Generalizations

6 August 2006

Philippe,

I worked through the last part of the paper in greater detail, and have come to the conclusion that the theorem is faulty. I started by attempting to generalize the proof to other series, in particular the Liouville and the Euler totient series. All of these have in common that they have $\zeta(s)$ in the denominator, implying poles on the critical line. The asymptotic behaviour of the finite differences for each is quite different – and this is what leads me to conclude that we cannot deduce the location of the poles based on the asymptotic behaviour, or that, at least, the part of the proof that attempts to do this is not correct.

Anyway, here's what follows (its quite long): a review of the definition of the series. A numeric exploration of the series (a rather surprising/interesting behaviour is seen for the Mobius series!). A saddle-point analysis to obtain asymptotic behaviour. The saddle point analysis appears to be consistent with the numeric work. An attempted generalization of the proof, which I'm unable to complete.

I wrote most of the below before I realized that I couldn't make the RH equivalence proof work, so there may be some inconsistent comments below.

Review of Dirichlet Series

Much of the proof depends on having $\zeta(s)$ in the denominator, and there are many Dirichlet series that achieve this. The canonical one, involving the Mobius function is

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

But there is also one for the Euler Phi function:

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}$$

The Liouville function:

$$\sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}$$

The von Mangoldt function

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$$

There are a dozen others that can be readily found in introductory textbooks and/or the web.

One may construct in a very straightforward way the Dirichlet series for $1/\zeta(s-a)$ for any complex a, as well as $1/\zeta(2s-a)$. These can be constructed by a fairly trivial application of Dirichlet convolution, together with the Mobius inversion formula. In short, one has an old, general theorem that

$$\sum_{n=1}^{\infty} \frac{(f*g)(n)}{n^s} = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \sum_{m=1}^{\infty} \frac{g(m)}{m^s}$$

where $f*g$ is the Dirichlet convolution of f and g:

$$(f*g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right)$$

Since Dirichlet convolution is invertible whenever $f(1) \neq 1$ (and/or $g(1) \neq 1$), one may multiply and divide Dirichlet series with impunity, more or less.

I haven't yet seen a way of building $1/\zeta(\alpha s+\beta)$, or more complicated expressions.

As regards to the proof, perhaps the most important generalization is that for the Dirichlet L-functions, since these are the ones for which the GRH applies. Specifically, one has

$$\sum_{n=1}^{\infty} \frac{\mu(n)\chi(n)}{n^s} = \frac{1}{L(s,\chi)}$$

where χ is the Dirichlet character.

Numeric Exploration

This section provides a numeric exploration of the finite differences of the various Dirichlet series given above. Consider first

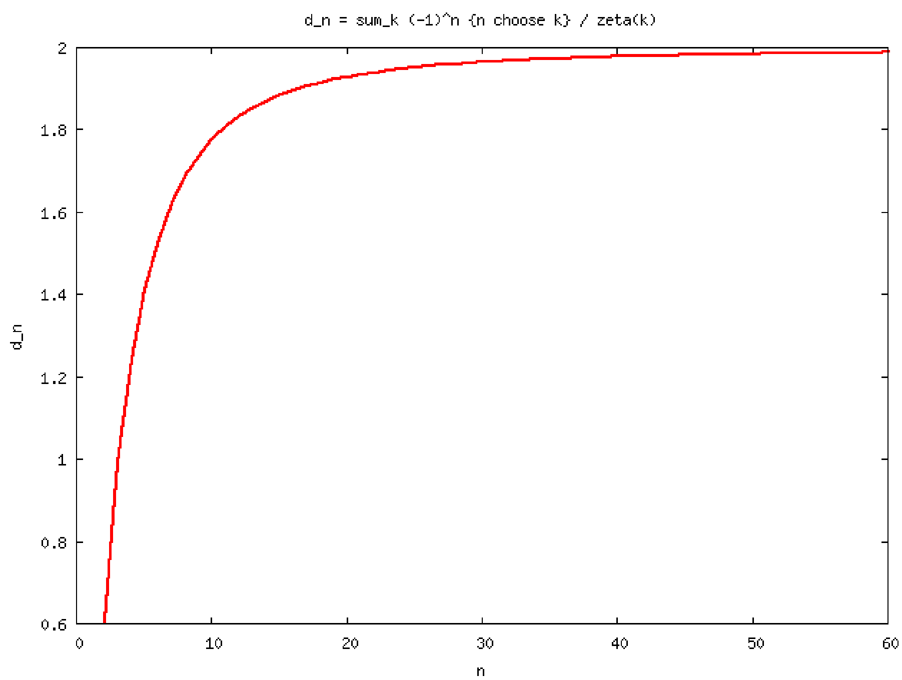
$$d_n = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \frac{1}{\zeta(k)}$$

