Arithmetic functions in short intervals

Brad Rodgers

Department of Mathematics
University of Michigan

Analytic Number Theory and Arithmetic Session
JMM Jan. 2017
A problem (context to come!):

Let $\omega(n) = \sum_{p | n} 1$ be the number of distinct prime factors of $n$.

Is it true that as $X \to \infty$, with $H = X^{\delta}$ with $\delta \in (0, 1)$,

$$
\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right)
= \frac{1}{X} \int_{X}^{2X} \left( \sum_{t \leq n \leq t+H} \omega(n) \right)^2 \, dt - \left( \frac{1}{X} \int_{X}^{2X} \sum_{t \leq n \leq t+H} \omega(n) \, dt \right)^2
= O_{\delta}(H) \quad ?
$$

Another guess might be that the variance is $\sim H \log \log X$. 

Brad Rodgers (UM)  
Short intervals  
Jan. 2017 2 / 14
A problem (context to come!):

Let \( \omega(n) = \sum_{p|n} 1 \) be the number of distinct prime factors of \( n \).

Is it true that as \( X \to \infty \), with \( H = X^\delta \) with \( \delta \in (0, 1) \),

\[
\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right) := \frac{1}{X} \int_X^{2X} \left( \sum_{t \leq n \leq t+H} \omega(n) \right)^2 \, dt - \left( \frac{1}{X} \int_X^{2X} \sum_{t \leq n \leq t+H} \omega(n) \, dt \right)^2
\]

\[= O_\delta(H) \ ? \]

Another guess might be that the variance is \( \sim H \log \log X \).
Primes in short intervals

A (weighted) count of primes in the interval \([t, t + H]\) is given by

\[
\sum_{t \leq n \leq t + H} \Lambda(n)
\]

where \(\Lambda(n) = \begin{cases} 
\log p & \text{if } n = p^k \\
0 & \text{otherwise}
\end{cases}
\)

This should be \(\approx H\), but if \(t\) is random the count will oscillate around that value.

Conjecture (Goldston-Montgomery)

For \(H = X^{\delta}\) with \(\delta \in (0, 1)\),

\[
\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t + H} \Lambda(n) \right) := 1
\]

\[
\sim H (\log X - \log H)
\]

Conflicts with the Cramér model! Must take into account 'off-diagonal' Hardy-Littlewood type contributions.

(Note: \(H (\log X - \log H) = (1 - \delta) H \log X\).)
Primes in short intervals

A (weighted) count of primes in the interval $[t, t+H]$ is given by

$$\sum_{t \leq n \leq t+H} \Lambda(n)$$

where $\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \\ 0 & \text{otherwise} \end{cases}$

This should be $\approx H$, but if $t$ is random the count will oscillate around that value.
A (weighted) count of primes in the interval \([t, t + H]\) is given by

\[
\sum_{t \leq n \leq t+H} \Lambda(n) \quad \left(\text{where } \Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \\ 0 & \text{otherwise} \end{cases}\right)
\]

This should be \(\approx H\), but if \(t\) is random the count will oscillate around that value.

**Conjecture (Goldston-Montgomery)**

For \(H = X^\delta\) with \(\delta \in (0, 1)\),

\[
\text{Var}_{t \in [X,2X]} \left( \sum_{t \leq n \leq t+H} \Lambda(n) \right) := \frac{1}{X} \int_X^{2X} \left( \sum_{t \leq n \leq t+H} \Lambda(n) - H \right)^2 dt 
\sim H(\log X - \log H)
\]
Primes in short intervals

A (weighted) count of primes in the interval \([t, t + H]\) is given by

\[
\sum_{t \leq n \leq t + H} \Lambda(n)
\]

where \(\Lambda(n) = \begin{cases} 
\log p & \text{if } n = p^k \\
0 & \text{otherwise}
\end{cases}
\)

This should be \(\approx H\), but if \(t\) is random the count will oscillate around that value.

**Conjecture (Goldston-Montgomery)**

For \(H = X^\delta\) with \(\delta \in (0, 1)\),

\[
\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t + H} \Lambda(n) \right) := \frac{1}{X} \int_X^{2X} \left( \sum_{t \leq n \leq t + H} \Lambda(n) - H \right)^2 dt 
\sim H(\log X - \log H)
\]

**Conflicts with the Cramér model! Must take into account ‘off-diagonal’ Hardy-Littlewood type contributions.**
Primes in short intervals

A (weighted) count of primes in the interval \([t, t + H]\) is given by

\[
\sum_{t \leq n \leq t + H} \Lambda(n)
\]

\[\text{where } \Lambda(n) = \begin{cases} 
\log p & \text{if } n = p^k \\
0 & \text{otherwise}
\end{cases} \]

This should be \(\approx H\), but if \(t\) is random the count will oscillate around that value.

