ (such $\lambda$ exists by Lemma $\omega$). Put $Z = (\log x)^{B/x}$, and
$$a = \prod_{p \leq Z} p$$
so that $a < \sqrt{x}$. Let $Q = x/(\log x)^B$ and consider
$$S = \sum_{\substack{\frac{a}{Z} < g \leq Q \\ \left(\frac{a}{Z}, g\right) = 1}} \psi(g; a, 0) = \sum_{\substack{\frac{a}{Z} < g \leq Q \\ \left(\frac{a}{Z}, g\right) = 1}} \left( \sum_{p \leq x} \log p \right) + O(\sqrt{x}),$$

since for $b > 2$, $b^a$ has at most 1 prime factors $> a/2$, and $p^b \leq x.$
If \( \frac{a}{2} < q \leq Q \) and \( p \equiv a \pmod{q} \) is prime with \( p \leq x \), then
\[
p = rq + a, \quad r = \frac{2(x-a)}{Q} \quad \text{and} \quad \frac{a}{2} - r + a < p \leq \min\left(x, \frac{Qr+a}{2}\right).
\]

Hence
\[
S = \sum_{r \leq \Delta} \psi\left(\max\left(x, \frac{a}{2} + r \cdot a\right); r, a\right) - \psi\left(\frac{a}{2} + r \cdot a, r, a\right) + O(x) \quad \text{(use Siegel-Walfisz)}
\]
\[
= \left(1 + O\left(\frac{1}{\log x}\right)\right) \sum_{r \leq \Delta} \frac{\max\left(x, \frac{a}{2} + r \cdot a\right) - \left(\frac{a}{2} - r + a\right)}{\phi(r)}.
\]

Now \( r \leq 2(\log x)^8 \), hence
\[
\phi(r) = r \prod_{p \mid r} (1 - \frac{1}{p}) = r e^{O\left(\sum_{p \mid r} \frac{1}{p}\right)} = r e^{O\left(\frac{\log r}{r}\right)} = r \left(1 + O\left(\frac{1}{r}\right)\right).
\]

Thus, as \( x \to \infty \),
\[
S \sim \sum_{r \leq \Delta} \frac{\max\left(x, \frac{a}{2} + r \cdot a\right) - \left(\frac{a}{2} - r + a\right)}{r}.
\]
\[
= \frac{\Delta}{\phi}\Phi\left(\frac{a}{2}, z\right) - \frac{\Delta}{2}\Phi\left(\frac{2x}{\Delta}, z\right) - \Phi\left(\frac{a}{2}, z\right) + \Delta \sum_{\substack{\frac{a}{2} < r \leq 2x/a \\ (r, a) = 1}} \frac{1}{r}.
\]
\[
= \Phi\left(\frac{a}{2}, z\right) - \frac{\Delta}{2}\Phi\left(\frac{2x}{\Delta}, z\right) + x \int_{x/\Delta}^{2x/\Delta} \frac{\Phi(t, z)}{t^2} \, dt + \int_{x/\Delta}^{2x/\Delta} \Phi(t, z) \, dt.
\]
\[
= \int_{x/\Delta}^{2x/\Delta} \Phi(t, z) \, dt.
\]

Uniformly for \( \frac{x}{\Delta} \leq t \leq \frac{2x}{\Delta} \),
\[
\frac{\log t}{\log z} = \frac{B \log \log x + O(1)}{(B \lambda) \log \log x} \sim \lambda,
\]
so that
\[
S \sim x \cdot \frac{\omega(\lambda) \log 2}{\log z} \int_{x/\Delta}^{2x/\Delta} \frac{dt}{t} = x \cdot \frac{\omega(\lambda) \log 2}{\log z}.
\]

On the other hand, we obtain an asymptotic for
\[
T := \sum_{\substack{\frac{a}{2} < q \leq Q \\ (a, q) = 1}} \frac{x}{\Phi(q)}.
\]

Break the sum into two parts: \( T_1 \), the sum over \( q \) with \( \omega(q) \geq 10 \log \log Q \), and \( T_2 \), the sum over those \( q \) with \( \omega(q) < 10 \log \log Q \).
By Proposition C,
\[ \# \{ \frac{\theta}{2} < \theta \leq Q : \omega(q) > 10 \log \log Q \} \ll \frac{Q}{(\log Q)^{5-\delta}} \ll \frac{Q}{(\log x)^{3}}. \]

Hence
\[ T_1 \ll \sum_{\frac{\theta}{2} < \theta \leq Q \atop \omega(q) > 10 \log \log Q} \frac{x \log \log Q}{Q} \ll \frac{x}{(\log x)^{3}}. \]

Also, for \((g, q) = 1\), \(P\{g\} = \frac{1}{2}\), so for \(\omega(q) \geq 10 \log \log Q\),
\[ \phi(q) = g \prod_{p \mid q} (1 - \frac{1}{p}) = g e^{O(\frac{\omega(q)}{\log q})} = g \left(1 + O\left(\frac{\log \log q}{\log q}\right)\right) \sim g, \]
so that
\[ T_2 \sim x \sum_{\frac{\theta}{2} < \theta \leq Q \atop \omega(q) > 10 \log \log q} \frac{1}{\theta} = x \sum_{\frac{\theta}{2} < \theta \leq Q \atop \omega(q) > 10 \log \log q} \frac{1}{\theta} + O\left(\frac{x}{(\log x)^{3}}\right) \]
\[ \sim x \left[ \frac{\phi(t, z)}{t} \right]_{\theta/2}^{Q} + \int_{\theta/2}^{Q} \frac{\phi(t, z)}{t} \, dt \right] \sim x \int_{\theta/2}^{Q} \frac{e^{-\gamma \log z}}{t} \, dt \sim \frac{x e^{-\gamma \log z}}{\log z}. \]

Therefore,
\[ \Delta(x; \frac{x}{(\log x)^{B}}) \geq \sum_{\frac{\theta}{2} < \theta \leq Q \atop (a, q) = 1} \left| \psi(x; g, q) - \frac{x}{\phi(q)} \right| \]
\[ \geq \sum_{\frac{\theta}{2} < \theta \leq Q \atop (a, q) = 1} \left| \psi(x; g, q) - \frac{x}{\phi(q)} \right| \]
\[ \sim \frac{x \log 2}{\log z} \left| \omega(\lambda) - e^{-\gamma} \right| \gg \frac{x}{\log \log x}. \]

Theorem (Friedlander, Granville, Hildebrand, Maier 1991)
\[ \forall \alpha > 0, \text{ if } Q = e^{\exp\left\{ - \frac{A}{2} \left( \frac{\log x}{\log \log x} \right)^{2} \right\}}, \text{ then} \]
\[ \Delta(x; Q) \gg \frac{x}{(\log x)^{A}}. \]