Using Linnik's Identity to Approximate the Prime Counting Function with the Logarithmic Integral

Nathan McKenzie
11/26/2011

Summary:This paper will show that summing Linnik's identity from 2 to n and arranging terms in a certain way will immediately suggest a method for approximating the prime power counting function. It will then prove this approximation is the Logarithmic Integral. The relationship between the two will then be shown to provide one way of reasoning about the difference between the Logarithmic Integral and the prime power counting function. The paper will then use a similar approach to approximate Mertens function. This approximation will lead to a further approximation for Chebyshev's function $\psi(n)$ *, the dominant term for which is simply n. The paper will conclude by showing how this approximation leads to one way of reasoning about the difference between n and* $\psi(n)$ *.*

Note: This paper will use variables x,y,z for sums rather than the usual i,j,k. This is done to stress visually the symmetries between sums and their integral approximations, and the role changing between the two plays as the sole source of error terms.

1. Linnik's Identity

We begin with Linnik's identity which says that

$$\frac{\Lambda(n)}{\log n}=\sum_{k=1}^{\lfloor \log_2 n\rfloor}\frac{-1^{k+1}}{k}d_k'(n) \tag{1.1}$$

where $\Lambda(n)$ is the Von Mangoldt function, and $d_k'(n)$ is the strict number of divisors function such that

$$d_1'(n)=\begin{cases}0: n=1\\ 1: n\neq 1\end{cases} \quad d_k'(n)=\sum_{j|n}d_1'(j)\,d_{k-1}'\left(\frac{n}{j}\right) \tag{1.2}$$

Combinatorially, $d_k'(n)$ is the count of solutions to the expression $a_1\cdot a_2\cdot\ldots\cdot a_k=n$, where $a_1, a_2, \ldots, a_k \geq 2$ and are integers, and order matters. So, as an example $d_2'(6)=2$ because it has solutions of both $2\cdot 3$ and $3\cdot 2$.

Now, $\frac{\Lambda(n)}{\log n}$ is a function equal to 1 at primes, $\frac{1}{2}$ at primes squared, $\frac{1}{3}$ at primes cubed, and so on, and 0 otherwise. So, if we sum Linnik's identity from 2 to some value n, we have

$$\pi(n)+\frac{1}{2}\pi(n^{\frac{1}{2}})+\frac{1}{3}\pi(n^{\frac{1}{3}})+\ldots=\sum_{j=2}^{n}\sum_{k=1}\frac{-1^{k+1}}{k}d_k'(j)$$

where $\pi(n)$ is the prime counting function.

Combinatorially, the sum $\sum_{j=2}^{n}d_k'(j)$ is the count of solutions to the expression $a_1\cdot a_2\cdot\ldots\cdot a_k\leq n$, where $a_1, a_2, \ldots, a_k\geq 2$ and are all natural numbers, and where order matters, so we can rewrite those sums as

$$\sum_{j=1}^{n}d_1'(j)=\sum_{x=2}^{n}1$$

$$\sum_{j=1}^{n}d_2'(j)=\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1$$

$$\sum_{j=1}^{n}d_3'(j)=\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1 \tag{1.3}$$

and so on. We can use this to rewrite our prime counting identity as

$$\pi(n)+\frac{1}{2}\pi(n^{\frac{1}{2}})+\frac{1}{3}\pi(n^{\frac{1}{3}})+\ldots=$$
$$\sum_{x=2}^{n}1-\frac{1}{2}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1+\frac{1}{3}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1-\frac{1}{4}\ldots \tag{1.4}$$

2. Approximating the Prime Counting Function

The statement

$$\sum_{x=2}^{n}1-\frac{1}{2}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1+\frac{1}{3}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1-\frac{1}{4}\ldots$$

immediately suggests a method for approximating the prime counting function. Our approximation $A_\Pi(n)$ will replace the sums in our original function with the following closely related integrals

$$A_\Pi(n)=\int_1^n dx-\frac{1}{2}\int_1^n\int_1^{\frac{n}{x}} dy\,dx+\frac{1}{3}\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\frac{1}{4}...$$

We would like to see if we can transform our approximation A_Π into a more recognizable function. That process will occupy the rest of this section.