**Conjecture (Goldston-Montgomery)**

For \(H = X^\delta\) with \(\delta \in (0, 1)\),

\[
\text{Var}_{t \in [X,2X]} \left( \sum_{t \leq n \leq t + H} \Lambda(n) \right) := \frac{1}{X} \int_X^{2X} \left( \sum_{t \leq n \leq t + H} \Lambda(n) - H \right)^2 \, dt
\]

\[\sim H(\log X - \log H)\]

Conflicts with the Cramér model! Must take into account ‘off-diagonal’ Hardy-Littlewood type contributions.

(Note: \(H(\log X - \log H) = (1 - \delta)H \log X\).)
Divisors in short intervals

Consider the same question for

\[ d_k(n) = \#\{(a_1, \ldots, a_k) : a_1 \cdots a_k = n\}, \quad \zeta(s)^k = \sum_n \frac{d_k(n)}{n^s} \]

Conjecture (Keating – R. – Roditty-Gershon – Rudnick)

For \( \log H / \log X \to \delta \) with \( \delta \in (0, 1 - 1/k) \),

\[
\operatorname{Var}_{x \leq n \leq x+H} \left( \sum_{x \leq n \leq x+H} d_k(n) \right) \sim a_k P_k(\delta) H (\log X)^{k^2-1},
\]

with \( P_k(\delta) \) an explicitly written down piecewise-polynomial, with different ‘phases’ on the intervals

\[ (0, \frac{1}{2}), (\frac{1}{2}, \frac{2}{3}), \ldots, (\frac{k-2}{k-1}, \frac{k-1}{k}) , \]

and \( a_k \) is a (well-understood) arithmetic constant.
Divisors in short intervals

Consider the same question for

\[ d_k(n) = \# \{(a_1, \ldots, a_k) : a_1 \cdots a_k = n\}, \quad \zeta(s)^k = \sum_n \frac{d_k(n)}{n^s} \]

**Conjecture (Keating – R. – Roditty-Gershon – Rudnick)**

For \( \log H / \log X \to \delta \) with \( \delta \in (0, 1 - 1/k) \),

\[
\text{Var}_{t \in [X, 2X]} \left( \sum_{x \leq n \leq x+H} d_k(n) \right) \sim a_k P_k(\delta) H (\log X)^{k^2 - 1},
\]

with \( P_k(\delta) \) an explicitly written down piecewise-polynomial, with different ‘phases’ on the intervals

\((0, \frac{1}{2}), (\frac{1}{2}, \frac{2}{3}), \ldots, (\frac{k-2}{k-1}, \frac{k-1}{k})\),

and \( a_k \) is a (well-understood) arithmetic constant.

Conflicts even with the naive application of ‘off-diagonal’ Hardy-Littlewood type contributions!
Divisors in short intervals

Some rigorous evidence for the conjecture:

- **Lester** for $\delta \in \left( \frac{k-2}{k-1}, \frac{k-1}{k} \right)$

- **Harper – Soundararajan, R. – Soundararajan, de la Bretèche – Fiorilli**: An arithmetic progression analogue for $\delta \in (0, 1/2)$. 
Divisors in short intervals

Some rigorous evidence for the conjecture:

- *Lester* for $\delta \in \left(\frac{k-2}{k-1}, \frac{k-1}{k}\right)$

- *Harper – Soundararajan, R. – Soundararajan, de la Bretèche – Fiorilli:* An arithmetic progression analogue for $\delta \in (0, 1/2)$. Probably possible to also analyze $\delta \in \left(\frac{k-3}{k-2}, \frac{k-2}{k-1}\right), \left(\frac{k-2}{k-1}, \frac{k-1}{k}\right)$
Divisors in short intervals

Some rigorous evidence for the conjecture:

- *Lester* for $\delta \in \left( \frac{k-2}{k-1}, \frac{k-1}{k} \right)$

- *Harper – Soundarajan, R. – Soundararajan, de la Bretèche – Fiorilli:* An arithmetic progression analogue for $\delta \in (0, 1/2)$. Probably possible to also analyze $\delta \in \left( \frac{k-3}{k-2}, \frac{k-2}{k-1} \right), \left( \frac{k-2}{k-1}, \frac{k-1}{k} \right)$

But this doesn’t cover all phase changes. Best evidence is an analogous result for $\mathbb{F}_q[T]$. 
One may ask similar questions over $\mathbb{F}_q[T]$, and often one can answer them when $q \to \infty$. 
One may ask similar questions over $\mathbb{F}_q[T]$, and often one can answer them when $q \to \infty$.

