A Gamma Distribution Hypothesis for Prime $k$-tuples

J. LaChapelle

Abstract

We conjecture average counting functions for prime $k$-tuples based on a gamma distribution hypothesis for prime powers. The conjecture is closely related to the Hardy-Littlewood conjecture for $k$-tuples. Possessing average counting functions along with the exact functions introduced in [6] allows to implicitly define pertinent $k$-tuple zeta functions. The $k$-tuple zeta functions in turn allow construction of $k$-tuple analogs of explicit formulæ.

1 Introduction

The motivation and eagerness to better understand prime $k$-tuples hardly needs introduction. For a small sample of the literature (which is by no means exhaustive) see [1], [2], [3], [4], [5] and references therein. Of course there would be no suspense if there existed Euler products for $k$-tuples like that for single primes. Their absence seems to indicate that there do not exist generating Dirichlet series whose summands are completely multiplicative in these cases. It is fair to say that this is at the heart of the difficulty in generalizing the single prime case.

This paper hopefully takes a step toward better understanding the distribution of prime $k$-tuples. It is based on two ideas. The first idea [6] uses the arithmetic function $\mu(n)\Lambda(n)/\log(n)$ to represent the exact prime counting function

$$\pi(x) = -\sum_{n\leq x} \mu(n)\frac{\Lambda(n)}{\log(n)}.$$

This simple representation can be readily extended to prime $k$-tuples. From the $k$-tuple counting function, one sees that prime $k$-tuples can be efficiently modeled on the pair-wise coprime $k$-lattice.

The second idea is the gamma distribution hypothesis: Counting prime powers is a random process following a non-homogenous gamma distribution [7]. From a probability perspective, it is well-known that the distribution of primes is not random even if the sample space is restricted to odd integers. Therefore it seems significant that the heuristics used to arrive at the gamma hypothesis imply prime powers are random variables (over the odd integers). Whether or not this is the case requires the test of time. Regardless, the resulting probability model — a non-homogenous Poisson process — yields quite accurate average counting functions associated with the primes [8].

Therefore it is natural to generalize to prime $k$-tuples. The conjectured generalizations follow from an ansatz for the joint distribution of $k$-tuples of prime powers up to some
cut-off integer $x$ along a suitable ray in the pair-wise coprime $k$-lattice

$$P_{(k)}(n; x) := C_{(k)} \frac{(-1)^{k(n-1)+n}}{(n!)^k} \gamma(k(n-1)+1, -\log(x))$$

where $C_{(k)}$ is a constant and $\gamma$ is the lower incomplete gamma function. For admissible $k$-tuples, the ansatz leads to accurate counting functions if we assume $C_{(k)}$ is the appropriate singular series, i.e. the prime $k$-tuple constant.

But the enumeration is secondary. The primary goal is to extract information about prime $k$-tuple distributions, which means we need to discover pertinent $k$-tuple zeta functions implicitly defined by

$$\log(\zeta_{(k)}(s)) := \sum_{n=1}^{\infty} \frac{\Lambda_{(k)}(n)}{\log^k(n_{(k)}) n_{(k)}^s}$$

where the objects $\Lambda_{(k)}(n)$ and $n_{(k)}$ in the summand are defined later. The log-zeta functions are just what one would guess from the structure of the $k$-tuple analogs of the first Chebyshev function.

Unfortunately we haven’t found an explicit representation of $\zeta_{(k)}(s)$ that would allow the prime $k$-tuple issue to be settled. But in the final section we conjecture that it is given by a sum over the pair-wise coprime lattice denoted by $\mathfrak{H}_k$

$$\zeta_{(k)}(s) = \sum_{n_k \in \mathfrak{H}_k} \frac{1}{n_{(k)}^s} ,$$

and that the singular behavior of $\log^{(k)}(\zeta_{(k)}(s))$ yields the prime $k$-tuple constant.

This point bears repeating: Possessing exact counting functions (in terms of standard arithmetic functions) and a model probability distribution facilitates constructing $k$-tuple zeta functions and, subsequently, explicit integral representations of certain counting functions. This opens the possibility to attack the problem of prime $k$-tuple distributions using more-or-less elementary methods borrowed from the single prime case.

## 2 Prime $k$-tuple conjecture

According to [12], events along a directed graph can be modeled by a suitable gamma distribution. Here we are counting events associated with admissible prime $k$-tuples that occur along a ray $r_k$ in the pair-wise coprime $k$-lattice determined by an admissible set $\mathcal{H}_k = \{0, h_2, \ldots, h_k\}$. According to the gamma hypothesis, this is a non-homogenous

---

1. Following a standard argument, $C_{(k)}$ can be ‘derived’ from the probability model as; (i) the probability that $k$ integers are not divisible by some prime $p$, multiplied by (ii) the probability that the components of an admissible $k$-tuple are all coprime to $p$.

