On the difference $\pi(x) - li(x)$

Lee, Christine

2008

MIMS EPrint: **2010.83**

ISSN 1749-9097

On the difference $\pi(x) - \mathrm{li}(x)$

Christine Lee

Introduction

Riemann's memoir [13] and the subsequent efforts to prove his statements have answered a lot of the questions we have about prime numbers. For the prime counting function

$$\pi(x) := \#\text{ of primes} \leq x,$$

we now have an explicit formula which expresses this quantity in terms of known functions. We also have the Prime Number Theorem [11, p.168], which tells us that the chance that a large number N is a prime number is roughly $1/\log N$. Riemann's ideas illustrate the connection between complex analysis and number theory, and paved the way to many results that had been unimaginable. However, his paper also spawned many more problems, all of considerable difficulties. Numerous examples can be given, of which the Riemann Hypothesis is the most prominent [5]. However, in this paper, we will mainly be concerned with the question described below:
Consider the explicit formula

$$J(x) = \mathrm{li}(x) - \sum_{\rho} \mathrm{li}(x^{\rho}) + \int_x^{\infty} \frac{du}{(u^2-1)u\log u} - \log 2,$$

where $J(x)$ is defined by

$$J(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) \ldots$$

When Möbius inversion is used to write $\pi(x)$ as a function of $J(x)$, we have an expression of $\pi(x)$ in terms of $\mathrm{li}(x)$ and the zeros of the Riemann zeta function. This is the main result of Riemann's paper. A natural question to ask at this point is how well $\mathrm{li}(x)$ approximates $\pi(x)$. More precisely, we want to know the size of the error $\pi(x) - \mathrm{li}(x)$, whether it is always positive or negative. The numerical evidence to this date seems to suggest that this difference is always negative. Indeed, the increasing difference of $\pi(x) - \mathrm{li}(x)$ from select values up to 10^{22} could remove the doubt from anyone who finds numerical evidence convincing enough:

x	$\pi(x)$	$\pi(x) - \text{li}(x)$
10^8	5761455	-753
10^9	50847534	-1700
10^{10}	455052511	-3103
10^{11}	4118054813	-11587
10^{12}	37607912018	-38262
10^{13}	346065536839	-108970
10^{14}	3204941750802	-314889
10^{15}	29844570422669	-1052618
10^{16}	279238341033925	-3214631
10^{17}	2623557157654233	-7956588
10^{18}	24739954287740860	-21949554
10^{19}	234057667276344607	-99877774
10^{20}	2220819602560918840	-222744643
10^{21}	21127269486018731928	-597394253
10^{22}	201467286689315906290	-1932355207

Table 1: Primes up to selected x, and the values $\pi(x) - \text{li}(x)$.

It has also been verified through heavy computation that $\pi(x)$ is less than $\text{li}(x)$ for *all* x up to 10^{14} [8]. Nevertheless, it remains a *fact*, proven by Littlewood in 1914 [10], that not only does $\pi(x)$ exceed $\text{li}(x)$ at infinitely many x's on the real line, their differences also reach arbitrarily large magnitude. Forty one years later, Skewes succeeded in obtaining the first unconditional upperbound

$$10^{10^{10^{10^3}}}$$

below which $\pi(x)$ is guaranteed to exceed $\text{li}(x)$ [14]. This remarkably large number has since been greatly reduced through the joint efforts of the computer and mathematics, of which Lehman's computational formula [9] is the principal figure. In the subsequent sections of this paper I shall gradually develop the reasoning which leads to this result. To that end, I will include a self-contained proof of Littlewood's and Lehman's theorem (see [11] and [9]), while only assuming some standard facts from number theory and complex analysis. Although the methods used are elementary, the proofs themselves are by no means trivial and are highly sensitive to the validity and the falsity of the Riemann Hypothesis. This results in completely different proofs for both cases. Following this exposition, I will present a brief summary in the computational advances in lowering the upper bound below which $\pi(x) - \text{li}(x) > 0$, with particular emphasis on the result of Chao and Plymen [4]. Their computation gives the smallest upper bound to date.

This article comprises the author's MSc dissertation (Manchester University 2008); some slight editorial changes have been made by the author's supervisor Professor Roger Plymen.

Notations

z, s	A complex number.
$\zeta(s)$	The Riemann Zeta function on the complex plane $\mathbb{C}$, defined by $\sum_{n=1}^{\infty} \frac{1}{n^s}$ in the half plane $\mathrm{Re}(s) > 1$.
ρ	A zero of the zeta function in the critical strip $0 < \mathrm{Re}(s) < \frac{1}{2}$ in $\mathbb{C}$.
$\Gamma(s)$	The Gamma function. Defined by $\int_0^{\infty} t^{s-1} e^{-t}\, dt$ on the half plane $\mathrm{Re}(s) > 0$.
β	The real part of ρ.
γ	The imaginary part of ρ.
x	A real number.
n	A nonnegative integer.
$\mu(n)$	$:= 0$ if n is not square free, $:= (-1)^k$, where k is the number of distinct prime factors of n. The Möbius function.
$\Lambda(n)$	$:= \log p$ if $n = p^k$ for a prime p, $:= 0$ otherwise.
$\log(x)$	:= The natural logarithm of x.
$\pi(x)$	The number of primes not exceeding x.
$\theta(x)$	$:= \sum_{p \text{ prime } \leq x} \log p$.
$J(x)$	$:= \sum_{k=1}^{\infty} \frac{1}{k} \pi(x^{1/k})$.
$\psi(x)$	$:= \sum_{n \leq x} \Lambda(n)$.
ϑ	A complex number satisfying $\lvert\vartheta\rvert \leq 1$.
$\lim_{n\to\infty} \inf x_n$	$:= \sup\{\inf\{x_m : m \geq n\} : n \geq 0\}$.
$\lim_{n\to\infty} \sup x_n$	$:= \inf\{\sup\{x_m : m \geq n\} : n \geq 0\}$.
$\lVert x\rVert$	The distance from x to the nearest integer.
$[\, x\,]$	The greatest integer not exceeding x.
$\{x\}$	The fractional part of x.
$f(z) = O(g(z))$	$\lvert f(x)\rvert \leq Cg(x)$ where C is an absolute constant.
$f(z) \ll g(z)$	$f(z) = O(g(z))$.
$f(z) \gg g(z)$	$g(z) = O(f(z))$.
$f(x) \asymp g(x)$	$cf(x) \leq g(x) \leq Cf(x)$ for some positive absolute constants c, C.
$f(x) \sim g(x)$	$\lim_{x\to\infty} f(x)/g(x) = 1$.

We list some preliminary lemmas.

The Riemann-Stieltjes Integral [11, p.486–494] [1]

Because we will frequently encounter sums taken over discrete step functions it is essential to develop a notion of integrals to work with them. This enables us to make very precise estimates for select portions of infinite series such as

$$\sum_{0<|\gamma|<\infty} \frac{1}{\gamma^2} .$$

For $a, b \in \mathbb{R}$ and $a < b$, let $\mathbf{x} = \{x_0, x_1, \ldots, x_n\}$ be a partition of the interval $[a, b]$:

$$a = x_0 \leq x_1 \leq \cdots x_n = b .$$

Let x'_k be any number such that $x_{k-1} \leq x'_k \leq x_k$, we can form the sum from $\mathbf{x}$ and $\mathbf{x}' = \{x'_1, x'_2, \ldots, x'_n\}$:

$$S(\mathbf{x}, \mathbf{x}') = \sum_{k=1}^{n} f(x'_k)(g(x_k) - g(x_{k-1})).$$

Definition. We say that the *Riemann-Stieltjes integral* $\int_a^b f(x)\, dg(x)$ exists and has the value I if for evey $\epsilon > 0$ there is a $\delta > 0$ such that

$$|S(\mathbf{x}, \mathbf{x}') - I| \leq \epsilon$$

whenever

$$\text{mesh}\{\mathbf{x}\} = \max_{1 \leq k \leq n} (x_k - x_{k-1}) \leq \delta .$$

Now we prove a criteria for the existence of the Riemann-Stieltjes integral $\int_a^b f(x)\, dg(x)$ for a pair of functions f and g.

Definition. We define the *variation* of g on an interval $[a, b]$ to be

$$\text{Var}_{[a,b]}(g) = \sup_{\mathbf{x}} \sum_{k=1}^{n} |g(x_k) - g(x_{k-1})| ,$$

where the supremum is taken over all partitions $\mathbf{x}$ on $[a, b]$.