A dictionary:

positive integers $\mathbb{N}$ $\longleftrightarrow$ $\mathbb{F}_q[T]$ monic
One may ask similar questions over $\mathbb{F}_q[T]$, and often one can answer them when $q \to \infty$.

**A dictionary:**

<table>
<thead>
<tr>
<th>Positive integers $\mathbb{N}$</th>
<th>$\leftrightarrow$</th>
<th>$\mathbb{F}_q[T]$ monic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log n$</td>
<td>$\leftrightarrow$</td>
<td>$\deg(f)$</td>
</tr>
<tr>
<td>Average over $[X, 2X]$</td>
<td>$\leftrightarrow$</td>
<td>Average over $\mathcal{M}_n := { f \text{ monic}, \deg(f) = n }$</td>
</tr>
</tbody>
</table>
One may ask similar questions over $\mathbb{F}_q[T]$, and often one can answer them when $q \to \infty$.

A dictionary:

- $\mathbb{Z}$ to $\mathbb{F}_q[T]$
- positive integers $\mathbb{N} \longleftrightarrow \mathbb{F}_q[T]$ monic monic
- $\log n \longleftrightarrow \deg(f)$
- average over $[X, 2X] \longleftrightarrow$ average over $\mathcal{M}_n := \{ f \text{ monic}, \deg(f) = n \}$
- primes $\longleftrightarrow$ irreducible polynomials
- $\Lambda(n) \longleftrightarrow \Lambda(f) := \begin{cases} \deg P, & \text{if } f = cP^k, \\ 0, & \text{otherwise.} \end{cases}$
- $d_k(n) \longleftrightarrow d_k(f)$
- $\omega(n) \longleftrightarrow \omega(f)$
One may ask similar questions over $\mathbb{F}_q[T]$, and often one can answer them when $q \to \infty$.

A dictionary:

- $\mathbb{Z}$ to $\mathbb{F}_q[T]$: $\mathbb{N} \leftrightarrow \mathbb{F}_q[T]$ monic
- $\log n \leftrightarrow \deg(f)$
- Average over $[X, 2X] \leftrightarrow$ average over $\mathcal{M}_n := \{f \text{ monic, } \deg(f) = n\}$
- Primes $\leftrightarrow$ irreducible polynomials
- $\Lambda(n) \leftrightarrow \Lambda(f) := \begin{cases} \deg P, & \text{if } f = cP^k, \\ 0, & \text{otherwise}. \end{cases}$
- $d_k(n) \leftrightarrow d_k(f)$
- $\omega(n) \leftrightarrow \omega(f)$
- $\{n : t \leq n \leq t + H\} \leftrightarrow I(f; h) := \{g \in \mathbb{F}_q[T] : \deg(f - g) \leq h\}$
- $H \leftrightarrow q^{h+1}$
Primes in short intervals of $\mathbb{F}_q[T]$

**Theorem (Keating-Rudnick)**

For $0 \leq h \leq n - 5$,

$$\text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f; h)} \Lambda(g) \right) \sim q^{h+1}(n - h - 2),$$

as $q \to \infty$. 

---

Brad Rodgers (UM)  
Short intervals  
Jan. 2017  
7 / 14
Primes in short intervals of $\mathbb{F}_q[T]$

**Theorem (Keating-Rudnick)**

For $0 \leq h \leq n - 5,$

$$\text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} \Lambda(g) \right) \sim q^{h+1}(n - h - 2),$$

as $q \to \infty$.

So up to first order:

$$q^{h+1}(n - h - 2) \longleftrightarrow H(\log X - \log H)$$
Divisors in short intervals of $\mathbb{F}_q[T]$

**Theorem (Keating – R. – Roditty-Gershon – Rudnick)**

For $0 \leq h \leq n - 5$,

$$\text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} d_k(g) \right) \sim q^{h+1} p_k(n, h),$$

as $q \to \infty$, where $p_k(n, h)$ is a piecewise polynomial in $n$ and $h$,
Divisors in short intervals of $\mathbb{F}_q[T]$

**Theorem (Keating – R. – Roditty-Gershon – Rudnick)**

For $0 \leq h \leq n - 5$,

$$\text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} d_k(g) \right) \sim q^{h+1} p_k(n, h),$$

as $q \to \infty$, where $p_k(n, h)$ is a piecewise polynomial in $n$ and $h$, with

$$p_k(n, h) \sim P_k(\delta)n^{k^2-1} \quad \text{as} \quad \frac{h}{n} \to \delta.$$
Prime divisors in short intervals, $\mathbb{Z}$

For $X = H^\delta$, with $\delta \in (0, 1)$, one might expect

$$\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right) \approx \text{Var}(\omega_1 + \cdots + \omega_H)$$

$$\sim H \log \log X,$$

where $\omega_1, \omega_2, \ldots, \omega_H$ are independent identically distributed random variables each with variance $\approx \log \log X$. 