If we name each of these nested integrals as $A_k(n)$, so

$$A_\Pi(n)=\sum_{k=1}\frac{-1^{k+1}}{k}A_k(n) \tag{2.1}$$

and evaluate into closed form, we have

$$A_1(n)=\int_1^n dx=n-1$$

$$A_2(n)=\int_1^n\int_1^{\frac{n}{x}} dy\,dx=\int_1^n \frac{n}{x}-1\,dx=n\log n-n+1$$

$$A_3(n)=\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx=\int_1^n\int_1^{\frac{n}{x}}\frac{n}{xy}-1\,dy\,dx=$$
$$\int_1^n \frac{n}{x}\log\frac{n}{x}-\frac{n}{x}+1\,dx=\frac{1}{2}n\log^2 n-n\log n+n-1$$

and, more generally,

$$A_k=-1^k(-1+n\sum_{j=0}^{k-1}\frac{(-\log n)^j}{j!}) \tag{2.2}$$

If we specify that a new value m is equal to $-\log n$, then we can rewrite this as

$$A_k(n)=-1^k(-1+e^{-m}\sum_{j=0}^{k-1}\frac{m^j}{j!}) \tag{2.3}$$

We turn now to two representations of the incomplete gamma function,

$$\Gamma(k,u)=\int_u^\infty t^{k-1}e^{-t} \qquad \Gamma(k,u)=(k-1)!\,e^{-u}\sum_{j=0}^{k-1}\frac{u^j}{j!}$$

Setting both representations equal to each other and rearranging terms, we arrive at the following identity

$$e^{-u}\sum_{j=0}^{k-1}\frac{u^j}{j!}=\frac{1}{(k-1)!}\int_u^\infty t^{k-1}e^{-t}\,dt$$

which lets us rewrite (2.3) as

$$A_k(n)=-1^k(-1+\frac{1}{(k-1)!}\int_{-\log n}^\infty t^{k-1}e^{-t}\,dt)$$

Let's rewrite our leading term

$$A_k(n)=-1^k(-\frac{(k-1)!}{(k-1)!}+\frac{1}{(k-1)!}\int_{-\log n}^\infty t^{k-1}e^{-t}\,dt)$$

which we can rewrite as the gamma function $\Gamma(n)$ like so

$$A_k(n)=-1^k(-\frac{\Gamma(k)}{(k-1)!}+\frac{1}{(k-1)!}\int_{-\log n}^\infty t^{k-1}e^{-t}\,dt)$$

The gamma function has the following integral representation

$$\Gamma(x)=\int_0^\infty t^{k-1}e^{-t}\,dt$$

which we can substitute into our main expression

$$A_k(n)=$$
$$-1^k(-\frac{1}{(k-1)!}\int_0^\infty t^{k-1}e^{-t}\,dt+\frac{1}{(k-1)!}\int_{-\log n}^\infty t^{k-1}e^{-t}\,dt)$$

This simplifies to give us one final representation for $A_k(n)$,

$$A_k(n)=\frac{-1^{k+1}}{(k-1)!}\int_{-\log n}^0 t^{k-1}e^{-t}\,dt \tag{2.4}$$

Now let's insert this back into (2.1)

$$A_\Pi(n)=\sum_{k=1}\frac{-1^{k+1}}{k}\cdot\frac{-1^{k+1}}{(k-1)!}\int_{-\log n}^0 t^{k-1}e^{-t}\,dt$$

This simplifies to

$$A_\Pi(n)=\sum_{k=1}\frac{1}{k!}\int_{-\log n}^0 t^{k-1}e^{-t}\,dt$$

Swapping summation and integration, we have

$$A(n)=\int_{-\log n}^0(\sum_{k=1}^\infty\frac{t^{k-1}}{k!})e^{-t}\,dt$$

which becomes

$$A_\Pi(n)=\int_{-\log n}^0(\frac{e^t-1}{t})e^{-t}\,dt=\int_{-\log n}^0\frac{1-e^{-t}}{t}\,dt=$$
$$-\int_{\log n}^0\frac{1-e^t}{-t}\,dt=\int_0^{\log n}\frac{e^t-1}{t}\,dt$$

The exponential integral Ei(x) is given by

$$Ei(x) = \int_0^x \frac{e^t - 1}{t} dt + \log x + \gamma$$

where $\gamma = .577215...$, the Euler-Mascheroni Constant,so

$$A_\Pi(n) = \int_0^{\log n} \frac{e^t - 1}{t} dt = Ei(\log n) - \log\log n - \gamma$$

The logarithmic integral is $li(x) = Ei(\log x)$, so this means, finally, that

$$A_\Pi(n) =$$

$$\int_1^n dx - \frac{1}{2}\int_1^n \int_1^{\frac{n}{x}} dy\,dx + \frac{1}{3}\int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\,dy\,dx - \frac{1}{4}...$$

$$= li(n) - \log\log n - \gamma \quad (2.5)$$

(As a somewhat interesting aside, it is easier to evaluate the looser fitting

$$\int_0^n dx - \frac{1}{2}\int_1^n \int_0^{\frac{n}{x}} dy\,dx + \frac{1}{3}\int_1^n \int_1^{\frac{n}{x}} \int_0^{\frac{n}{xy}} dz\,dy\,dx - \frac{1}{4}...$$

which is just $\frac{n-1}{\log n}$.)

