The Prime Number Theorem and the Series $\sum_{n=1}^{\infty} \frac{\mu(n)}{n}$

R. C. Daileda

October 16, 2008

Throughout what follows, we let Λ denote the Mangoldt function

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^a, \ p \text{ prime}, \ a \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

and μ the familiar Möbius function. We also introduce the summatory functions

$$\begin{aligned} \psi(x) &= \sum_{n \leq x} \Lambda(n) \\ m(x) &= \sum_{n \leq x} \frac{\mu(n)}{n}. \end{aligned}$$

It is well known that the Prime Number Theorem is equivalent to the statement that

$$\lim_{x \to \infty} \frac{\psi(x)}{x} = 1.$$

It is the purpose of this note to prove that this asymptotic result further implies that $\lim_{x \to \infty} m(x) = 0$. We now state this formally.

Theorem 1. *If* $\lim_{x \to \infty} \frac{\psi(x)}{x} = 1$ *then*

$$\lim_{x \to \infty} m(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0.$$

The proof requires two preliminaries, which we state here. The first can be found, for example, in *Introduction to Analytic Number Theory* by Tom Apostol and the second is a standard result.

Lemma 1. *For all* $x > 0$, $|m(x)| \leq 1$.

Lemma 2 (Partial Summation)**.** *Let* $\{a_n\}$ *and* $\{b_n\}$ *be sequences of complex numbers and set*

$$A(x) = \sum_{n \leq x} a_n$$

for $x \geq 1$ *and* $A(x) = 0$ *for* $x < 1$*. If* $m \geq 0$ *and* $N \geq 1$ *are integers then*

$$\sum_{n=m+1}^{N} a_n b_n = \sum_{n=m+1}^{N-1} A(n) \left(b_n - b_{n+1} \right) + A(N) b_N - A(m) b_{m+1}.$$

Taking $m = 0$ in the lemma and manipulating the resulting expression yields

$$\sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N} A(n)\left(b_n - b_{n+1}\right) + A(N) b_{N+1}. \tag{1}$$

It is in this form that we will need the lemma later.

Proof of Theorem 1. We begin by observing that if $I(n) = [1/n]$ and $u(n) = 1$ for all n then the convolution identity $I = \mu * u$ implies

$$I(n) = \frac{I(n)}{n} = \sum_{d|n} \frac{\mu(d)}{n}$$

so that

$$\begin{aligned} 1 &= \sum_{n \le x} \sum_{d|n} \frac{\mu(d)}{n} \\ &= \sum_{d \le x} \frac{\mu(d)}{d} \sum_{k \le x/d} \frac{1}{k} \\ &= \sum_{d \le x} \frac{\mu(d)}{d} \left(\log \frac{x}{d} + \gamma + O\left(\frac{d}{x}\right)\right) \\ &= m(x) \log x - \sum_{d \le x} \frac{\mu(d) \log d}{d} + O(1) \end{aligned}$$

where γ is Euler's constant and we have appealed to Lemma 1 to produce the error term. Thus

$$m(x) = \frac{1}{\log x} \sum_{d \le x} \frac{\mu(d) \log d}{d} + O\left(\frac{1}{\log x}\right)$$

and to establish the theorem it is therefore sufficient to prove that

$$\lim_{x \to \infty} \frac{1}{\log x} \sum_{d \le x} \frac{\mu(d) \log d}{d} = 0. \tag{2}$$

At this point let's pause for some motivation for what comes next. In order to prove (2) we would like to express the arithmetic function $\mu \log$ as a convolution and apply the usual interchange of order of summation trick. There are certainly many ways to do this, but we seek an identity that relates to our hypothesis about ψ. What we have assumed is that

$$0 = \lim_{x \to \infty} \frac{\psi(x)}{x} - 1 = \lim_{x \to \infty} \frac{1}{x} \sum_{n \le x} \left(\Lambda(n) - 1\right),$$

i.e. that we have an asymptotic formula for the average of $\Lambda - u$. This suggests that we try to express $\mu \log$ in terms of $\Lambda - u$. This is quite easy, for from the identity $\Lambda = -u * \mu \log$ we have

$$\begin{aligned} \mu \log &= -\mu * \Lambda \\ &= -\mu * \Lambda + I - I \\ &= -\mu * \Lambda + \mu * u - I \\ &= \mu * (u - \Lambda) - I. \end{aligned}$$