Lemma 0.1. *The Riemann-Stieltjes integral $\int_a^b f(x)\, dg(x)$ exists if f is continuous on $[a, b]$ and g is of bounded variation.*

Proof. Since f is continuous on $[a, b]$, for every $\epsilon > 0$ we can find a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $x, y \in [a, b]$ satisfying $|x - y| < \delta$. We show that for every $\epsilon > 0$, there exists a δ such that

$$|S(\mathbf{x}, \mathbf{x}') - S(\mathbf{y}, \mathbf{y}')| \leq 2\epsilon\, \text{Var}_{[a,b]}(g)$$

[1] obtained directly from [11] with minor changes in notation.

whenever $\text{mesh}\{\mathbf{x}\}, \text{mesh}\{\mathbf{y}\} < \delta$. Consider the partition $\mathbf{z}$ which is the union of $\mathbf{x}$ and $\mathbf{y}$. We can write

$$\begin{aligned} &S(\mathbf{x}, \mathbf{x}') \\ &= \sum_{k=1}^{n} f(x'_k)(g(x_k) - g(x_{k-1})) \\ &= \sum_{k=1}^{n} f(x'_k)(g(x_k) - g(z_{k\ell}) + g(z_{k\ell}) - g(z_{k(\ell-1)}) + \cdots + g(z_{k1}) - g(z_{x_{k-1}}))\,, \end{aligned}$$

where $\{z_{k1}, z_{k2}, \ldots, z_{k\ell}\}$ is the partition of $\mathbf{z}$ on the interval (x_{k-1}, x_k). The absolute value of difference can be written as

$$\begin{aligned} &|S(\mathbf{x}, \mathbf{x}') - S(\mathbf{y}, \mathbf{y}')| \\ &= \left| \sum_{i=1}^{I} (f(x'_i) - f(y'_i))(g(z_i) - g(z_{i-1})) \right| \end{aligned}$$

Since $\text{mesh}\{x\}$ and $\text{mesh}\{y\}$ are less than δ, $|f(x'_i) - f(y'_i)| < \epsilon$ for all $1 \le i \le I$, and

$$\begin{aligned} &\le \epsilon \left| \sum_{i=1}^{I} (g(z_i) - g(z_{i-1})) \right| \\ &\le \epsilon \sum_{i=1}^{I} |g(z_i) - g(z_{i-1})| \\ &\le \epsilon \operatorname{Var}_{[a,b]}(g), \end{aligned}$$

and we are done since we have a Cauchy sequence. □

The reason for defining this integral is so that we can write

$$\sum_{k=1}^{n} a_k f(k) = \int_0^n f(x)\, dA(x)\,,$$

where $A(k) - A(k-1) = a_k$ and $A(x) = \sum_{k=0}^{j-1} a_k$ for $x \in [j-1, j)$. Certainly the endpoints of the integral are flexible as long as they don't change the value of the the sum.

We also need the formula for integration by parts.

Lemma 0.2. *If $\int_a^b f\, dg$ exists for functions f, g, then $\int_a^b g\, df$ also exists, and*

$$\int_a^b g\, df = f(b)g(b) - f(a)g(a) - \int_a^b f\, dg.$$

Proof. For any partition $\mathbf{x}$ on $[a, b]$, we expand the sum $S(\mathbf{x}, \mathbf{x}')$ according to the definition for $\int_a^b f\, dg$:

$$S(\mathbf{x}, \mathbf{x}') = \sum_{k=1}^{n} f(x'_k)(g(k) - g(k-1))\,,$$

since $x_0 = a$ and $x_n = b$, we have, after choosing $x'_1 = a$ and $x'_n = b$,

$$= f(a)(g(x_1) - g(a)) + f(x_2)(g(x_2) - g(x_1)) + \cdots + f(b)(g(b) - g(x_{n-1}))$$
$$= f(b)g(b) - f(a)g(a) - \sum_{k=1}^{n} g(x_k)(f(x'_k) - f(x'_{k-1}))\,.$$

thus if we take the sums over all possible partitions of $[a, b]$ we have

$$\int_a^b g\,df = f(b)g(b) - f(a)g(a) - \int_a^b f\,dg\,.$$ □

The final, and the most useful lemma developing the notion of the Riemann-Stietljes integral is the following:

Lemma 0.3. *If g' is continuous on $[a, b]$, then*

$$Var_{[a,b]}g = \int_a^b |g'(x)|\,dx\,.$$

If in addition f is Riemann integrable, then

$$\int_a^b f(x)\,dg(x) = \int_a^b f(x)g'(x)\,dx\ .$$

Proof. Suppose g' is continuous on $[a, b]$, then by the mean value theorem there is a $x'_k \in [x_{k-1}, x_k]$ such that $g'(x'_k)(x_k - x_{k-1}) = g(x_k) - g(x_{k-1})$, thus

$$\sum_{k=1}^{n} |g(x_k) - g(x_{k-1})| = \sum_{k=1}^{n} |g'(x'_k)(x_k - x_{k-1})|$$

Consider two partitions $\mathbf{x}, \mathbf{y}$, where $\mathbf{x}$ is a subpartition of $\mathbf{y}$, then

$$\sum_{k=1}^{n} |g(x_k) - g(x_{k-1})| \leq \sum_{k=1}^{n} |g(y_k) - g(y_{k-1})|$$

by the triangle inequality. Certainly the Riemann-Stieltjes integral $\int_a^b |g'(x)|dx$ exists by Lemma 0.1 since g' is continuous. It is also the supremum of the right hand side of the equation above, and we have the first statement. The second statement is given by a similar substitution.

$$\sum_{k=1}^{n} f(x'_k)(g(x_k) - g(x_{k-1}))$$
$$= \sum_{k=1}^{n} f(x'_k)g'(x''_k)(x_k - x_{k-1})$$
$$= \sum_{k=1}^{n} f(x'_k)g'(x'_k)(x_k - x_{k-1}) + \sum_{k=1}^{n} f(x'_k)(g'(x''_k) - g'(x'_k))(x_k - x_{k-1})$$

where $x''_k, x'_k \in [x_k, x_{k-1}]$. Since g' is continuous, as mesh$\{\mathbf{x}\}$ tends to zero, the second term of the sum tends to

$$\epsilon \sum_{k=1}^{n} f(x'_k)(x_k - x_{k-1}) = \epsilon M,$$

and M is the value of the Riemann integral of f on the interval $[a, b]$, and we are done. □

Backlund's Formula [9, p. 399][2]

The Riemann-Stieltjes integral will be used in combination with the following formula for the number of zeta zeros up to a height T to estimate sums involving the zeta zeros. We will also assume for the rest of the paper that whenever we encounter a sum taken over all the zeros of the zeta function, each term of the sum is arranged according to the increasing order of the imaginary parts.

Lemma 0.4. *Let $N(T)$ be the number of zeta zeros ρ with $0 < \gamma < T$, then*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + Q(T) \qquad \text{for} \qquad (T > 2)$$

where

$$|Q(T)| < 0.137\log T + 0.443\log\log T + 4.35\,.$$

With this formula and the Stieltjes integral, we can prove the following.

Lemma 0.5.

$$\sum_{\gamma > T}\frac{1}{\gamma^n} < T^{1-n}\log T\,,$$

where the γ's are arranged in increasing order.

Proof.

$$\sum_{\gamma>T}\frac{1}{\gamma^n} = \int_T^\infty \frac{1}{t^n}\,dN(t)$$
$$= \int_T^\infty \frac{1}{2\pi t^n}\log\frac{t}{2\pi}\,dt + \int_T^\infty \frac{1}{t^n}Q'(t)\,dt$$
$$= \frac{1}{2\pi}\left(\frac{1}{t^{n-1}(-n+1)}\log\frac{t}{2\pi} - \frac{1}{t^{n-1}(-n+1)^2}\right)\Bigg|_T^\infty$$
$$+ \frac{1}{t^n}Q(t)\Bigg|_T^\infty + n\int_T^\infty \frac{1}{t^{n+1}}Q(t)\,dt$$

Since $|Q(t)| < 0.137\log T + 0.443/2\log T + 1.533\log T < 2\log T$ for $T > 2\pi e$,

$$< T^{1-n}\log T\left(\frac{1}{2\pi(n-1)} + \frac{\log 2\pi}{2\pi\log T(n-1)} + \frac{1}{2\pi(n-1)^2\log T}\right)$$
$$+ \frac{2\log T}{T^n} + \frac{1}{t^n}2\log t\Bigg|_T^\infty - 2\int_T^\infty \frac{1}{t^n}d\log t$$

where since $n \geq 2$,

$$\leq T^{1-n}\log T\left(\frac{1}{2\pi} + \frac{\log 2\pi}{2\pi\log T} + \frac{1}{4\pi^2\log T}\right) + T^{1-n}\log T\left(\frac{4}{T} + \frac{1}{T\log T}\right).$$

The terms multiplying $T^{1-n}\log T$ is less than 0.574, so

$$< T^{1-n}\log T.$$

□

[2] Here we are using the formulation from [1] as used by Lehman because it is the most directly accessible. There are different formulations for other uses which can be found in number theory texts such as [11].

Here is an estimate of the sum:

$$\sum_{0<\gamma<\infty} \frac{1}{\gamma^2} < 0.025\,. \tag{0.1}$$

Cauchy's Integral Formula

We will also use some standard lemmas from complex analysis.

Definition. Let $A \subset B \subseteq \mathbb{C}$ and $f : A \to \mathbb{C}$ be analytic. A *meromorphic extension* of f is a meromorphic function $g : B \to \mathbb{C}$ such that $g|_A = f$.

Lemma 0.6. *The meromorphic extension of an analytic function to a larger domain is unique; i.e., with the notation above, if $h : B \to \mathbb{C}$ is meromorphic and has the property that $h|_A = f$, then $g = h$ on B.*

Lemma 0.7. *Suppose U is an open subset of the complex plane $\mathbb{C}$, and $f : U \to \mathbb{C}$ is a holomorphic function defined on U, and the closed disk D centered at z_0 with radius r is completely contained in U. Let C be the circle forming the boundary of D. Then for every z in the interior of D,*

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(s)}{s-z}\, ds$$

and

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_C \frac{f(s)}{(s-z)^{n+1}} ds\ .$$

Explicit Formulas

We shall use the following explicit formulas without proof. [3] The logarithmic integral is defined as follows.