By examining d_n in the range of $2 \leq n \leq 1000$, one would be tempted to incorrectly conclude that $\lim_{n\to\infty} d_n = 2$; this behaviour is shown in the graph below. However, by exploring the range $1000 \leq n \leq 50000$, one discovers that d_n is oscillatory. The oscillations seem bounded, and thus the numeric work suggests that

$$d_n = O(\log n)$$

or possibly better, but presumably not $O(1)$.

Figure 1: d_n for smaller n

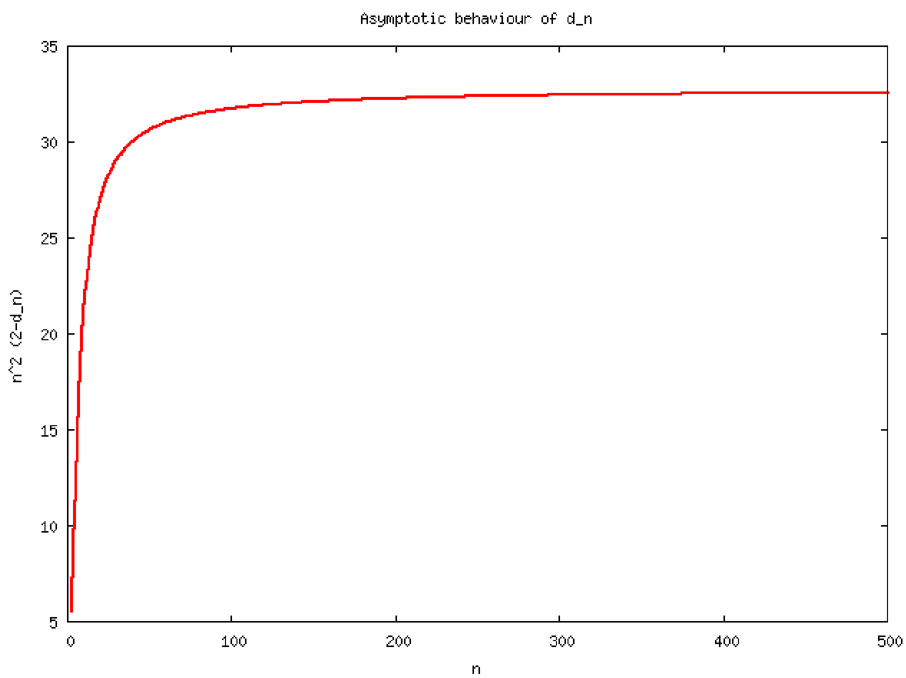


The graphic above shows

$$d_n = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \frac{1}{\zeta(k)}$$

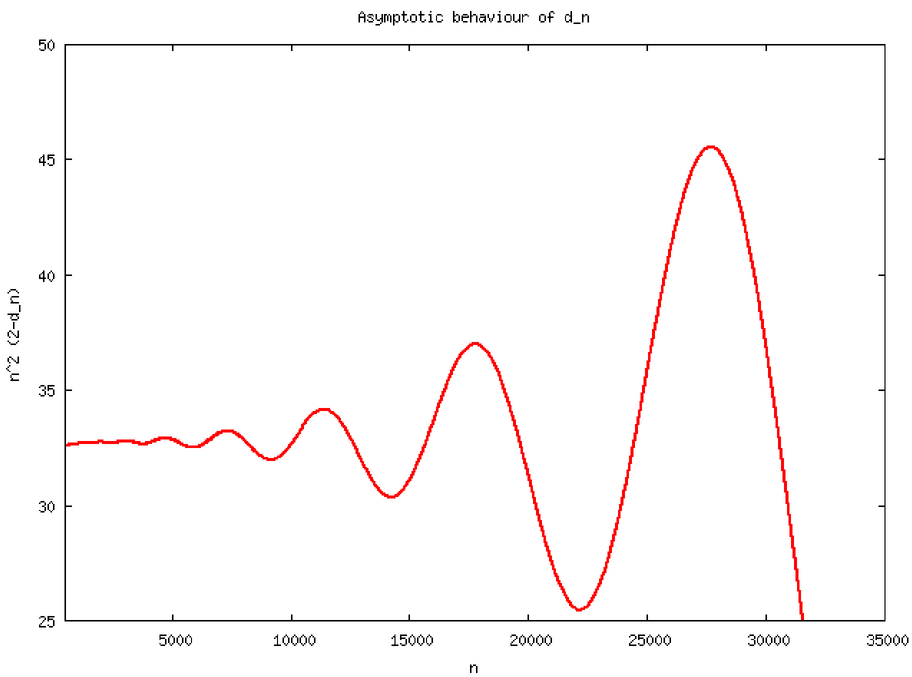
in the range of $2 \leq n \leq 60$. It strongly suggests an asymptotic approach to $d_n \to 2$, which can be seen to be incorrect if larger n are explored.