3. Error in this Approximation

So this provides one way of thinking about the connection between the logarithmic integral and the prime counting function. Linnik's identity let's us say that the prime counting function can be expressed like so,

$$\pi(n) + \frac{1}{2}\pi(n^{\frac{1}{2}}) + \frac{1}{3}\pi(n^{\frac{1}{3}}) + ... =$$
$$\sum_{x=2}^{n} 1 - \frac{1}{2}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor} 1 + \frac{1}{3}\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor} 1 - \frac{1}{4}...$$

and this identity can be approximated as

$$li(n) - \log\ \log n - \gamma =$$
$$\int_1^n dx - \frac{1}{2}\int_1^n \int_1^{\frac{n}{x}} dy\,dx + \frac{1}{3}\int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\,dy\,dx - \frac{1}{4}...$$

From this perspective, the error term in the Prime Number Theory can be reasoned about as the volumes under these various hyperbolic curves that aren't contained within the integer lattice volumes completely under each curve. Which is to say

$$li(n) - \pi(n) - \frac{1}{2}\pi(n^{\frac{1}{2}}) - \frac{1}{3}\pi(n^{\frac{1}{3}}) - ... =$$
$$-\ \log\ \log n - \gamma$$
$$-\frac{1}{2}\Big(\int_1^n \int_1^{\frac{n}{x}} dy\,dx - \sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor} 1\Big)$$
$$+\frac{1}{3}\Big(\int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\,dy\,dx - \sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor} 1\Big)$$
$$-\frac{1}{4}... \quad (3.1)$$

It might be interesting to connect this statement to Riemann's explicit prime counting formula

$$\pi(n) + \frac{1}{2}\pi(n^{\frac{1}{2}}) + \frac{1}{3}\pi(n^{\frac{1}{3}}) + ... =$$
$$li(n) - \sum_\rho li(n^\rho) - \log 2 + \int_n^\infty \frac{dt}{t(t^2-1)\log t}$$

where ρ are the zeroes of the Riemann Zeta function.

We can obviously restate this as

$$li(n) - \pi(n) - \frac{1}{2}\pi(n^{\frac{1}{2}}) - \frac{1}{3}\pi(n^{\frac{1}{3}}) - ... =$$
$$\sum_\rho li(n^\rho) + \log 2 - \int_n^\infty \frac{dt}{t(t^2-1)\log t}$$

If we define the following ancillary function

$$E(n) = \log 2 - \int_x^\infty \frac{dt}{t(t^2-1)\log t} + \log\log n + \gamma$$

then, comparing this paper's results to Riemann's function, we can say that

$$\sum_\rho li(x^\rho) + E(x) =$$
$$-\frac{1}{2}\Big(\int_1^n \int_1^{\frac{n}{x}} dy\,dx - \sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor} 1\Big)$$
$$+\frac{1}{3}\Big(\int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} dz\,dy\,dx - \sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor} 1\Big)$$
$$-\frac{1}{4}\Big(\int_1^n \int_1^{\frac{n}{x}} \int_1^{\frac{n}{xy}} \int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx - \sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor} 1\Big)$$
$$+\frac{1}{5}... \quad (3.2)$$

Given that *E(n)* is *O(log log n),* it can be largely ignored, leaving us with the observation, left without any further comment, that the Zeta Zeroes perfectly describe these non-lattice differences.

With a bit of work, this equation can be further transformed into

$$\sum_{\rho} li(x^{\rho})+E(x)=$$

$$-\frac{1}{2}\left(\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\sum_{x=2}^{n}\frac{n}{x}-1+\sum_{x=2}^{n}\left\{\frac{n}{x}\right\}\right)$$
$$+\frac{1}{3}\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\frac{n}{xy}-1+\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\left\{\frac{n}{xy}\right\}\right)$$
$$-\frac{1}{4}\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\frac{n}{xyz}-1+\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\left\{\frac{n}{xyz}\right\}\right)$$
$$+\frac{1}{5}\ldots \tag{3.3}$$

where {n} is the fractional part function. This representation is noteworthy in that all of the familiar high frequency / discontinuous aspects of the prime counting function seem to be captured by the fractional part sums; the other nested sums represent fairly smooth, predictable values.