2. The zeta functions implicitly defined by the log-zeta functions will not be studied here, but they are of obvious interest.

3. We emphasize that our explicit formulae are left as integral representations. Since we do not determine the complex analytic properties of the $k$-tuple zeta functions, we cannot express the integrals in terms of residues.
Poisson process. There is no reason to expect the counting measure along \( r_k \) to coincide with the single prime case since the events correspond to coprime integer sets rather than simple integers, but we will assume the two are related by a normalization constant \( C_{(k)} \).

**Ansatz 2.1** Let \( H_k = \{0, h_2, \ldots, h_k\} \) be admissible. The expected number of joint events associated with counting admissible prime-power \( k \)-tuples is

\[
\overline{N}_{(k)}(x) := \sum_{n=1}^{\infty} P_{(k)}(n; x) := C_{(k)} \sum_{n=1}^{\infty} \frac{(-1)^{k(n-1)+n}}{(n!)^k} \gamma(k(n-1) + 1, -\log(x))
\]

where \( C_{(k)} \) is a normalization and \( \gamma \) is the lower incomplete gamma function.

Note that the series converges absolutely for finite \( x \);

\[
\lim_{n \to \infty} \frac{(n!)^2}{((n+1)!)^2} \left| \frac{\gamma(kn + 1, -\log(x))}{\gamma(k(n-1) + 1, -\log(x))} \right| = \lim_{n \to \infty} \frac{1}{(n+1)^2} \left| (-1)^k \log(x)^k \right| = 0 .
\]

The analysis in [8] suggests that

\[
\overline{N}_{(k)}(x) \approx \sum_{n \leq x} \lambda_{(k)}(n) - \sum_{n \nmid x} \lambda_{(k)}(n) .
\]

where

\[
\lambda_{(k)}(n) := \Lambda(n) \cdots \Lambda(n + h_k)/\log(n) \cdots \log(n + h_k) .
\]

Since \( \int_1^x |d\gamma(n, -\log(t))/dt| \, dt = \gamma(n, -\log(x)) \) and \( \overline{N}_{(k)}(x) \) converges absolutely, we write the summand as an integral and then interchange the sum and integral to get

\[
\overline{N}_{(k)}(x) \approx C_{(k)} \int_{x(k)}^{x} \log^{-k}(t) \, dt \text{ – hypergeometric terms}
\]

for all \( k \in \mathbb{N}_+ \) where

\[
x(k) := (x(x + h_2) \cdots (x + h_k))^{1/k} .
\]

Then, similar reasoning used in [8] implies that the hypergeometric terms are responsible for \( \sum_{n \mid x} \lambda_{(k)}(n) \) while

\[
\overline{J}_{(k)}(x) := C_{(k)} \int_2^{x(k)} \log^{-k}(t) \, dt \approx \sum_{n \leq x} \lambda_{(k)}(n)
\]

is the average \( k \)-tuple analog of Riemann’s counting function if we assume \( C_{(k)} \) to be the prime \( k \)-tuple constant. In particular, for prime doubles

\[
\overline{N}_{(2)}(x) = C_{(2)} \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2} \gamma(2n - 1, -\log(x)) , \quad x - 2 i \in \mathbb{N}_+ \\
=: \overline{J}_{(2)}(x) - \omega_{(2)}(x)
\]
where we will assume
\[ C(2) = C_2 \prod_{p \mid i} \frac{p - 1}{p - 2} \] (2.9)
with \( C_2 \) the twin prime constant. It is interesting to note that \( \omega(2)(x) \) is a Bessel function exhibiting oscillatory behavior.

Given this heuristic motivation, we conjecture:

**Conjecture 2.1** Given an admissible \( \mathcal{H}_k \), the average number of admissible prime \( k \)-tuples up to some cut-off integer \( x \) is
\[ \overline{\pi_{(k)}(x)} = \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \overline{J_{(k)}(x^{1/m})}. \] (2.10)

Compare this to the Hardy-Littlewood \( k \)-tuple conjecture.

Similarly, the gamma distribution hypothesis leads naturally to the average prime double Chebyshev function.

**Definition 2.1**
\[ \overline{Ch_{(2)}(x)} := C_{(k)} \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2} \gamma(2n, -\log(x)) \]
\[ =: \overline{\psi_{(2)}(x)} - \overline{d_{(2)}(x)}. \] (2.11)

This has obvious extensions to higher \( k \)-tuples.