Figure 2: Asymptotic behaviour of d_n



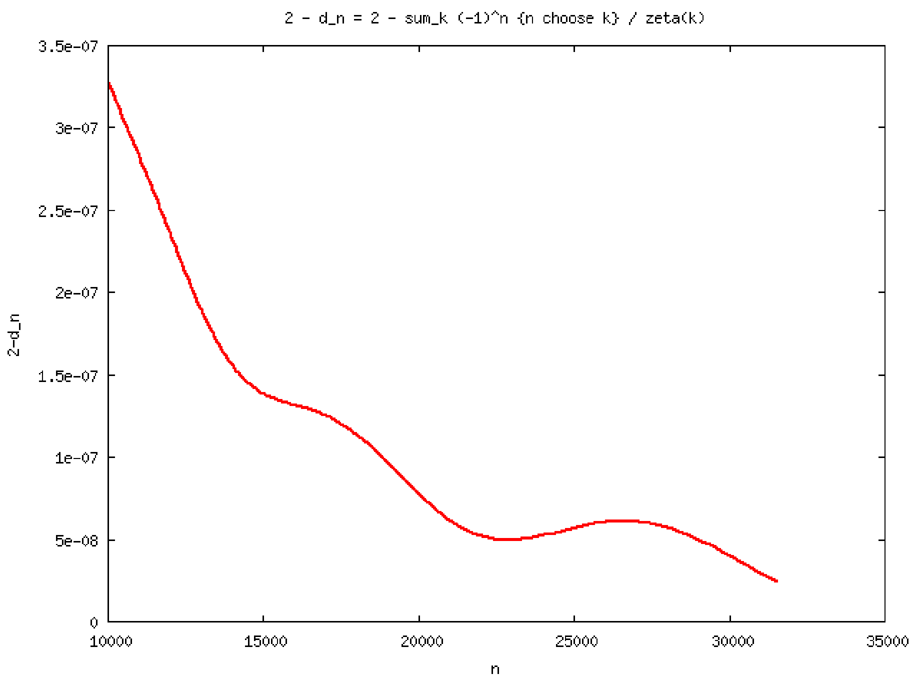
The above figure shows a graph of $n^2(2-d_n)$ in the range of $2 \leq n \leq 500$. As with the previous graphic, it strongly but incorrectly suggests that $d_n \to 2$in the limit of large n. That this is not the case can be discovered by pursuing larger n.

Figure 3: Asymptotic behaviour of d_n



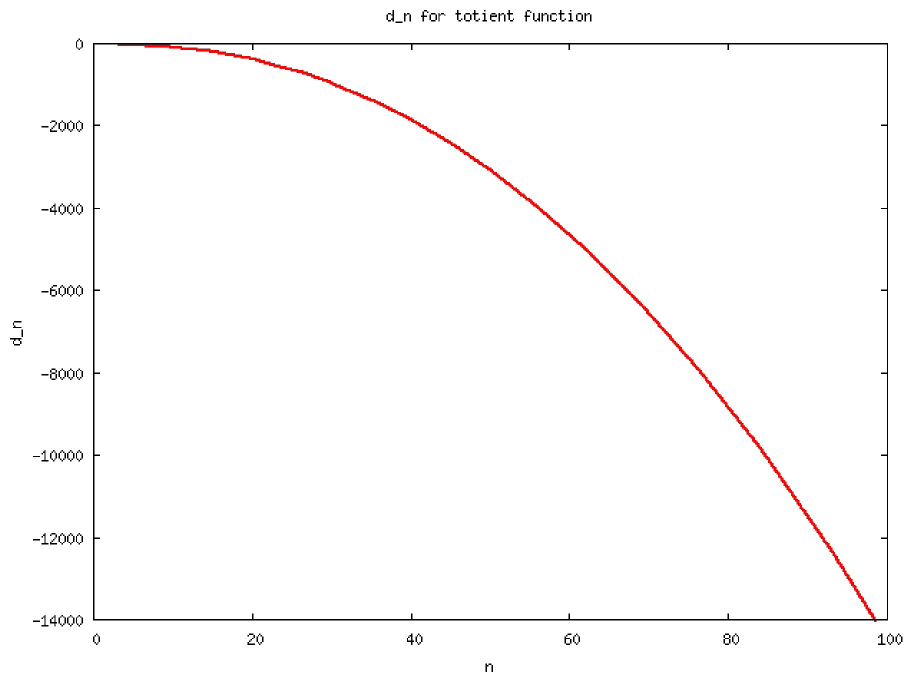
This graphic charts the value of $n^2(2-d_n)$ in the range of $500 \leq n \leq 30000$. Rather than approaching a limit for large n, there are a series of oscillations that seem to grow ever larger. The amplitude of these oscillations appears to be of $O\left(n^2\right)$ or possibly $O\left(n^2 \log n\right)$. In either case, such oscillations indicate that $d_n$cannot be approaching a constant for large n, instead suggesting that d_n is itself oscillatory for large n, as the next graphic shows.