We therefore have

$$\begin{aligned}
\sum_{n\le x}\frac{\mu(n)\log n}{n} &= \sum_{n\le x}\sum_{d|n}\frac{(1-\Lambda(d))\,\mu\left(\frac{n}{d}\right)}{n}-1\\
&= \sum_{d\le x}(1-\Lambda(d))\,\frac{1}{d}\sum_{k\le x/d}\frac{\mu(k)}{k}-1\\
&= \sum_{d\le x}(1-\Lambda(d))\,\frac{1}{d}m\left(\frac{x}{d}\right)-1. \qquad (3)
\end{aligned}$$

If we take $N=[x]$, $a_n=1-\Lambda(n)$ and $b_n=\frac{1}{n}m\left(\frac{x}{n}\right)$ then by (1) the expression (3) is equal to

$$\sum_{d=1}^{N}(d-\psi(d))\left(\frac{1}{d}m\left(\frac{x}{d}\right)-\frac{1}{d+1}m\left(\frac{x}{d+1}\right)\right)+(N-\psi(N))\,\frac{1}{N+1}m\left(\frac{x}{N+1}\right)-1. \qquad (4)$$

Since

$$\begin{aligned}
\left|\frac{1}{d}m\left(\frac{x}{d}\right)-\frac{1}{d+1}m\left(\frac{x}{d+1}\right)\right| &= \left|\frac{1}{d+1}\left(m\left(\frac{x}{d}\right)-m\left(\frac{x}{d+1}\right)\right)+\frac{1}{d(d+1)}m\left(\frac{x}{d}\right)\right|\\
&\le \frac{1}{d}\sum_{x/(d+1)<m\le x/d}\frac{1}{m}+\frac{1}{d^2},
\end{aligned}$$

where again we have appealed to Lemma 1, we find that the absolute value of (4) is bounded by

$$\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\sum_{x/(d+1)<m\le x/d}\frac{1}{m}+\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\frac{1}{d}+\left|1-\frac{\psi(N)}{N}\right|+1$$

and hence this provides an upper bound for the sum $\sum_{n\le x}\frac{\mu(n)\log n}{n}$. Recalling our hypothesis and our goal, we find now that it suffices to prove that

$$\lim_{x\to\infty}\frac{1}{\log x}\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\sum_{x/(d+1)<m\le x/d}\frac{1}{m}=\lim_{x\to\infty}\frac{1}{\log x}\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\frac{1}{d}=0. \qquad (5)$$

Let $\epsilon>0$. Since $\lim_{x\to\infty}\frac{\psi(x)}{x}=1$ there is a constant C so that $\left|1-\frac{\psi(x)}{x}\right|\le C$ for all x and a $y>0$ so that $\left|1-\frac{\psi(x)}{x}\right|\le\epsilon$ for all $x\ge y$. Thus if $x>y+1$

$$\begin{aligned}
\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\sum_{x/(d+1)<m\le x/d}\frac{1}{m} &\le C\sum_{d\le y}\sum_{x/(d+1)<m\le x/d}\frac{1}{m}+\epsilon\sum_{d\le x}\sum_{x/(d+1)<m\le x/d}\frac{1}{m}\\
&\le C\sum_{x/(y+1)<m\le x}\frac{1}{m}+\epsilon\sum_{m\le x}\frac{1}{m}. \qquad (6)
\end{aligned}$$

Since $\sum_{m\le z}\frac{1}{m}=\log z+\gamma+O(1/z)$ (in which the implied constant can be taken to be 2) for all $z>0$ we have

$$\sum_{x/(y+1)<m\le x}\frac{1}{m}=\log(y+1)+O(1)\le 2\log(y+1)$$

as long as y is large enough (which we can have arranged earlier) so that (6) is less than or equal to

$$2C\log(y+1) + 2\epsilon\log x.$$

Dividing by $\log x$ we thus have

$$\frac{1}{\log x}\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\sum_{x/(d+1)<m\le x/d}\frac{1}{m}\le\frac{2C\log(y+1)}{\log x}+2\epsilon.$$

for all $x > y+1$. If we now choose $x_0 > y$ so that $\dfrac{2C\log(y+1)}{\log x} < \epsilon$ for all $x \geq x_0$ then we find that

$$\frac{1}{\log x}\sum_{d=1}^{N}\left|1-\frac{\psi(d)}{d}\right|\sum_{x/(d+1)<m\le x/d}\frac{1}{m}<3\epsilon$$

for $x \geq x_0$. Since $\epsilon > 0$ was arbitrary, this proves the first limit of (5) is 0. That the second is also zero is proven in an entirely analogous fashion and is left to the reader. As noted above, this completes the proof of the theorem.

□