**Conjecture 2.2**
\[ \overline{\theta_{(2)}(x)} = \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \overline{\psi_{(2)}(x^{1/m})}. \] (2.12)

Note that \( \overline{\psi_{(2)}(x)} = C(2) \text{Ei}(\log(x)) \) follows from (2.11) after repeating the arguments leading to (2.7). Hence \( \overline{\theta_{(2)}(x)} \sim C(2) (x / \log(x)) \) which, given the assumption that \( C(2) \) is the prime double constant, is consistent with the Hardy-Littlewood twin prime conjecture.

### 3 Explicit formulae

Having both exact and average summatory functions allows to deduce associated \( k \)-tuple zeta functions and subsequent explicit formulae. Here we will confine attention to prime doubles but indicate the generalization to higher \( k \)-tuples.

Define the prime double zeta function *implicitly* by

**Definition 3.1**
\[ \log \left( \zeta_{(2)}(s) \right) := \sum_{n=1}^{\infty} \frac{\lambda_{(2)}(n)}{n^{s/2}(n + 2i)^{s/2}} =: \sum_{n=1}^{\infty} \frac{\Lambda_{(2)}(n)}{\log^2(n_{(2)}) n_{(2)}^s}, \quad \Re(s) > 1. \] (3.1)
It follows that
\[
\log'(\zeta(2)(s)) = \frac{\zeta'(2)(s)}{\zeta(2)(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(2)(n)}{\log(n) n^{s}}.
\] (3.2)

Using this log-zeta function, along with the gamma hierarchy from [7] as a guide, we construct an explicit formula for \( \sum_{n=2}^{x} \frac{\Lambda(2)(n)}{\log(n)} \):

**Proposition 3.1** Put \( \tilde{x} = x + \epsilon \) with \( x \in \mathbb{N}_+ \) and \( 0 < \epsilon < 1 \). Let \( \sigma_{a} \) be the abscissa of absolute convergence of \( \frac{\lambda(2)(n) \log(n)}{n^{s/2}(n+2)^{s/2}} \). Then, for \( c > \sigma_{a} \),

\[
\psi(2)(x) = \lim_{\epsilon \to 0} \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(\tilde{x}(s))) \, d\log'(\zeta(2)(s)) , \quad c > \sigma_{a}
\]

\[
= \sum_{n=2}^{x} \frac{\Lambda(2)(n)}{\log(n)} . \quad (3.3)
\]

**Proof:** First integrate \((3.3)\) by parts. The boundary term does not contribute because i) a comparison test yields a finite \( \sigma_{a} \) (in fact \( \sigma_{a} = 1 \)) so \( \lim_{x \to \infty} |\log'(\zeta(2)(c + it))| < \infty \) for \( c > \sigma_{a} \); and ii) \( \lim_{t \to \infty} |\text{Ei}(\log(\tilde{x}(s)))| = 0 \) since

\[
\lim_{t \to \infty} |\text{Ei}(\log(x^{(c+it)}))| = \lim_{t \to \infty} \left| \frac{x^{(c+it)}}{(c+it) \log(x)} \left( 1 + O\left( \frac{1}{(c+it) \log(x)} \right) \right) \right| \\
\leq \frac{x}{\log(x)} \lim_{t \to \infty} \left| \frac{1}{(c+it)} \left( 1 + O\left( \frac{1}{(c+it)} \right) \right) \right| = 0 .
\] (3.4)

Next, following standard arguments, use the truncating integral

**Lemma 3.1**

\[
\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{x^{s}}{s} ds = \begin{cases} 
1 + O\left( \frac{x^{s}}{T \log(x)} \right) & x > 1 \\
O\left( \frac{x^{s}}{T \log(x)} \right) & 0 < x < 1
\end{cases} . \quad (3.5)
\]

**proof:** We include the proof for completeness. For \( x > 1 \) integrate over a rectangle with left edge \((L - iT, L + iT)\) such that \( L < c \). We have

\[
\lim_{L \to -\infty} \left| \int_{L-iT}^{L+iT} \frac{x^{s}}{s} ds \right| \leq \lim_{L \to -\infty} \int_{-T}^{T} \frac{x^{L}}{|L+it|} dt < \lim_{L \to -\infty} \frac{T x^{L}}{L} = 0 .
\] (3.6)

The top and bottom contribute

\[
\left| \int_{-\infty \pm iT}^{\pm iT} \frac{x^{s}}{s} ds \right| \leq \int_{-\infty}^{0} \frac{1}{(c-r) \pm iT} dr \leq x \int_{-\infty}^{0} \frac{1}{(c-r) \pm iT} dr \\
= x \int_{-\infty}^{0} \frac{1}{(c-r) \pm iT} dr < x \int_{-\infty}^{0} \frac{1}{T} dr \\
= \frac{x}{T \log(x)} .
\] (3.7)
Finally, the pole at $s = 0$ contributes $\text{Res} = 1$.