Figure 4: $d_n$for large n



This figure shows a plot of $2-d_n$ for the range $10^4 \leq n \leq 3\times 10^4$. A nascent oscillatory behaviour can be seen.

Figure 5: Graph of d_n^{φ}



A graph of d_n^{φ} for $3 \le n \le 100$ shows rapidly increasing behaviour.

The finite differences for the other series appear to show a similar pattern, except that the scale of the leading order is different. Consider, for example,

$$d_n^{\varphi} = \sum_{k=3}^{n} (-1)^k \binom{n}{k} \frac{\zeta(k-1)}{\zeta(k)}$$

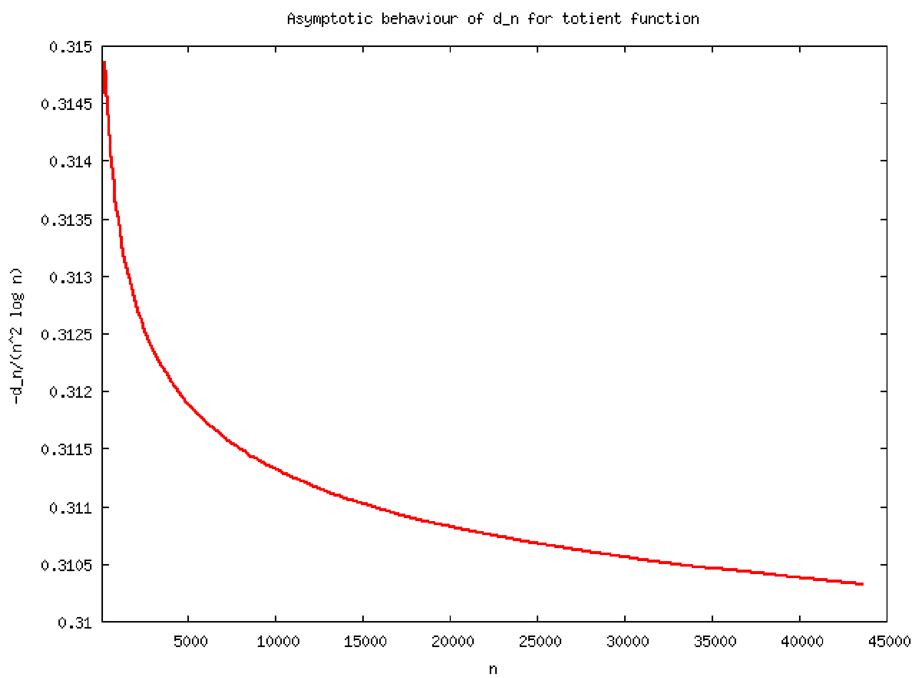
with the superscript φ indicating that the corresponding Dirichlet series involves the totient function φ. For smaller values of n, numeric analysis suggests that $d_n^{\varphi} = O\left(n^2 \log n\right)$, as the following graphics illustrate.

Let

$$d_n^{\lambda} = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \frac{\zeta(2k)}{\zeta(k)}$$

be the series corresponding to the Liouville function.