Definition.

$$\mathrm{li}(e^z) := \int_{-\infty+iy}^{x+iy} \frac{e^t}{t} dt$$

where $z = x + iy$, $y \neq 0$.

For $x > 1$, $\mathrm{li}(x)$ is then defined as follows:

$$\mathrm{li}(x) := \frac{1}{2}[\mathrm{li}(x + i0) + \mathrm{li}(x - i0)]$$

In this way, we recover the classical definition of $\mathrm{li}(x)$ as an integral principal value:

$$\mathrm{li}(x) = \lim_{\epsilon\to 0} \left(\int_0^{1-\epsilon} \frac{e^t}{t} dt + \int_{1+\epsilon}^{x} \frac{e^t}{t} dt \right)$$

[3]Interested readers can check out [6], which discusses the proofs and the history of these explicit formulas in detail.

Lemma 0.8.

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) + \sum_{k=1}^{\infty} \frac{x^{-2k}}{2k} \ . \tag{0.2}$$

Lemma 0.9.

$$J(x) = \mathrm{li}(x) - \sum_\rho \mathrm{li}(x^\rho) + \int_x^\infty \frac{du}{(u^2-1)u\log u} - \log 2 \ . \tag{0.3}$$

We use the following formulation of the Riemann Hypothesis for Littlewood's theorem:
The supremum of the real parts of the zeros of the zeta function does not exceed $\frac{1}{2}$*; i.e. write* $\rho = \beta + i\gamma$*, then* $\beta \leq \frac{1}{2}$ *.*
This alternative formulation is used by Lehman:
The real parts of the zeros of the zeta function lie on the critical line; i.e. write $\rho = \beta + i\gamma$*, then* $\beta = \frac{1}{2}$*.*

1 Littlewood's theorem

How do we even know whether $\pi(x)$ exceeds $\mathrm{li}(x)$ at some point? Let us look at the explicit formulas. We can write[4]

$$\pi(x) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} J(x^{1/k}) \ . \tag{1.1}$$

By (0.3), we have

$$\pi(x) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \left[\mathrm{li}(x^{1/k}) - \sum_\rho \mathrm{li}(x^{\rho/k}) + \int_{x^{1/k}}^\infty \frac{du}{(u^2-1)u\log u} \right]^5 . \tag{1.2}$$

when the difference is written in this way the nature of the difference $\pi(x) - \mathrm{li}(x)$ is difficult to obtain. Therefore, this formula is not very useful when the goal is just to prove that the difference changes sign at *some point*. Instead, we will make do with very rough estimates of $\pi(x)$. This difference in approach will be important later when we actually want to find a precise value for the first x where the sign switch occurs.

Since we prefer not to work directly with the formula for $\pi(x)$, however accurate it is, we consider other functions which are related to $\pi(x)$. It turns out that the prime number theorem can be easily deduced from the asymptotic formula $\psi(x) \sim x$ [6, p.76–77]. We consider the function $\theta(x)$, which gives the following equation

$$\psi(x) = \sum_{k=1}^{\infty} \theta(x^{1/k}) \ .$$

[4] This can be seen by substituting the definition for $J(x)$ into the sum and noting that $\sum_{d|n} \mu(d) = 0$ for $n > 1$, or by the sieve method.

[5] The constant term $\log 2$ disappears because $\sum_{k=1}^{\infty} \frac{\mu(k)}{k} = 0$.

On the other hand[6],

$$\begin{aligned}&(1-\delta)\log x(\pi(x)-x^{1-\delta}) \le \theta(x) \le \log x\ \pi(x)\\ &\Rightarrow \frac{\pi(x)\log x}{\theta(x)} \le \frac{\log x}{\theta(x)}x^{1-\delta} + \frac{1}{1-\delta} \text{ and } \frac{\pi(x)\log x}{\theta(x)} \ge 1\end{aligned}$$

for any constant δ such that $0 \le \delta \le 1$. Since we can deduce $\theta(x) \sim x$ from $\psi(x) \sim x$, the term $\frac{\log x}{\theta(x)}x^{1-\delta}$ disappears because

$$\frac{\log x}{\theta(x)}x^{1-\delta} \sim \frac{\log x}{x}x^{1-\delta} = 0 \quad \text{as } x \to \infty\,,$$

and we can choose δ such that $\frac{\pi(x)\log x}{\theta(x)}$ is as close to 1 as we would like as x tends to infinity, thus completing the proof. What is noteworthy here is the connection between $\psi(x)$ and $\pi(x)$ via $\theta(x)$. From the last two inequalities, we have

$$\pi(x) \le x^{1-\delta} + \frac{\theta(x)}{\log x}\left(\frac{1}{1-\delta}\right)\ .$$

By definition of $\psi(x)$ we also have that

$$\psi(x) = \theta(x) + x^{1/2} + O(x^{1/3})\ .$$

We reconsider the explicit formula of $\psi(x)$:

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) + \sum_{k=1}^{\infty}\frac{x^{-2k}}{2k}\ .$$

It is apparent now, that if only the first term in the sum of the right hand side of the explicit formula for $\pi(x)$ "counts", namely, that

$$\pi(x) = \mathrm{li}(x) - \sum_\rho \mathrm{li}(x^\rho) + \int_x^\infty \frac{du}{(u^2-1)u\log u} - \log 2 + \text{error}, \tag{1.3}$$

there is an analog between the difference $\psi(x)-x$ and $\pi(x)-\mathrm{li}(x)$. Write $c = 1-\delta$, all we know is that

$$\pi(x) - \mathrm{li}(x) = \frac{1}{c}\frac{\theta(x)}{\log x} + O(x^c) - \mathrm{li}(x)$$

Substitute $\theta(x) = \psi(x) - x^{1/2} + O(x^{1/3})$, we have

$$= \frac{1}{c}\frac{\psi(x)-x}{\log x} + \left(\frac{x - x^{1/2} + O(x^{1/3})}{\log x} + O(x^c) - \mathrm{li}(x)\right)$$

6

$$\begin{aligned}\theta(x) &= \sum_{p \text{ prime } \le x} \log p \le \sum_{p \text{ prime } \le x} \log x = \pi(x)\log x\,, \text{ and}\\ &\ge \sum_{x^{1-\delta} \le p \text{ prime } \le x} \log p \ge \sum_{x^{1-\delta} \le p \text{ prime } \le x} (1-\delta)\log x \ge (1-\delta)\log x(\pi(x) - \pi(x^{1-\delta}))\end{aligned}$$

If $\psi(x) - x$ oscillates with a magnitude that overpowers the terms in the parenthesis, then we have that $\pi(x) - \mathrm{li}(x)$ oscillates as well. We can make a further simplification if we note that [11]

$$\frac{d}{du}\left[\frac{u}{\log u}\right] = \frac{1}{\log u} - \frac{1}{(\log u)^2},$$

and

$$\mathrm{li}(x) = \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{1}{(\log u)^2}\, du$$

so

$$\begin{aligned}\pi(x) - \mathrm{li}(x) &= \frac{1}{c}\frac{\psi(x) - x}{\log x} - \frac{x^{1/2}}{\log x} + \frac{O(x^{1/3})}{\log x} + \frac{2}{\log 2} - \int_2^x \frac{1}{(\log u)^2}\, du + O(x^c)\\ &= \frac{1}{c}\frac{\psi(x) - x}{\log x} + O\left(\frac{x^{1/2}}{\log x}\right) - \int_2^x \frac{1}{(\log u)^2}\, du + O(x^c)\,.\end{aligned}$$

Now we see that it's not enough that $\psi(x) - x$ changes sign, it also has to change sign with magnitude greater than $x^{1/2}$. The upper bound of the error term $\pi(x) - \frac{\theta(x)}{\log x}$ indicated by $O(x^c)$ does not suffice anymore, we actually need a bound on the size of the error term minus $\int_2^x \frac{1}{(\log u)^2} du$.