Now, for $x < 1$ integrate over the right edge $(R - iT, R + iT)$ with $c < R$. Then

$$
\lim_{R \to \infty} \left| \int_{R - iT}^{R + iT} \frac{x^s}{s} ds \right| \leq \lim_{R \to \infty} \int_{-T}^{T} e^{-R|\log(x)|} \frac{dt}{|R + it|} < \lim_{R \to \infty} \frac{Te^{-R|\log(x)|}}{R} = 0.
$$

(3.8)

The top and bottom contribute the same order as for $x > 1$, so the well-known lemma is established. □

Hence, for $c > \sigma_a$,

$$
- \lim_{\epsilon \to 0} \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \log' (\zeta(2)(s)) \frac{x^{s}}{s} ds
= \lim_{\epsilon \to 0} \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \sum_{n=1}^{\infty} \frac{\lambda(2)(n) \log(n(2)) x^{s}}{n^{s/2}(n+2i)^{s/2}} ds
= \lim_{\epsilon \to 0} \sum_{n \leq \lfloor \tilde{x} \rfloor} \frac{\Lambda(2)(n)}{\log(n(2))}
= \sum_{n=2}^{x} \frac{\Lambda(2)(n)}{\log(n(2))},
$$

(3.9)

where the third equality follows from the lemma. (Justifying the interchange of the sum and integral is straightforward, and interchange of the $T$-limit and sum is allowed because the summand contains $O(n^{-c})$ with $c > 1$.) □

Clearly this result only has teeth if one possesses an explicit representation of $\zeta(2)(s)$. But if a suitable representation of $\zeta(2)(s)$ can be found and it enjoys analytic properties similar to $\zeta(s)$, then we might expect

$$
\psi(2)(x) = \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(x^s(2))) d \log' (\zeta(2)(s))
= \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(x^s(2))) \log'' (\zeta(2)(s)) ds
$$

(3.10)

would lead to something like

$$
\psi(2)(x) \approx C(2) \text{Ei}(\log(x(2))) - C(2) \sum_{n(2)} \text{Ei}(\log(x^{\rho(2)})) + \text{small terms}
$$

(3.11)

where the sum would include conjugate pairs of nontrivial zeros of $\zeta(2)(s)$. And if nontrivial zeros of $\zeta(2)(s)$ are confined within its critical strip, then the same proof strategy used for the PNT would appear to apply to prime doubles. We could then bootstrap our
way to successive prime \(k\)-tuples. Evidently, if this scenario plays out the prime double constant \(C_2\) will have to come from \(\log''(\zeta(2)(s))\).

Remark that for higher \(k\)-tuples one should define

**Definition 3.2**

\[
\log \left( \zeta_k(s) \right) := \sum_{n=1}^{\infty} \frac{\lambda_k(n)}{n_k^s} = \sum_{n=1}^{\infty} \frac{\Lambda_k(n)}{\log^k(n_k) n_k^s}
\]

such that

\[
\log^{(k-1)'} \left( \zeta_k(s) \right) = (-1)^{k-1} \sum_{n=1}^{\infty} \frac{\lambda_k(n)}{n_k^s} \log^{k-1}(n_k)
\]

Here \(\Lambda_k(n) := \lambda_k(n) \log^k(n_k)\) and \(n_k := (n + h_2 \cdots (n + h_k))^{1/k}\).

Then to construct an explicit formula at level \(k\), consider

\[
\frac{(-1)^k}{C_k} \lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \varphi_k(x_k^s) d \log^{(k-1)'} \left( \zeta_k(s) \right)
\]

where

\[
\varphi_k(x) - \zeta_k(x) := C_k \sum_{n=1}^{\infty} (-1)^{(k-1) - 1} \left( \gamma(kn, - \log(x_k)) \right)
\]

so that

\[
\varphi_k(x) = C_k \text{Ei}(\log(x_k))
\]

Assuming integration by parts to be valid, Perron’s formula would apply yielding

\[
\varphi_k(x) = \sum_{n=2}^{x} \frac{\Lambda_k(n)}{\log(n_k)}
\]

Finally, assuming favorable analytic properties for \(\zeta_k(s)\) similar to \(\zeta(s)\), one would expect to find \(\varphi_k(x) \sim C_k x / \log(x)\).