But (I think) more likely is

$$\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right) = O(\delta(H)).$$
Prime divisors in short intervals, $\mathbb{Z}$

For $X = H^\delta$, with $\delta \in (0, 1)$, one might expect

$$\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right) \approx \text{Var}(\omega_1 + \cdots + \omega_H)$$

$$\sim H \log \log X,$$

where $\omega_1, \omega_2, \ldots, \omega_H$ are independent identically distributed random variables each with variance $\approx \log \log X$.

But (I think) more likely is

$$\text{Var}_{t \in [X, 2X]} \left( \sum_{t \leq n \leq t+H} \omega(n) \right) = O_\delta(H).$$
Prime divisors in short intervals, $\mathbb{F}_q[T]$

**Proposition (R.)**

For $0 \leq h \leq n - 5$,

\[
\text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} \omega(g) \right) \sim q^{h+1} \sum_{1 \leq \lambda_1 \leq \lambda_2 \leq n-h-2} \left( \frac{1}{n - \lambda_1} - \frac{1}{n - \lambda_2 + 1} \right)^2 \\
:= \rho(n, h)
\]

as $q \to \infty$.

And for $\frac{h}{n} \to \delta \in (0, 1)$,

\[
\rho(n, h) = O_\delta(1).
\]
Some simpler examples

Proposition

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} 1 \right) = 0 \]

Why? There is no variation in the short interval sums!

1 is regular/structured across a short interval.

Theorem

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} (f;h) \mu(g) \right) \sim q h + 1 \]

Heuristically why? This is just a formal statement of the idea that for \( \mu(f) \), only diagonal terms contribute when we expand the variance!

\( \mu(f) \) is oscillatory/random across a short interval.
Some simpler examples

**Proposition**

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f,h)} 1 \right) = 0 \]

Why? There is no variation in the short interval sums!

1 is **regular/structured** across a short interval.

**Theorem**

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f,h)} \mu(g) \right) \sim q^{h+1} \]

Heuristically why? This is just a formal statement of the idea that for \( \mu(f) \), only diagonal terms contribute when we expand the variance!

\( \mu(f) \) is **oscillatory/random** across a short interval.
Some simpler examples

**Proposition**

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} 1 \right) = 0 \]

Why? There is no variation in the short interval sums!

1 is **regular/structured** across a short interval.

**Theorem**

\[ \text{Var}_{f \in \mathcal{M}_n} \left( \sum_{g \in I(f;h)} \mu(g) \right) \sim q^{h+1} \]

Heuristically why? This is just a formal statement of the idea that for \( \mu(f) \), only diagonal terms contribute when we expand the variance!

\( \mu(f) \) is **oscillatory/random** across a short interval.
A structured/random decomposition

**Theorem**

For $0 \leq h \leq n - 5$,

$$d_k(f) = u(f) + v(f)$$

with the decomposition depending on $n$ and $h$, where

$$\operatorname{Var}_{f \in M_n} \left( \sum_{g \in I(f; h)} u(f) \right) = o(q^{h+1}) \quad \text{(regular/structured part)}$$

and

$$\operatorname{Var}_{f \in M_n} \left( \sum_{g \in I(f; h)} v(f) \right) \sim \frac{q^{h+1}}{q^{n+1}} \sum_{g \in M_n} |v(f)|^2 \quad \text{(oscillatory/random part)}$$
A structured/random decomposition

- This decomposition holds not just for $d_k(f)$, but *in general* for functions that depend only on the size and multiplicity of prime factors (e.g. $\Lambda(f), \omega(f)$, etc.).
- Can be given explicitly in terms of symmetric function theory.
A structured/random decomposition

- This decomposition holds not just for $d_k(f)$, but *in general* for functions that depend only on the size and multiplicity of prime factors (e.g. $\Lambda(f), \omega(f)$, etc.).
- Can be given explicitly in terms of symmetric function theory.
- Evaluating the oscillatory party makes use of the some deep equidistribution theorems of N. Katz.
A structured/random decomposition

- This decomposition holds not just for $d_k(f)$, but in general for functions that depend only on the size and multiplicity of prime factors (e.g. $\Lambda(f), \omega(f)$, etc.).
- Can be given explicitly in terms of symmetric function theory.
- Evaluating the oscillatory party makes use of the some deep equidistribution theorems of N. Katz.
- For me a pleasant surprise: D. Hast and V. Matei are able to use the result to turn back to a computation in algebraic geometry that seems difficult otherwise.
Thanks for your attention!