We use the Stieltjes integral to write

$$\begin{aligned}\pi(x) &= \int_2^x \frac{1}{\log u}\, d\theta(u)\\ &= \left[\frac{1}{\log u}\theta(u)\right]_2^x + \int_2^x \theta(u)\frac{1}{u(\log u)^2}\, du\,.\end{aligned}$$

So now the error term plus the integral becomes

$$\int_2^x \frac{\theta(u) - u}{u(\log u)^2}\, du\,,$$

and with integration by parts,

$$= \left[\frac{\theta_1(u) - u^2/2}{u(\log u)^2}\right]_2^x + \int_2^x \frac{\theta_1(u) - u^2/2}{u^2(\log u)^2}\left(1 - \frac{2}{(\log u)^3 u^2}\right) du$$

where $\theta_1(x)$ is the integral of $\theta(x)$ from 2 to x. Since $\theta(x)$ can be written in terms of $\psi(x)$, we make use of the explicit formula for $\psi(x)$ to find the integral of $\theta(x)$. The only problem is that there is an oscillatory sum over the zeta zeros in the expression of $\psi(x)$.

$$\int_2^x \psi(u)\, du = \frac{x^2}{2} - \sum_\rho \frac{x^{\rho+1}}{\rho+1} - \frac{\zeta'}{\zeta}(0)x + \frac{\zeta'}{\zeta}(-1) + O(x^{-1/2}),\ \text{but}$$

$$\left|\sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)}\right| \le \sum_\rho \frac{x^{\beta+1}}{\gamma^2}\,,$$

and the numerator of each term is bounded by $x^{\Theta+1}$, where Θ denotes the supremum of the real parts of the zeta zeros. Also, $\sum_{\gamma>0} \frac{1}{\gamma^2}$ converges by (0.1) , so the oscillatory term is $O(x^{\Theta+1})$. If we assume that the Riemann hypothesis is true, we can immediately integrate the explicit formula for $\psi(x)$ and obtain

$$\begin{aligned}\theta_1(x) &= \int_2^x \psi(x) - x^{1/2} + O(x^{1/3})\, dx \\ &= \frac{x^2}{2} + O(x^{3/2})\,.\end{aligned}$$

Finally,

$$\begin{aligned}\int_2^x \frac{\theta(u) - u}{u(\log u)^2}\, du &= \frac{x^2/2 + O(x^{3/2}) - x^2/2}{x(\log x)^2} + O(1) \\ &\quad + \int_2^x \frac{u^2/2 + O(u^{3/2}) - u^2/2}{u^2(\log u)^2} \left(1 - \frac{2}{(\log u)^3 u^2}\right) du \\ &= O\left(\frac{x^{1/2}}{(\log x)^2}\right)\end{aligned}$$

after estimating the last integral. Now we can be assured that every term in the expression for $\pi(x) - \mathrm{li}(x)$ is bounded by $\frac{x^{1/2}}{\log x}$. We drop the now irrelevant constant c and write

$$\pi(x) - \mathrm{li}(x) = \frac{\psi(x) - x}{\log x} + O\left(\frac{x^{1/2}}{\log x}\right)\,. \tag{1.4}$$

We introduce the main statement that leads to Littlewood's Theorem and the notation $f(x) = \Omega_\pm(g(x))$.

Theorem 1.1. *If the Riemann Hypothesis is true, then [11, p.477]*

$$\frac{1}{(e^\delta - e^{-\delta})x} \int_{e^{-\delta x}}^{e^{\delta x}} (\psi(u) - u) du = -2x^{1/2} \sum_{\gamma>0} \frac{\sin \gamma\delta}{\gamma\delta} \cdot \frac{\sin(\gamma \log x)}{\gamma} + O(x^{1/2}) \tag{1.5}$$

uniformly for $x \geq 4, 1/2x \leq \delta \leq 1/2$.

This formula fully reveals the oscillatory nature of the difference $\psi(x) - x$. Depending δ and $\log x$, multiplying by γ sends $\sin\gamma\delta$ to different values on the sine curve. At a casual glance, the interactions between each term of the sum seem quite complicated, but with the help of Dirichlet's theorem, the sum proves to be quite manageable. Since x is taken over all real numbers, we might even guess that the difference on the left hand side switches sign infinitely many times because of the periodic nature of sine. Furthermore, it does so while achieving a magnitude that is at least $x^{1/2}$. We make this notion rigorous by introducing the notation $\Omega_\pm$.

Definition. Let $f(x), g(x)$ be real-valued functions, we write $f(x) = \Omega_+(g(x))$ if $\limsup_{x\to\infty} f(x)/g(x) > 0$, $f(x) = \Omega_-(g(x))$ if $\liminf_{x\to\infty} f(x)/g(x) > 0$, and $f(x) = \Omega_\pm(g(x))$ if both statements are true.

We can now state Littlewood's theorem:

Theorem 1.2. *(Littlewood) [11, p.478–479]*
Assuming the Riemann Hypothesis, as $x \to \infty$,

$$\psi(x) - x = \Omega_{\pm}(x^{1/2} \log\log\log x) \tag{1.6}$$

and

$$\pi(x) - \mathrm{li}(x) = \Omega_{\pm}(x^{1/2}(\log x)^{-1} \log\log\log x) \tag{1.7}$$

The first statement is derived from (1.5), the second statement is a direct consequence of the first and (1.4). The factor $\log\log\log x$ is the lower bound on the magnitude of the oscillation of $\sum_{\gamma>0} \frac{\sin\gamma\delta}{\gamma\delta} \cdot \frac{\sin(\gamma \log x)}{\gamma}$. As we will see when we prove the first statement, assuming the Riemann Hypothesis means the error term $O(x^{1/2})$ in (1.5) is not easily overwhelmed. Therefore, the case where the Riemann hypothesis is true is the more difficult one to prove.

Proof. (Theorem 1.1)
The formula (1.5) is the average of $\psi(x) - x$ over an interval. If this value is positive then there must be a point within the interval where $\psi(x)$ exceeds x, therefore it is sufficient to prove that the sign-switching behavior occurs for this integral.

Use the explicit formula for $\psi(x)$ and write

$$\int_0^x \psi(u) - u\, du = -\sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)} - \frac{\zeta'}{\zeta}(0)x + O(1).^7 \tag{1.8}$$

Replacing x with $e^{\pm\delta}x$, we have

$$\begin{aligned}\int_{e^{-\delta}x}^{e^{\delta}x} (\psi(u) - u)\, du &= -\int_0^{e^{-\delta}x} (\psi(u) - u)\, du + \int_0^{e^{\delta}x} (\psi(u) - u)\, du + O(1) \\ &= -\sum_\rho \frac{(e^\delta x)^{\rho+1} - (e^{-\delta}x)^{\rho+1}}{\rho(\rho+1)} - \frac{\zeta'}{\zeta}(0)(e^\delta x - e^{-\delta}x) + O(1)\,.\end{aligned}$$

Dividing both sides by $e^\delta x - e^{-\delta}x = 2\sinh(\delta)x$,

$$= -\frac{\delta}{2\sinh(\delta)} \sum_\rho \frac{(e^{\delta(\rho+1)} - e^{-\delta(\rho+1)})x^\rho}{\delta\rho(\rho+1)} + O(1)\,. \tag{1.9}$$

We invoke the Riemann Hypothesis and get that

$$\begin{aligned} e^{\delta(\rho+1)} &= e^{\delta(1/2+i\gamma+1)} \\ &= e^{3/2\delta} \cdot e^{\delta i\gamma} \\ &= \left(\sum_{n=0}^{\infty} \frac{(3/2\delta)^n}{n!}\right) e^{\delta i\gamma} \\ &= (1 + O(\delta))e^{\delta i\gamma}\end{aligned}$$

[7] The $O(1)$ comes from the sum over the trivial zeros of the zeta function, $\sum_{n=0}^{\infty} \int \frac{x^{-2n}}{2n} = \sum_{n=0}^{\infty} \frac{x^{-2n+1}}{(-2n+1)2n} = O(1)$.

since $\delta \leq 1/2 < 1$. We use the same estimate for $e^{-\delta(\rho+1)}$. Replacing both terms with those estimates we get

$$(1.9) = -\frac{\delta}{\sinh\delta}\sum_{\rho}\frac{(1+O(\delta))(e^{\delta i\gamma}-e^{-\delta i\gamma})}{2\delta\rho(\rho+1)}x^{\rho} + O(1)\,.$$

Since $\frac{d}{d\delta}\left[\sinh(\delta)\right] = \frac{e^{\delta}+e^{-\delta}}{2}$ which is positive for all δ, $\sinh(\delta)$ is strictly increasing and so it reaches its smallest point at $\delta = \frac{1}{2x}$. Moreover, $x \geq 4$, so the smallest δ can be is zero, and we have that $\frac{1}{\sinh\delta}$ is $O(1)$. We bound the other terms in the error term in the sum multiplied by $O(\delta)$:

$$|e^{\delta i\gamma}-e^{-\delta i\gamma}| = O(1)\,,\quad |x^{\rho}| = O(x^{1/2})\,,\quad \left|\sum_{\rho}\frac{1}{\rho(\rho+1)}\right| \leq \sum_{\rho}\frac{1}{\gamma^2} = O(1)\,.$$

Thus the error term is bounded by $x^{1/2}$. We reconsider $\delta/\sinh(\delta)$ and take the Taylor expansion of $\sinh(\delta)$ to obtain $\delta/\sinh(\delta) = 1 + O(\delta^2)$. We once again try to bound the error term

$$O(\delta^2)\cdot -ix^{1/2}\sum_{\rho}\frac{\sin\gamma\delta}{\delta}\cdot\frac{x^{i\gamma}}{\rho(\rho+1)}\,, \tag{1.10}$$

where $\sin\delta\gamma = \frac{e^{i\delta\gamma}-e^{-i\delta\gamma}}{2i}$. Since $O(\delta^2) = O(1)$, we need only to deal with the sum multiplied by $\frac{1}{\delta}$. If $\gamma \leq \frac{1}{\delta}$ then $\gamma\delta \leq 1$ and $\frac{\sin\gamma\delta}{\gamma\delta} \leq 1$, below the curve $x = y$. Thus we can split the sum into two parts

$$\begin{aligned}\frac{1}{\delta}\sum_{\rho}\frac{\sin\gamma\delta}{\rho(\rho+1)} &\leq \frac{1}{\delta}\sum_{\rho}\frac{1}{\gamma^2}\\ &\leq \sum_{0\leq\gamma\leq 1/\delta}\frac{1}{\gamma} + \frac{1}{\delta}\sum_{\gamma\geq 1/\delta}\frac{1}{\gamma^2}\,.\end{aligned} \tag{1.11}$$