### 4 Searching for \(\zeta_k(s)\)

The manipulations in the previous section point to a possible representation for \(\zeta_k(s)\) which we formulate as another ansatz. First note that

\[
\log \left( \zeta_2(s) \right) > \sum_{p_2 \in \mathbb{P}_2} \sum_{\omega, \omega' : \omega = \omega'} \frac{1}{\omega^s} \frac{1}{(p_2^2 + 2i)^{\omega s/2}} \quad \Re(s) > 1
\]

\[
> \sum_{p_2 \in \mathbb{P}_2} \sum_{\omega, \omega' : \omega = \omega'} \frac{1}{\omega^s} \frac{1}{(p_2^2 + 2i)^{\omega s/2}} \quad \Re(s) > 1
\]

\[
= \sum_{p_2 \in \mathbb{P}_2} \sum_{\omega^2} \frac{1}{\omega^2 p_2^{\omega^2}}
\]

\[
= - \sum_{p_2 \in \mathbb{P}_2} \log \left( 1 - p_2^{-s} \right)
\]

(4.1)
This suggests to define \( Z_{(k)}(s) := \prod_{p_k \in \Phi_k} \left( 1 - p_k^{-s} \right)^{-1} \) which motivates

**Ansatz 4.1** Let \( \mathfrak{N}_k \) be the pair-wise coprime \( k \)-lattice. Then

\[
\zeta_{(k)}(s) = \sum_{n_k \in \mathfrak{N}_k} \frac{1}{n_k^s} .
\]  

(4.2)

This in turn motivates

**Conjecture 4.1** \( \zeta_{(k)}(s) \) is meromorphic on \( \mathbb{C} \), and the singular part of \( (-1)^k \log^{(k)'}(\zeta_{(k)}(s)) \) is given by

\[
\frac{1}{s-1} \frac{(-1)^k}{2\pi i} \int \log^{(k)'}(\zeta_{(k)}(s)) \, ds = \frac{C_{(k)}}{s-1} .
\]  

(4.3)

This conjecture is equivalent to the Hardy-Littlewood prime \( k \)-tuple conjecture in the following sense. If we believe the gamma hypothesis, then being a sum of \( 1/n_{(k)}^s \) along a ray in the pair-wise coprime \( k \)-lattice strongly suggests \( \zeta_{(k)}(s) \) has a first order pole at \( s = 1 \) and there are no other poles, while its zeros are determined by conspiring projections of almost periodic exponentials.

Together with the explicit formula at level \( k \) the conjecture implies

\[
\varphi_{(k)}(x) = \lim_{T \to \infty} \frac{(-1)^k}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(x_{(k)}^s)) \log^{(k)'}(\zeta_{(k)}(s)) \, ds \bigg|_{s=1} = C_{(k)} \text{Ei}(\log(x_{(k)})) .
\]  

(4.4)

Equivalently,

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n \leq N} \Lambda_{(k)}(n) = C_{(k)} .
\]  

(4.5)

Of course, possessing poles and zeros of \( \zeta_{(k)}(s) \) would be tantamount to evaluating the exact summatory functions. And their singular part would presumably furnish the prime \( k \)-tuple constants. The goal would be to express the integral in Proposition 3.1 as a sum over \( k \)-tuple zeta residues in the usual way; which would presumably verify Conjecture 2.1 and validate the gamma hypothesis.

**References**

[1] G.H Hardy and J.E. Littlewood, Some problems of ‘Partitio numerorum’; III: On the expression of a number as a sum of primes, *Acta Math.*, **44**(1), (1923), 1–70.

[2] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.*, **167**, (2008), 481–547

[3] D.A. Goldston, J. Pintz, and C.Y. Yildirim, Primes in tuples I, *Ann. Math.*, **170**(2), (2009), 819–862.
[4] J. Korevaar, Prime pairs and the zeta function, *Jour. Approx. Theor.*, **158**, (2009), 69–96.

[5] Yitang Zhang, Bounded gaps between primes, *Ann. Math.*, **179**(3), (2014), 1121–1174.

[6] J. LaChapelle, Exact Summatory Functions for Prime $k$-tuples, math.NT/1406.5533 (2014).

[7] J. LaChapelle, Functional Integration on Constrained Function Spaces II: Applications, arXiv:math-ph/1405.0461 (2014).

[8] J. LaChapelle, Evidence of a Gamma Distribution for Prime Numbers, arXiv:math.NT/1307.1717 (2013).