Figure 6: Asymptotic behaviour of d_n^{φ}

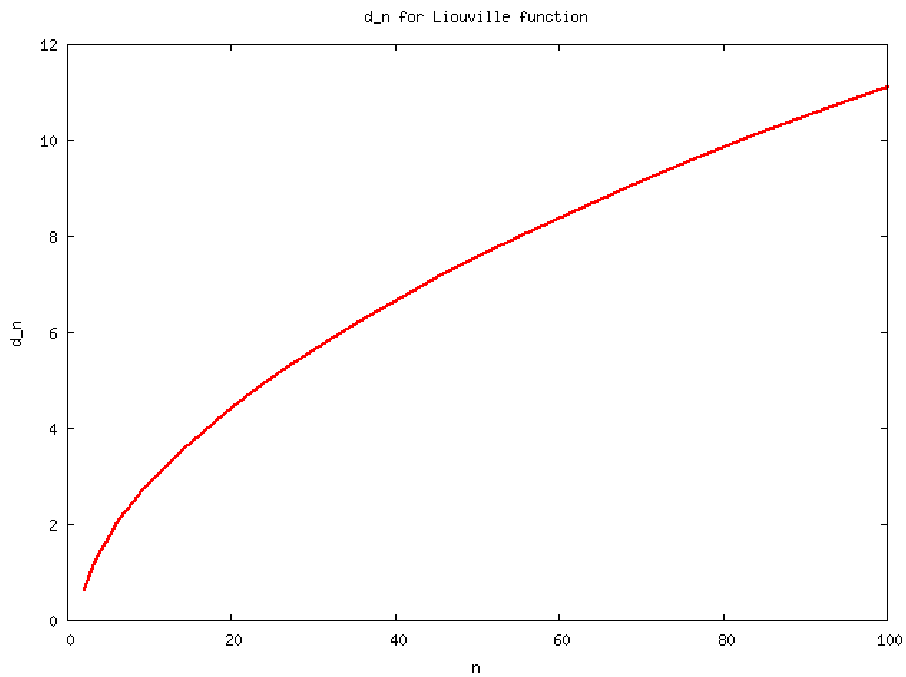


This graphic shows $-d_n^{\varphi}/n^2 \log n$ for $100 \leq n \leq 16000$, suggesting a form for the asymptotic behaviour of d_n^{φ}. This figure has a logarithmic shape, suggesting that

$$d_n^{\varphi} = A\,(1 + B \log n)\, n^2 \log n + O\left(n^2 \log^3 n\right)$$

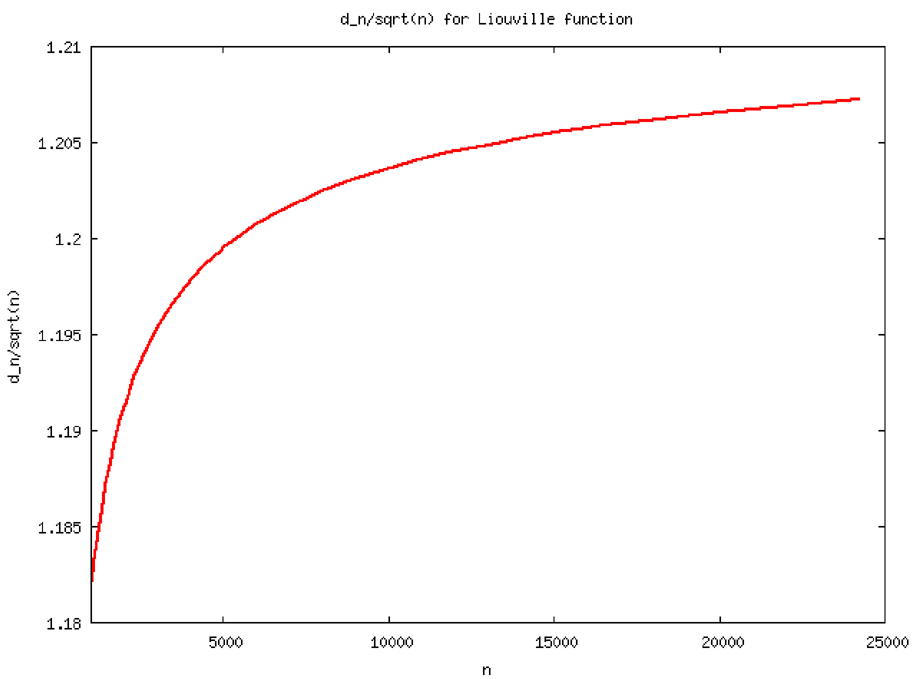
for some constants A, B. Numerically, it seems that $A \approx 0.316$, $B \approx 7 \times 10^{-4}$.