To evaluate the first sum we form the Stieltjes integral and use the more general form of Lemma 0.4:

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T)\,.$$

$$\begin{aligned}\sum_{0\leq\gamma\leq 1/\delta}\frac{1}{\gamma} &= \int_0^{1/\delta}\frac{1}{t}\,dN(t)\\ &= \int_0^{1/\delta}\frac{1}{2\pi t}\log\frac{t}{2\pi}\,dt + \int_0^{1/\delta}\frac{1}{t}\,d(O(\log t))\,.\end{aligned}$$

We use integration by parts for the second term, choosing $dv = dO(\log t)$, $v = O(\log t)$, and get

$$\int_0^{1/\delta}\frac{1}{t}\,d(O(\log t)) = \left[\frac{O(\log t)}{t}\right]_0^{1/\delta} + \int_0^{1/\delta}O(\log t)\frac{1}{t^2}\,dt$$

that $\log t$ does not evaluate at 0 is not an issue because the interval of the Stieltjes integral is flexible and we can always choose the lower bound to be slightly smaller than the imaginary component of the first zeta zero above the real axis.

$$\begin{aligned}
&= O\left(\frac{\log 1/\delta}{1/\delta}\right) + \left\{\text{a term} \leq \int_0^{1/\delta} \frac{\log t}{t^2}\, dt\right\} \\
&= O\left(\frac{\log 1/\delta}{1/\delta}\right) + O\left(\left[\log t \cdot \frac{-1}{2\pi}\right]_0^{1/\delta} + \int_0^{1/\delta} \frac{1}{t^2}\, dt\right) \\
&= O\left(\frac{\log 1/\delta}{1/\delta}\right) + O(\delta)\,,
\end{aligned}$$

and the first term is just

$$\int_0^{1/\delta} \frac{1}{2\pi t} \log \frac{t}{2\pi}\, dt = \left[\frac{\left(\log \frac{t}{2\pi}\right)^2}{8\pi^2}\right]_0^{1/\delta} = O((\log(1/\delta))^2)\,.$$

Thus

$$\sum_{0<\gamma\leq 1/\delta} \frac{1}{\gamma} = O((\log(1/\delta))^2).$$

We write similarly,

$$\begin{aligned}
\frac{1}{\delta} \sum_{\gamma>1/\delta} \frac{1}{\gamma^2} &= \int_{1/\delta}^{\infty} \frac{1}{t^2}\, dN(t) \\
&= \int_{1/\delta}^{\infty} \frac{1}{2\pi t^2} \log \frac{t}{2\pi}\, dt + \int_{1/\delta}^{\infty} \frac{1}{t^2}\, d(O(\log t)) \\
&= \frac{\log \frac{1/\delta}{2\pi}}{1/\delta} + O(\delta) + \frac{\log 1/\delta}{(1/\delta)^2} + O\left(\frac{\log 1/\delta}{1/\delta}\right) \\
&= O(\log(1/\delta))\,.
\end{aligned}$$

and the entire sum over ρ is $O((\log 1/\delta)^2)$. We have

$$(1.10) = O\left(x^{1/2} \frac{(\log 1/\delta)^2}{(1/\delta)^2}\right) = O(x^{1/2})\,,\ \text{and}$$

$$(1.9) = -ix^{1/2} \sum_{\rho} \frac{\sin \gamma\delta}{\delta} \cdot \frac{x^{i\gamma}}{\rho(\rho+1)} + O(x^{1/2}).$$

Now

$$\left|\frac{1}{\rho} - \frac{1}{i\gamma}\right| = \left|\frac{-1/2}{i1/2\gamma - \gamma^2}\right| = O\left(\frac{1}{\gamma^2}\right)\,,$$

so

$$(1.9) = -ix^{1/2} \sum_{\rho} \frac{\sin \gamma\delta}{\delta} \cdot \frac{x^{i\gamma}}{\rho+1} \left(\frac{1}{i\gamma} + O\left(\frac{1}{\gamma^2}\right)\right) + O(x^{1/2})\,.$$

We apply the estimate we have used before for the sum over ρ to deduce that the error term from replacing $\frac{1}{\rho}$ with $\frac{1}{i\gamma}$ is $O(x^{1/2})$. We can do the same thing again with the $(\rho+1)$ term in the denominator and obtain

$$(1.9) = -x^{1/2}\sum_{\rho}\frac{\sin\gamma\delta}{\gamma\delta}\cdot\frac{x^{i\gamma}}{i\gamma} + O(x^{1/2})\,.$$

The zeros of the zeta function come in conjugate pairs so when we take the sum over $\gamma > 0$, we add a factor of 2, and we are done. □

Estimating the sum in (1.5) requires knowing some information about the imaginary parts of the zeros of the zeta function. The constant δ is fixed but $\sin\gamma\delta$ oscillates depending on the position of γ modulo 2π. If $\gamma\log x$ is "close enough" to $\gamma\delta$ for $|\gamma| < N$, then we need only to estimate the simpler sum

$$\delta\sum_{\gamma<N}\left|\frac{\sin\gamma\delta}{\gamma\delta}\right|^2$$

which, for small enough δ, will obey $\sin\gamma\delta \approx \gamma\delta$, and reduce to the still simpler

$$\frac{1}{\delta}\sum_{\gamma<N}1\,.$$

The first obstacle here is to pick an x such that $\gamma\log x$ is close to $\gamma\delta$. That we can do this is guaranteed by the following lemma due to Dirichlet. It turns out that one does not need to know more than the number of zeros up to a certain height.

Lemma 1.3. *(Dirichlet)*
If $x_1,\ldots,x_K$ are real numbers, and N is a positive integer, then there is a positive integer $n \le N^K$ such that $\|x_k n\| < 1/N$ for $1 \le k \le K$.

Proof. We know that $n = 0, 1, \ldots, N^K$ N^K+1, different values. Let $\{x_k n\}$ denote the fractional part of $x_k n$, then

$$p(n) = (\{x_1 n\}, \{x_2 n\}, \ldots, \{x_K n\}) \in [0,1)^K\,.$$

We can partition $[0,1]^K$ into N^K smaller cubes with sides of length $1/N$. If we let n run over N^K+1 possibilities then there would be N^K+1 number of $p(n)'s$. By the pigeonhole principle there must be two $p(n)'s$ in the same hypercube, say for n_1 and n_2. Write, for each k, $x_k n_1 = [x_k n_1] + \{x_k n_1\}$, then

$$\begin{aligned}\|x_k n_1 - x_k n_2\| &= \|([x_k n_1] - [x_k n_2]) + (\{x_k n_1\} - \{x_k n_2\})\|\\ &= \|\{x_k n_1\} - \{x_k n_2\}\| \le |\{x_k n_1\} - \{x_k n_2\}| < 1/N\,,\end{aligned}$$

since $|p(n_1) - p(n_2)| \le 1/N$. Assume $n_2 > n_1$, we take $n = n_2 - n_1$ and we are done. □

Now we are ready to prove Littlewood's theorem.

Proof. (Littlewood's Theorem) We again assume RH and use equation (1.5). Let N be an integer which we will take to be very large. Let $T = N\log N$ be the height to which we would like to consider the zeta zeros, and let

$$\{\gamma_1(\log N)/2\pi, \gamma_2(\log N)/2\pi, \ldots, \gamma_{N(T)}(\log N)/2\pi\}$$

be a collection of $N(T)$ real numbers, where $N(T)$ is the number of zeta zeros with imaginary parts $\leq T$. By Lemma 1.3, there exists an integer $1 \leq n \leq N^{N(T)}$ such that $\|\gamma_k(\log N)/2\pi \cdot n\| < 1/N$ and $1 \leq k \leq N(T)$.
Take $x = N^n e^{\pm 1/N}$, $\delta = 1/N$, we use the inequality

$$|\sin 2\pi\alpha \pm \sin 2\pi\beta| \leq 2\pi\|\alpha \pm \beta\|.$$[8]

By this identity, we have

$$\begin{aligned}\left|\sin\gamma\log x \pm \sin\frac{\gamma}{N}\right| &\leq 2\pi\left\|\frac{\gamma\log x \pm \gamma/N}{2\pi}\right\| \\ &= 2\pi\left\|\frac{\gamma(n\log N \pm 1/N) \pm \gamma/N}{2\pi}\right\| \\ &= 2\pi\left\|\frac{\gamma\log Nn}{2\pi}\right\| \leq \frac{2\pi}{N}\end{aligned}$$

for all γ up to height T. Consider the right hand side of (1.5) and substitute δ for $1/N$,

$$-2x^{1/2}\sum_{\gamma>0}\frac{\sin\gamma/N}{\gamma/N}\cdot\frac{\sin\gamma\log x}{\gamma} + O(x^{1/2}).$$