4. Mertens Function

There is the following identity for the Möbius function in terms of the strict count of divisor functions from (1.2), derived in a fashion similar to Linnik's identity

$$\mu(n)=\sum_{k=0} -1^k d_k{}'(n)$$

If we sum this identity from 2 to n, we have

$$M(n)=\sum_{k=0} -1^k \sum_{j=2}^{n} d_k{}'(j)$$

where $M(n)$ is the Mertens function.

Following similar techniques from our first section that led from (1.3) to (1.4), this means we can write Mertens function as

$$M(n)=1-\sum_{x=2}^{n}1+\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1+\ldots \tag{4.1}$$

This once again immediately suggests an approximation for Mertens function of the form

$$A_M(n)=1-\int_1^n dx+\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx+\ldots \tag{4.2}$$

These nested integrals are the same integrals of the form $A_k(n)$ that we saw before in (2.2); in fact,

$$A_M(n)=1+\sum_{k=1} -1^k A_k(n)$$

Substituting in our final integral representation for $A_k(n)$ from (2.4), we have

$$A_M(n)=1+\sum_{k=1} -1^k \frac{-1^{k+1}}{(k-1)!}\int_{-\log n}^{0} t^{k-1}e^{-t}\,dt$$

which simplifies to

$$A_M(n)=1-\sum_{k=1}\frac{1}{(k-1)!}\int_{-\log n}^{0} t^{k-1}e^{-t}\,dt$$

Rearranging summation and integration, we have

$$A_M(n)=1-\int_{-\log n}^{0}\left(\sum_{k=1}\frac{t^{k-1}}{(k-1)!}\right)e^{-t}\,dt$$

which leads to

$$A_M(n)=1-\int_{-\log n}^{0} e^{t}e^{-t}\,dt=1-\int_{-\log n}^{0} dt=1-\log n$$

So,

$$1-\log n=1-\int_1^n dx+\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx+\ldots \tag{4.3}$$

Thus, the difference between Mertens function and our approximation can be expressed as

$$M(n)-\log n+1=$$
$$-\left(\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1\right)$$
$$+\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1\right)$$
$$-\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor}1\right)$$
$$+\ldots \tag{4.4}$$

which, given the smallness and smoothness of log n, means that the non-lattice volumes are largely just the inverse of Mertens function itself. As with the prime counting function, this can be further developed into

$$M(n)-\log n+1=$$
$$-\left(\int_1^n\int_1^{\frac{n}{x}} dy\,dx-\sum_{x=2}^{n}\frac{n}{x}-1+\sum_{x=2}^{n}\{\frac{n}{x}\}\right)$$
$$+\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\frac{n}{xy}-1+\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\{\frac{n}{xy}\}\right)$$
$$-\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\frac{n}{xyz}-1+\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\{\frac{n}{xyz}\}\right)$$
$$+\ldots$$

where all of the discontinuities and high frequency change are contained in the fractional part sums.

5. Chebyshev's Psi Function

Mertens gave the following formula for Chebyshev's function $\psi(n)$

$$\psi(n)=\sum_{j=2}^{n}\log j\, M\left(\frac{n}{j}\right) \tag{5.1}$$

Now, we know from the previous section that we can write Merten's function as

$$M(n)=1-\sum_{y=2}^{n}1+\sum_{y=2}^{n}\sum_{z=2}^{\lfloor\frac{n}{y}\rfloor}1-\sum_{y=2}^{n}\sum_{z=2}^{\lfloor\frac{n}{y}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{yz}\rfloor}1+\ldots$$

so we can rewrite $\psi(n)$ as

$$\psi(n)=\sum_{x=2}^{n}\log x\left(1-\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}1+\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}1-\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor}1+\ldots\right) \tag{5.2}$$

This expression for $\psi(n)$ immediately suggests its own approximation, namely

$$A_\psi(n)=$$
$$\int_1^n\log x\left(1-\int_1^{\frac{n}{x}}dy+\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}dz\,dy-\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}}dw\,dz\,dy+\ldots\right)dx \tag{5.3}$$