Figure 7: d_n^λ for the Liouville function



This figure shows the basic behavior for the finite differences d_n^λ corresponding to the Liouville function. Numeric work suggests that the divergence is similar to but stronger than $\sqrt{n}$. This conclusion is contradicted by higher-order work.

Figure 8: Asymptotic form of d_n^λ



This figure shows a graph of $d_n^\lambda/\sqrt{n}$ for the range of $1000 \leq n \leq 25000$. A numerical study seems to indicate that this graph is not logarithmic, but instead is converging to a value of 1.208 or so. It should be noted that this graph shows no hint of oscillation, which is in sharp contrast to the graph of d_n^μ, which has a strongly manifested oscillation for this range of n.

Saddle-point analysis

The asymptotic behaviour of the d_n can be obtained by performing a saddle-point evaluation of the corresponding Norlund-Rice integral. That is, one writes

$$d_n = \frac{(-1)^n}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(s) \frac{n!}{s(s-1)(s-2)\cdots(s-n)} ds$$

where $f(s) = 1/\zeta(s)$ for d_n^μ, and $f(s) = \zeta(s-1)/\zeta(s)$ for d_n^φ, and so on. By assuming the Riemann Hypothesis, it has been shown that

$$\zeta(s) = O\left(|t|^\varepsilon\right)$$

and

$$\frac{1}{\zeta(s)} = O\left(|t|^\varepsilon\right)$$

for any $\varepsilon > 0$ and $s = \sigma + it$, as $t \to \infty$. This can be substituted directly into the integral. One should also take account of the fact that $1/\zeta(s)$ has a zero at $s = 1$ that cancels one of the poles in the integrand. Thus, one obtains,

$$\begin{aligned} d_n^\mu &= O(1) \int_{c-i\infty}^{c+i\infty} \frac{|s|^\varepsilon}{s} \frac{n!}{(s-2)(s-3)\cdots(s-n)} ds \\ &= O(1) \int_{c-i\infty}^{c+i\infty} e^{N\omega(s)} ds \\ &= O(1) \sqrt{\frac{2\pi}{N\omega''(s_0)}} e^{N\omega(s_0)} \end{aligned}$$

with s_0 the saddle point. Careful analysis shows that $N = \log n$, so that the saddle-point approximation is valid, and that the saddle point is located at

$$s_0 = \frac{1-\varepsilon}{\log n} \left[1 + \frac{1-\gamma}{\log n} + O\left(\frac{1}{\log^2 n}\right)\right]$$

This saddle point is located very near the origin; one must take the integration contour so that it passes from the lower-right quadrant, crosses $\sigma = 1/2$ to approach the origin, and departs to the upper-right quadrant.

The two parts of the saddle point formula are

$$e^{N\omega(s_0)} = \left(\frac{e \log n}{1-\varepsilon}\right)^{1-\varepsilon} + O(1)$$

and

$$\sqrt{\frac{2\pi}{N\omega''(s_0)}} = \frac{\sqrt{2\pi(1-\varepsilon)}}{\log n} \left(1 + O\left(\frac{1}{\log n}\right)\right)$$

Combining these, one obtains

$$d_n^\mu = O(1) \cdot \sqrt{\frac{2\pi}{1-\varepsilon}} \left(1 + \varepsilon \log\left(\frac{1-\varepsilon}{\log n}\right)\right) \left(1 + O\left(\frac{1}{\log n}\right)\right)$$

Pulling all of these factors together, one concludes that

$$d_n^{\mu} = O(\log\log n)$$

An alternate analysis

There is an alternate approach to the analysis which leads to some curious observations about the relationship between the Riemann Hypothesis and the location of saddle points. One may take as the starting point the well known formula

$$\zeta(s) = \frac{\exp(\log 2\pi - 1 - \gamma/2)\, s}{2(s-1)\Gamma(1+s/2)} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}$$

Here, the product extends over the zeros ρ of the Riemann zeta function. The derivation of this formula does not require the assumption of the RH, and so this seems to provide a more general starting point for a saddle-point analysis. To proceed, one must *assume* that the only saddle point is located near the origin. As will be made clear later, this assumption is equivalent to assuming RH.