We consider its difference and sum with a similar sum of $\sin\gamma/N$ over ρ and estimate the error:

$$\begin{aligned}&\left|(1.5) \mp 2x^{1/2}N^{-1}\sum_{\gamma>0}\left(\frac{\sin\gamma/N}{\gamma/N}\right)^2\right| \\ &= 2x^{1/2}\sum_{0<\gamma<T}\left|\frac{\sin\gamma/N}{\gamma/N}\cdot\frac{\sin\gamma\log x \mp \sin\gamma/N}{\gamma}\right| && \text{(S1)} \\ &+ 2x^{1/2}\sum_{\gamma>T}\left|\frac{(\sin\gamma/N)^2}{\gamma/N\cdot\gamma}\right| && \text{(S2)} \\ &+ 2x^{1/2}\sum_{\gamma>T}\left|\frac{\sin\gamma/N}{\gamma/N}\cdot\frac{\sin\gamma\log x}{\gamma}\right|. && \text{(S3)}\end{aligned}$$

We consider (S1) first,

$$\text{(S1)} \leq x^{1/2}\sum_{0<\gamma<T}\left|\frac{\sin\gamma/N}{\gamma^2}\right|\frac{2\pi/N}{1/N} = O(x^{1/2}).$$

[8] Use the sum-to-product formulas we get

$$|\sin 2\pi\alpha \pm \sin 2\pi\beta| \leq |2\sin\pi(\alpha\pm\beta)| \leq 2\pi\|\alpha\pm\beta\|.$$

as we know from the Introduction that $\sum_{\gamma>0} 1/\gamma^2 < \infty$. For (S2), (S3) we note that

$$\sum_{\gamma>T} \frac{1}{\gamma^2} \ll \frac{\log T}{T} \ll \frac{\log N + \log\log N}{N \log N} \ll \frac{1}{N} .$$

by Lemma 0.5. Thus these two terms are both $O(x^{1/2})$ as well, and we may write

$$(1.5) = \pm 2x^{1/2} N^{-1} \sum_{\gamma>0} \left(\frac{\sin \gamma/N}{\gamma/N} \right)^2 + O(x^{1/2}) \tag{1.12}$$

when $x = N^n e^{\pm 1/N}$ and $\delta = 1/N$. Now we show that this quantity is strictly greater than $x^{1/2} \log\log\log x$ for suitable choices of N.
First we consider the sum of sines. This can be split into two:

$$N^{-1} \sum_{\gamma>0} \left(\frac{\sin \gamma/N}{\gamma/N} \right)^2 = N^{-1} \left[\sum_{\gamma<T} \left(\frac{\sin \gamma/N}{\gamma/N} \right)^2 + \sum_{\gamma>T} \left(\frac{\sin \gamma/N}{\gamma/N} \right)^2 \right] .$$

As was discussed previously,

$$\begin{aligned} &\asymp N^{-1}(N \log N + O(N)) \\ &\asymp \log N + O(1) . \end{aligned}$$

We know that $N(T) \asymp T \log T \asymp N (\log N)^2$, so that

$$\begin{aligned} &\log x \leq N^{N(T)} \log N + 1/N \\ &\Rightarrow \log x = O(N^{N(T)} \log N) \\ &\Rightarrow \log\log x = O(N(T) \log N + \log\log N) . \end{aligned}$$

subsituting the asymptotic estimate for $N(T)$, we get

$$\begin{aligned} &\Rightarrow \log\log\log x \ll \log N + 3 \log\log N \\ &\Rightarrow \log\log\log x < C \log N . \end{aligned}$$

for some absolute constant $C > 0$. Now we divide the expression (1.12) by $x^{1/2} \log\log\log x$ to get

$$\frac{(1.12)}{x^{1/2} \log\log\log x} \asymp \frac{\pm 2x^{1/2} \log N + O(x^{1/2})}{x^{1/2} \log\log\log x} .$$

As $x \to \infty$, the term bounded by $x^{1/2}$ goes to zero, and we have

$$\asymp \pm \frac{2}{C} . \qquad \square$$

By splitting the sum into two parts, we see that the contributions of the zeros over a certain height, namely N, are trivial, and for a very specific value of x, this sum can be approximated by a term of order $\log N$. Passing the calculation to the integrated version of the explicit formula for $\psi(x)$ is helpful since $\sum_{\gamma>0} \frac{1}{\gamma^2}$ converges.

A point of interest here is whether this formula yields a numerical value for x. Although a choice of x is explicitly stated in the proof, the use of the number from Dirichlet's lemma is merely existential. Also from this choice we can already see how big this number would be, as verified by Skewes efforts to find a numerical value. One might also be interested in a computational formula based on Littlewood's method. However, his method involves assuming that the contribution from the square factors of x is always non-neligible, making the switch from $\psi(x)-x$ to $\pi(x)-\mathrm{li}(x)$ imprecise. He also selected his x by assuming that the list of zeta zeros is an arbitrary list. This assumption is not quite accurate as we will see from the much smaller upper bound found by Lehman by actually entering the precise values for the zeta zeros in his computational formula. Furthermore, finding an efficient algorithm for the elusive Dirichlet's number will be required for a direct implementation.

What about the case when RH is not true? As Littlewood said in his paper, "one already knows more."[9] Indeed, the proof in the case where the Riemann Hypothesis is false is much simpler. The inspiration comes from being able to express the Mellin transform of the related functions in terms of the zeta function itself. Let Θ be the supremum of the real part of the zeros of the zeta function, and $\epsilon > 0$. We can write

$$\frac{1}{s-\Theta+\epsilon} + \frac{\zeta'}{\zeta}(s) + \frac{1}{s-1} = \int_1^\infty (x^{\Theta-\epsilon} - \psi(x) + x) x^{-s-1}\, dx\,.$$

If $x^{\Theta-\epsilon} > \psi(x)-x$ for all x, then the expression on the left is analytic for $s > \Theta-\epsilon$, which contradicts the assumption that there is a zero of $\zeta(s) > \Theta - \epsilon$ by the definition of supremum. A rigorous proof of the above statement requires deeper theoretical foundation than the case where the Riemann Hypothesis is true. In particular, we will need to develop more theory about the general Dirichlet series and examine their Mellin transforms. Also in contrast to Littlewood's theorem, this method of proof gives not even a number x in principal where the sign change might occur.

Consider $\zeta(s)$ on the half plane $\mathrm{Re}(s) > 1$, we can write

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1} \quad \text{(The Euler Product)}$$

Since each of the factors is a holomorphic function in the half plane, and the product on the right converges to to $\zeta(s)$ locally uniformly on compact sets, we can write

$$\log \zeta(s) = \sum_{p \text{ prime}} \log \left(1 - \frac{1}{p^s}\right)^{-1}$$

and differentiate the sum on the right term by term to obtain the logarithmic

[9] "on sait dèjá plus".

derivative of $\zeta(s)$:

$$\begin{aligned}
\frac{\zeta'}{\zeta}(s) &= \frac{d}{ds}\left[\sum_{p \text{ prime}} \log\left(1-\frac{1}{p^s}\right)^{-1}\right] \\
&= -\sum_{p \text{ prime}} \log p \left(\frac{p^{-s}}{1-p^{-s}}\right) \\
&= -\sum_{p \text{ prime}} \log p \left(\frac{1}{p^s}+\frac{1}{p^{2s}}+\cdots\right) \\
&= -\sum_{n=1}^{\infty} \Lambda(n) n^{-s}\,,
\end{aligned}$$

since $\Lambda(n)$ vanishes at any integer which is not a prime power. Then we use the Stieltjes integral to write the partial sum

$$\begin{aligned}
\sum_{n=0}^{N} \Lambda(n) n^{-s} &= \int_1^N x^{-s}\, d\psi(x) \\
&= |\psi(x)x^{-s}]_1^N - \int_1^N \psi(x)\, dx^{-s} \\
&= \psi(N)N^{-s} + s\int_1^N \psi(x)x^{-s-1}\, dx\,.
\end{aligned}$$

Since we know $\psi(N) \sim N$ as $N \to \infty$, the first term goes to zero, and we have only the integral left when $\mathrm{Re}(s) > 1$, and

$$-\frac{\zeta'}{\zeta}(s) = s\int_1^{\infty} \psi(x)x^{-s-1}dx\,. \tag{1.13}$$

Since the meromorphic continuation of a function is unique, wherever the integral on the right hand side is defined outside of the half plane $\mathrm{Re}(s) > 1$, the meromorphic continuation of $\frac{\zeta'}{\zeta}(s)$ must also be defined. We already have some information on the real parts of the zeros of the Riemann zeta function, so this restricts the magnitude of the function multiplying x^{-s-1}. Suppose that $\psi(x) - x \leq x^{\Theta-\epsilon}$ for all x and an $\epsilon > 0$ after a certain X_ϵ, then we can consider

$$\int_0^{\infty} (x^{\Theta-\epsilon} - \psi(x) + x)x^{-s-1}\, dx\,. \tag{1.14}$$

On the left hand side we have

$$\frac{1}{s-\Theta+\epsilon} + \frac{1}{s}\frac{\zeta'}{\zeta}(s) + \frac{1}{s-1}\,. \tag{1.15}$$

If we can prove that the integral converges where there is a Riemann zeta zero, there would be a contradiction. The other statement $\psi(x) - x = \Omega_-(x^{\Theta-\epsilon})$ can be proved in an entirely analogous way. There are several issues here. First is the validity of the formula like (1.13) if we were to use the same set up to prove the equivalent statement for $\pi(x) - \mathrm{li}(x)$. Second is determining where the integral (1.14) converges. To prove these things for each instance of Dirichlet series would be cumbersome, so we will include the results for general Dirichlet series here, starting with a definition.