The inner sum of integrals is our approximation for $A_M(n)$ (4.2) with n taking the value $\frac{n}{x}$. We know from (4.3) that this approximation is just $1-\log\frac{n}{x}$, so we can replace it with

$$A_\psi(n)=\int_1^n\log x\left(1-\log\frac{n}{x}\right)dx$$

which we can rewrite as

$$A_\psi(n)=\int_1^n\log x\cdot(1-\log n+\log x)\,dx$$

$$A_\psi(n)=\int_1^n\log x\,dx-\int_1^n\log x\log n\,dx+\int_1^n\log x\log x\,dx$$

$$A_\psi(n)=(n\log n-n+1)-(n\log^2 n-n\log n+\log n)$$
$$+(n\log^2 n-2n\log n+2n-2)$$

$$A_\psi(n)=n-\log n-1$$

So

$$n-\log n-1=$$
$$\int_1^n\log x\left(1-\int_1^{\frac{n}{x}}dy+\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}dz\,dy-\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}}dw\,dz\,dy+\ldots\right)dx \tag{5.4}$$

For what follows, it is useful to multiply the log value through for both our expression for $\psi(n)$ and its approximation. So, we have

$$\psi(n)=\left(\sum_{x=2}^{n}\log x\right)-\left(\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\log y\right)$$
$$+\left(\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\log z\right)-\left(\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor}\log w\right)+\ldots \tag{5.5}$$

and

$$n-\log n-1=\left(\int_1^n\log x\,dx\right)-\left(\int_1^n\int_1^{\frac{n}{x}}\log y\,dy\,dx\right)$$
$$+\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\log z\,dz\,dy\,dx\right)-\left(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}}\log w\,dw\,dz\,dy\,dx\right)+\ldots \tag{5.6}$$

Our real interest here is relating $\psi(n)$ to *n.* Thus, if we subtract our representation for $\psi(n)$ from its approximation, we have

$$
\begin{aligned}
&n-\psi(n)= \\
&-\log n-1 \\
&+\Big(\int_1^n \log x\,dx-\sum_{x=2}^{n}\log x\Big) \\
&-\Big(\int_1^n\int_1^{\frac{n}{x}} \log y\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\log y\Big) \\
&+\Big(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} \log z\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\log z\Big) \\
&-\Big(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} \log w\,dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor}\log w\Big) \\
&+\ldots
\end{aligned}
\tag{5.7}
$$

The function $\psi(n)$ can be written in terms of the zeroes of the Riemann Zeta function as

$$\psi(n)=n-\sum_{\rho}\frac{n^{\rho}}{\rho}-\frac{\zeta'(0)}{\zeta(0)}-\frac{1}{2}\log(1-n^{-2})$$

Obviously, we can rewrite this as

$$n-\psi(n)=\sum_{\rho}\frac{n^{\rho}}{\rho}+\frac{\zeta'(0)}{\zeta(0)}+\frac{1}{2}\log(1-n^{-2})$$

If we compare this result to (5.7), and we define the following O(log n) function

$$E(n)=\log 2\pi+\frac{1}{2}\log(1-n^{-2})+\log n+1$$

then that means

$$
\begin{aligned}
&\sum_{\rho}\frac{n^{\rho}}{\rho}+E(n)= \\
&\Big(\int_1^n \log x\,dx-\sum_{x=2}^{n}\log x\Big) \\
&-\Big(\int_1^n\int_1^{\frac{n}{x}} \log y\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\log y\Big) \\
&+\Big(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}} \log z\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\log z\Big) \\
&-\Big(\int_1^n\int_1^{\frac{n}{x}}\int_1^{\frac{n}{xy}}\int_1^{\frac{n}{xyz}} \log w\,dw\,dz\,dy\,dx-\sum_{x=2}^{n}\sum_{y=2}^{\lfloor\frac{n}{x}\rfloor}\sum_{z=2}^{\lfloor\frac{n}{xy}\rfloor}\sum_{w=2}^{\lfloor\frac{n}{xyz}\rfloor}\log w\Big) \\
&+\ldots
\end{aligned}
\tag{5.8}
$$

6. Odds and Ends

Although this paper won't cover the topic, a wide range of noteworthy arithmetic summatory functions can be expressed in a fashion similar to (5.1) and (5.2). Consequently, the approach taken in section 5 can also be applied to them.

Special thanks to J. Mangaldan from math.stackexchange.com for providing the explicit method for transforming (2.2) into (2.5); I knew (2.2) seemed to be the Logarithmic Integral, but I wasn't quite sure why.

Special thanks to Professor Mark Coffey, who provided helpful feedback and a few notes about (3.3) at http://www.icecreambreakfast.com/primecount/note%20on%20McKenzie%20sums.pdf , which was a response to a previous threadbare write up of mine at http://www.icecreambreakfast.com/primecount/mckenzie_math.old.pdf