Making this assumption that the only saddle point is near the origin, one then obtains for its location

$$s_0 = \frac{1}{\log n} + O\left(\frac{1}{\log^2 n}\right)$$

The only direct requirement on the location of the zeros was that they satisfy

$$\sum_{\rho} \frac{1}{\rho(\rho - s)} = O(1)$$

Proceeding with the saddle-point analysis, one obtains that

$$d_n = O(1)$$

Note that this result seems to be obtainable with only extremely weak assumptions about the location of the zeros of the Riemann zeta. This makes sense: the saddle point is essentially a local property, and local properties shouldn't really depend on the fine points of distant parts of the integrand. Thus, it is tempting to conclude that the result that $d_n = O(1)$ follows without requiring RH; but this would be wrong. This result was based on the assumption that the only saddle point was near the origin, and this assumption must surely require RH, as will be established in a later section.

Saddle-point analysis for the Totient series

This section summarizes a saddle-point analysis for the totient function finite difference series

$$d_n^{\varphi} = \sum_{k=3}^{n} (-1)^k \binom{n}{k} \frac{\zeta(k-1)}{\zeta(k)}$$

The integrand of the corresponding Norlund-Rice integral is regular at $s = 1$ and has a simple pole at $s = 0$, and a double pole at $s = 2$. In most other respects, it resembles the previous integral. It can be seen to have a saddle point near $s = 2$ which determines the asymptotic behaviour. Evaluating the integrand for the location of the saddle point, one obtains

$$s_0 = 2 + \frac{2}{\log n} + \frac{(1-\varepsilon)/2 - \gamma}{\log^2 n} + O\left(\frac{1}{\log^3 n}\right)$$

The two parts of the formula are

$$e^{N\omega(s_0)} = O(1) \cdot n^2 \log^2 n$$

and

$$N\omega''(s_0) = -2\log^2 n + \frac{1-\varepsilon}{4} + O\left(\frac{1}{\log n}\right)$$

Combining these, one obtains

$$d_n^{\varphi} = O\left(n^2 \log n\right)$$

This result is in good agreement with the numerical calculations for the range $10^2 < n < 10^4$.

Saddle-point analysis for the Liouville series

This section summarizes a saddle-point analysis for the Liouville function finite difference series

$$d_n^{\lambda} = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \frac{\zeta(2k)}{\zeta(k)}$$

The integrand of the corresponding Norlund-Rice integral is regular at $s = 1$ and has a simple pole at $2s = 1$. In all other respects, it resembles the previous integrals. Thus, a priori we expect a saddle point near $s = 1/2$ to determine the asymptotic behaviour. Indeed, proceeding as before, one obtains

$$s_0 = \frac{1}{2} + \frac{2+\varepsilon/4}{\log n} + O\left(\frac{1}{\log^2 n}\right)$$

The two parts of the saddle-point formula may be obtained from

$$e^{N\omega(s_0)} = O(1) \cdot \sqrt{n}\, \frac{\log n}{2+\varepsilon/4}$$

and

$$N\omega''(s_0) = \frac{1-\varepsilon/4}{2} \log^2 n + O(1)$$

which may be combined to obtain

$$d_n^{\lambda} = O\left(\sqrt{n}\right)$$

This result is in good agreement with the numerical calculations for the range $10^2 < n < 10^4$.

Equivalence to the Riemann Hypothesis

Based on the above, I'm concluding that the theorem regarding the equivalence to the RH is faulty. The Dirichlet series for μ, ϕ, and λ all have poles on the critical line. The general arguments for the asymptotic form of the d_n^{μ}, d_n^{ϕ} and d_n^{λ}, etc. are the same, but the actual form of the asymptotic behaviour is quite different. I don't see how this behaviour can be turned around to deduce the location of the poles. Before I realized this, I wrote the section below, attempting to generalize the proof; apparently it went astray, and I was unable to proceed.

0.1 Generalization of the proof

Based on the above, I suggest modifying the proof so that, in one direction, it runs as follows

Consider the general Dirichlet series

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

for some arithmetic function a_n, for example, $a_n = \mu(n)$. Defining

$$d_n = \sum_{k=2}^{n} (-1)^k \binom{n}{k} f(k)$$

one finds immediately that

$$d_n = \sum_{m=1}^{\infty} a_m \left[\left(1 - \frac{1}{m}\right)^n - 1 + \frac{n}{m} \right]$$

The leading order of a term in this sum is $a_m n^2/m^2$ and so convergence is guaranteed for $a_m = O\left(m^{1-\varepsilon}\right)$ for some $\varepsilon > 0$. This condition is satisfied by both the Mobius and the Liouville sums, for which $|\mu(m)| \leq 1$ and $|\lambda(m)| \leq 1$. The analogous sum for the totient function starts at k=3, and thus has the form

$$d_n = \sum_{m=1}^{\infty} a_m \left[\left(1 - \frac{1}{m}\right)^n - 1 + \frac{n}{m} - \frac{n(n-1)}{2m^2} \right]$$

More generally, the "Ramanujan conjecture" states that any series obeying RH will satisfy a similar bound; this is discussed further below.