Definition. A *Dirichlet series* is a series of the form $\alpha(s) = \sum_{n=1}^{\infty} a_n n^{-s}$.

Theorem 1.4. *Any Dirichlet series $\alpha(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ has an abscissa of convergence σ_c with the property that $\alpha(s)$ converges for all s with* $\mathrm{Re}(s) > \sigma_c$*, and for no s with* $\mathrm{Re}(s) < \sigma_c$.

Proof. We prove that for any s_0 where the Dirichlet series converges, the series is also convergent in the sector

$$S = \{s : \sigma \geq \sigma_0, |t - t_0| \leq C(\sigma - \sigma_0)\},$$

where C is any constant greater than 0.
Let $R(N) = \sum_{n>N} a_n n^{-s_0}$ be the tail of $\alpha(s_0)$. Write $a_n = (R(n-1) - R(n))n^{s_0}$, then

$$\sum_{n=M+1}^{K} a_n n^{-s} = \sum_{n=M+1}^{K} (R(n-1) - R(n))n^{s_0 - s}.$$

we take the first and the final term of the sum

$$\begin{aligned}
&= R(M)M^{s_0-s} - R(K)K^{s_0-s} \\
&+ \left[\sum_{n=M+1}^{K} (R(n-1) - R(n))n^{s_0-s} - R(M)M^{s_0-s} + R(K)K^{s_0-s}\right] \\
&= R(M)M^{s_0-s} - R(K)K^{s_0-s} - \sum_{n=M+1}^{K} R(n-1)((n-1)^{s_0-s} - n^{s_0-s}) \\
&= R(M)M^{s_0-s} - R(K)K^{s_0-s} - \sum_{n=M+1}^{K} R(n-1)(s_0 - s)\int_{n-1}^{n} x^{s_0-s-1}\,dx.
\end{aligned}$$

Since $\alpha(s_0)$ converges, with any $\epsilon > 0$, there exists an N' such that $R(N) < \epsilon$ for all $N > N'$. Let M, L be such that $M, K > N'$, we have

$$\begin{aligned}
\left|\sum_{n=M+1}^{K} a_n n^{-s}\right| = \left|R(M)M^{s_0-s} - R(K)K^{s_0-s}\right| &< 2\epsilon \\
+ \left|\sum_{n=M+1}^{K} R(n-1)(s_0 - s)\int_{n-1}^{n} x^{s_0-s-1}\,dx\right| &< \epsilon|(s_0 - s)|\int_{M}^{\infty} x^{\sigma_0-\sigma-1}\,dx.
\end{aligned}$$

Since $\sigma_0 < \sigma$ we have

$$\int_{M}^{\infty} x^{\sigma_0-\sigma-1}dx = -\frac{M^{\sigma_0-\sigma}}{\sigma_0 - \sigma} < \frac{1}{\sigma_0 - \sigma}.$$

Thus

$$\left|\sum_{n=M+1}^{K} a_n n^{-s}\right| < \epsilon\left(2 + \frac{|s_0 - s|}{\sigma_0 - \sigma}\right).$$

For $s \in S$, $|t - t_0| \leq C(\sigma - \sigma_0)$

$$\begin{aligned}
|s - s_0| &\leq |\sigma - \sigma_0 + i(t - t_0)| \\
&\leq |\sigma - \sigma_0| + |t - t_0| \\
&\leq (C + 1)(\sigma - \sigma_0).
\end{aligned}$$

Thus $\left|\sum_{n=M+1}^{K} a_n n^{-s}\right|$ can be made arbitrarily small with the right choice of N' for which $M, K > N'$. We can let K tend to infinity to obtain our statement. Now it is a straightforward matter to deduce the theorem since for any s_0 where $\alpha(s_0)$ converges, there exists a constant C such that a point s with $\mathrm{Re}(s) > \mathrm{Re}(s_0)$ belongs to the set S. We take the infimum over all s for which $\alpha(s)$ to obtain σ_c. □

Theorem 1.5. *Let* $A(x) = \sum_{n \le x} a_n$. *If* $\sigma_c \ge 0$, *then the equation*

$$\sum_{n=1}^{\infty} a_n n^{-s} = s \int_1^{\infty} A(x) x^{-s-1} \, dx \tag{1.16}$$

holds for $\mathrm{Re}(s) > \sigma_c$, *and*

$$\lim_{x \to \infty} \sup \frac{\log |A(x)|}{\log x} = \sigma_c \,. \tag{1.17}$$

Proof. We can use the Stieltjes integral to write

$$\sum_{n=1}^{\infty} a_n n^{-s} = \int_1^{\infty} x^{-s} \, dA(x) = \left[A(x) \frac{1}{x^s} \right]_1^{\infty} + s \int_0^{\infty} A(x) x^{-s-1} \, dx \,.$$

This is reminiscent of the formula derived for $\psi(x)$, except this is in a general context. Evaluating the first term requires estimating the magnitude of $A(x)$. Let ϕ denote the left hand side of (1.17), if $\theta > \phi$, then $A(x) = O(x^{\theta})$. Thus if $\phi = \sigma_c$, the term $A(x)x^{-s-1}$ goes to 0 as $x \to \infty$ if $\mathrm{Re}(s) > \sigma_c$. We need only to prove (1.17). The series $\alpha(s)$ diverges when $\mathrm{Re}(s) < \sigma_c$, thus the right hand side of (1.16) cannot converge, thereby $\phi \ge \sigma_c$. So we show $\phi \le \sigma_c$. For any $\epsilon > 0$, let $R(N) = \sum_{n > N} a_n n^{-\sigma_c - \epsilon}$ and write $a_n = (R(n-1) - R(n)) n^{-\sigma_c - \epsilon}$, we repeat the process from before

$$A(N) = \sum_{n \le N} a_n = -R(N) N^{\sigma_c + \epsilon} + (\sigma_c + \epsilon) \int_0^N R(u) u^{\sigma_c + \epsilon - 1} \, du$$

Since $\alpha(\sigma_c + \epsilon)$ converges $R(N)$ is bounded for all N. Thus $A(x) = O(x^{\sigma_c + \epsilon})$ for every $\epsilon > 0$, and $A(x) = O(x^{\sigma_c})$. □

Now we can prove

Theorem 1.6.

$$\psi(x) - x = \Omega_{\pm}(x^{\Theta - \epsilon}) \,.$$

Proof. We combine (1.14), (1.15) and write

$$\frac{1}{s - \Theta + \epsilon} + \frac{1}{s} \frac{\zeta'}{\zeta}(s) + \frac{1}{s-1} = \int_0^{\infty} (x^{\Theta - \epsilon} - \psi(x) + x) x^{-s-1} \, dx \quad \text{for } \mathrm{Re}(s) > 1. \tag{1.18}$$

By assumption $\psi(x) - x < x^{\Theta-\epsilon}$ for all $x > X_\epsilon$. Thus the integral can be split into two

$$\int_0^{X_\epsilon} (x^{\Theta-\epsilon} - \psi(x) + x)x^{-s-1}\,dx + \int_{X_\epsilon}^{\infty} (x^{\Theta-\epsilon} - \psi(x) + x)x^{-s-1}\,dx\,.$$

The first integral is entire, and the second integral $\leq \int_{X_\epsilon}^{\infty} x^{\Theta-\epsilon-s-1}dx$, which converges when $s > \Theta - \epsilon$. Thus the function defined by the right hand side of (1.18) is entire on the half plane $\mathrm{Re}(s) > \Theta - \epsilon$. By the uniqueness of meromorphic extension, the left hand side of (1.18) must also converge on this half plane as well. This is a contradiction since Θ is the supremum of all zeros of the zeta function. Therefore the left hand side has a pole in the half plane $\Theta - \epsilon$. We apply the same argument with the assumption $\psi(x) - x > -x^{\Theta-\epsilon}$ for all $x > X_\epsilon$ to

$$\frac{1}{s+\Theta+\epsilon} - \frac{1}{s}\frac{\zeta'}{\zeta}(s) - \frac{1}{s-1} = \int_1^{\infty} (x^{\Theta-\epsilon} + \psi(x) - x)x^{-s-1}\,dx\,. \qquad \square$$

We consider the following integral of $\mathrm{li}(x)$, and choose $v = \mathrm{li}(x)$, $du = x^{-s-1}$

$$s\int_2^{\infty} \mathrm{li}(x)x^{-s-1}\,dx = \big[-\,\mathrm{li}(x)x^{-s}\big]_2^{\infty} + \int_2^{\infty} \frac{1}{x^s \log x}\,dx\,.$$