Define

$$D(x) = \sum_{m=1}^{\infty} a_m \left[e^{-x/m} - 1 + \frac{x}{m} \right]$$

which has similar convergence properties to the previous sum. The estimate for the behaviour of this sum proceeds essentially unmodified, although I present some additional detail here to make this clear.

The aim is to show that d_n and $D(x)$ are substantially similar, and that in particular, both quantities have the same asymptotic behaviour for $n \to \infty$ (that is, $x \to \infty$). This will be done by comparing the two directly. Let $m_0 = n^{1-\delta}$ for some δ small but fixed. Write

$$d_n - D(n) = \sum_{m<m_0} + \sum_{m \geq m_0} a_m e^{-n/m} \left[e^{n/m + n\log(1-1/m)} - 1 \right]$$

For large m, namely, those appearing in the second sum, one easily finds

$$e^{n/m+n\log(1-1/m)} - 1 = O\left(n/m^2\right)$$

as $m \to \infty$ with n held fixed. From this, one obtains

$$\begin{aligned} \sum_{m \geq m_0} a_m e^{-n/m} \left[e^{n/m+n\log(1-1/m)} - 1 \right] &= O(n) \sum_{m \geq m_0} \frac{a_m}{m^2} + O\left(\frac{1}{n^{2-3\delta}}\right) \\ &= O\left(n^\delta\right) \end{aligned}$$

as $n \to \infty$. This result requires no additional constraints on the a_m other than the weak condition already given.

For the other sum, one has

$$\begin{aligned} \sum_{m<m_0} a_m e^{-n/m} \left[e^{n/m+n\log(1-1/m)} - 1 \right] &= O(1) \sum_{m<m_0} a_m e^{-n/m} \\ &= O\left(e^{-n/m_0}\right) \sum_{m<m_0} a_m \end{aligned}$$

The remaining sum is at most quadratic in m_0, and is dominated by the exponential term $e^{-n/m_0} = e^{-n^\delta}$. For fixed positive δ, this term decreases exponentially. Thus one concludes that

$$d_n - D(n) = O\left(n^\delta\right)$$

Again, this result requires no particularly special assumptions.

To show that d_n and $D(x)$ have an identical asymptotic behaviour for non-integer real x, one may consider the derivative $D'(x)$ and show that it is suitably bounded. Proceeding along the same lines as before, and starting with

$$D'(x) = -\sum_{m=1}^{\infty} \frac{a_m}{m} \left[e^{-x/m} - 1 \right]$$

one deduces that

$$D'(x) = O\left(n^\delta\right) + O(1) \sum_{m<m_0} \frac{a_m}{m}$$

At this point I notice an error in my notes, and, as it is late at night, I'd rather not guess my way out of it. I don't believe its serious. The remainder of the proof, regarding the Mellin transform goes through unmodified.

The only fact about Selberg class functions used in this proof seems to be the so-called "Ramanujan conjecture", namely, a necessary condition for RH to hold is that $a_n = O(n^{r+\varepsilon})$ for some real r and any $\varepsilon > 0$. In the above, it was required that $r < 2$, but other r's are easily accommodated. Simply changing the definition to

$$d_n = \sum_{k=n_0}^{n} (-1)^k \binom{n}{k} f(k)$$

allows for any $r < n_0$.

Comments

Just as you have shown the so-called "Poisson-Mellin-Newton" cycle, there appears to be a similar cycle, connection the so-called "Riesz means" defined by Hardy & Littlewood in the paper you sent me, with the so-called "Perron's formula" expressing a finite sum as a Mellin transform. There appear to be various bounds on finite sums that are known to number theory, and I am hoping that by means of Perron's formula, these can be converted over to the asymptotic behaviour of the Norlund-Rice integral that we desire. Maybe. But this is only after mulling things over while walking the dog, it might not work.

On further review of the Selberg class, it defines a lot of ingredients that seem to be necessary in order for GRH to hold. So far, we've used only one of these. Perhaps the others come into play in this half of the proof, which might suggest that this half of the proof is much harder?

Discussion

Its not clear how Li's criterion impacts on the Selberg class, and vice-versa. It may be an edifying exercise to connect the two together. Alternately, I'd like to translate Li's criterion to this proof, but don't quite see how.