The first term evaluates to 0, and with a subsitution $u = \log x$

$$= \int_{\log 2}^{\infty} \frac{1}{e^{(s-1)u}u}\,du\,.$$

Write $w = (s-1)u$,

$$\begin{aligned}
&= \int_{(s-1)\log 2}^{\infty} \frac{1}{e^w w}\,dw \\
&= \int_{(s-1)\log 2}^{1} \frac{e^{-w}-1}{w}\,dw + \int_1^{\infty} \frac{e^{-w}}{w}\,dw + \int_{(s-1)\log 2}^{1} \frac{1}{w}\,dw \\
&= \int_{(s-1)\log 2}^{1} \frac{e^{-w}-1}{w}\,dw + \int_1^{\infty} \frac{e^{-w}}{w}\,dw - \log(s-1) - \log\log 2\,.
\end{aligned}$$

By integration by parts,

$$\int_0^1 \frac{e^{-w}-1}{w}\,dw + \int_1^{\infty} \frac{e^{-w}}{w}\,dw = \int_0^{\infty} e^{-w}\log w\,dw = \Gamma'(1) = C$$

for C a constant. Use this in our formulation of the Mellin transform of $\mathrm{li}(x)$ to obtain

$$= -\int_0^{(s-1)\log 2} \frac{e^{-w}-1}{w}\,dw - C - \log\log 2 - \log(s-1)\,. \tag{1.19}$$

Note that the integrand is discontinuous at 0, but we can always ignore a single point of discontinuity. The integral thus converges everywhere and is an entire

function.
Now we consider $J(x)$ and write it slightly differently,

$$J(x) = \sum_{n\leq x} \frac{\Lambda(n)}{\log n}\,.$$

By Theorem 1.5 we have

$$s\int_2^\infty J(x)x^{-s-1}dx = \log\zeta(s) \quad (\sigma > 1).$$

Thus we can write

$$\int_2^\infty (x^{\Theta-\epsilon} - J(x) + \mathrm{li}(x))x^{-s-1}\,dx = \frac{1}{s-\Theta+\epsilon} - \frac{1}{s}\log(\zeta(s)(s-1)) + \frac{r(s)}{s}$$

for $\mathrm{Re}(s) > 1$. The function $A(x) = x^{\Theta-\epsilon} - J(x) + \mathrm{li}(x)$ is bounded in a finite interval $1 \leq x \leq X$, and is positive by our assumption that $J(x) - \mathrm{li}(x) < x^{\Theta+\epsilon}$. In comparison to x^{-s-1} the integral converges if $-s-1+\Theta-\epsilon < -1$, or for $\mathrm{Re}(s) > \Theta - \epsilon$ by Landau's theorem. By the uniqueness of meromorphic continuation the function on the right hand side must converge in the half plane as well, but this is a contradiction since by assumption there exists a zero of the zeta function with real part greater than $\Theta-\epsilon$. Thus $J(x) - \mathrm{li}(x)$ must exceed $x^{\Theta-\epsilon}$. On the other hand if $\psi(x) - x > -x^{\Theta-\epsilon}$ for all $x > X_0(\epsilon)$, then we can repeat the argument with the following integral

$$\int_2^\infty (x^{\Theta-\epsilon} + J(x) - \mathrm{li}(x)x^{-s-1})\,dx = \frac{1}{s-\Theta+\epsilon} + \frac{1}{s}\log(\zeta(s)(s-1)) - \frac{r(s)}{s}\,.$$

Now the switch to $\pi(x)$ is the reason why this argument will not work if we assume that the Riemann Hypothesis is true. By definition $J(x) = \pi(x) + O(x^{1/2}/\log x)$, and we have

$$J(x) - \mathrm{li}(x) = \Omega_+(x^{\Theta-\epsilon})\,.$$

This means

$$\lim_{x\to\infty}\sup\frac{J(x)-\mathrm{li}(x)}{x^{\Theta-\epsilon}} > 0$$
$$\Leftrightarrow \inf\left\{\sup\left\{\frac{J(x)-\mathrm{li}(x)}{x^{\Theta-\epsilon}} : x > n\right\} : n > 0\right\} > 0\,.$$

We use $J(x) = \pi(x) + O(x^{1/2}/\log x)$,

$$\Leftrightarrow \inf\left\{\sup\left\{\frac{\pi(x)-\mathrm{li}(x)+O(x^{1/2}/\log x)}{x^{\Theta-\epsilon}} : x > n\right\} : n > 0\right\} = \lambda > 0\,.$$

Let ϵ be such that $\Theta - \epsilon > 1/2$, it is always possible to find this number because the Riemann Hypothesis is assumed to be false. Then the difference between $\pi(x) - \mathrm{li}(x)/x^{\Theta-\epsilon}$ and $O(x^{1/2}/\log x)/x^{\Theta-\epsilon}$ is greater than or equal to λ. Pick δ such that $\delta < \lambda$, then there exists some N such that

$$\frac{-O(x^{1/2}/\log x)}{x^{\Theta-\epsilon}} > -\delta \quad \text{for all } x > N,$$

and there exists some $x > N$ such that

$$\begin{aligned}\frac{\pi(x) - \mathrm{li}(x) + O(x^{1/2}/\log x)}{x^{\Theta-\epsilon}} &> \lambda \\ &> \lambda - \delta > 0\,.\end{aligned}$$

We have that

$$\pi(x) - \mathrm{li}(x) = \Omega_+(x^{\Theta-\epsilon})$$

is true for $0 < \epsilon < \Theta - 1/2$, but also trivially true for all $\epsilon' > \epsilon$, since then $x^{\Theta-\epsilon'}$ is smaller than $x^{\Theta-\epsilon}$. The case for

$$\pi(x) - \mathrm{li}(x) = \Omega_-(x^{\Theta-\epsilon})$$

proceeds in exactly the same way.

This result is stronger than Littlewood's theorem in the sense that the bound $\Omega_\pm(x^{1/2}\log\log\log x)$ can be easily derived from this result. To derive Littlewood's theorem, we used that $\pi(x)$ is equal to $\frac{\theta(x)}{\log x}$ within an order that is bounded by $x^{1/2}$, and then similarly, that $\theta(x)$ is $\psi(x)$ within an error bounded by $x^{1/2}$. This error comes from the number of repeated factors in an integer, and we proceed with assuming that this error always exists. Also the growth of this difference is very little ($\log\log\log x$), nevertheless it is there. For the proof assuming the falsity of the Riemann Hypothesis, there is a similar switch involving a term of order $x^{1/2}$ which is easily overpowered by x^{Θ}. The result on Dirichlet series (Theorem 1.5) is merely a statement about the magnitude of the difference compared to x^{-s-1}, and it is easier to work with $J(x)$ since there is an easily identifiable representation of $J(x)$ in terms of the series of the zeta function.

2 Lehman's approach

Instead of considering the difference $\psi(x) - x$, Lehman chose to work with the explicit formula for $\pi(x) - \mathrm{li}(x)$ as originally conceived in Riemann's paper. The goal is to develop a computational formula for the difference, which depends, to sufficient accuracy, on only a finite number of zeta zeros. It is easy to see that for any positive function $K(y)$ and any interval (a, b),

$$\int_a^b K(y)\{\pi(y) - \mathrm{li}(y)\}\,dy > 0 \tag{2.1}$$

implies that $\pi(y) - \mathrm{li}(y)$ is positive within (a, b). Lehman showed that with a suitable selection of $K(y)$, the computation of this integral only required the first 12,000 zeta zeros to yield a dramatically improved upper bound of Skewes' number. Since Lehman's publication, other mathematicians have used his formula with minor modifications to yield the best upper bound that is currently known.

The crucial theorem due to Lehman is the following:

Theorem 2.1. *Let $\rho = \beta + i\gamma$ denote a zero of the zeta function in the critical strip, and let A be a positive number such that $\beta = \frac{1}{2}$ for all zeros ρ with imaginary* *parts* $0 \leq \gamma \leq A$. *Let* α, η, *and* ω *be positive numbers such that* $\omega - \eta > 1$ *and the conditions*

$$4A/\omega \leq \alpha \leq A^2, \tag{2.2}$$
$$2A/\alpha \leq \eta < \omega/2, \tag{2.3}$$

hold. Let

$$K(y) = \sqrt{\frac{\alpha}{2\pi}} e^{-\alpha y^2/2}. \tag{2.4}$$

Then for $2\pi e < T \leq A$ *,*

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega) u e^{-u/2} \{\pi(e^u) - \mathrm{li}(e^u)\}\, du = -1 - \sum_{0<|\gamma|\leq T} \frac{e^{i\gamma w}}{\rho} e^{-\gamma^2/2\alpha} + R\,, \tag{2.5}$$

where

$$|R| \leq \frac{3.05}{\omega - \eta} + 4(\omega+\eta)e^{-(\omega-\eta)/6} + \frac{2e^{-\alpha\eta^2/2}}{\sqrt{2\pi\alpha}\eta} + 0.08\sqrt{\alpha}e^{-\alpha\eta^2/2} \tag{2.6}$$
$$+ e^{-T^2/2\alpha}\left(\frac{\alpha}{\pi T^2}\log\frac{T}{2\pi} + \frac{8\log T}{T} + \frac{4\alpha}{T^3}\right) \tag{2.7}$$
$$+ A\log A e^{-A^2/2\alpha+(\omega+\eta)/2}(4\alpha^{-1/2} + 15\eta)\,. \tag{2.8}$$

If we ignore conditions (2.2) and (2.3) for the moment and consider the error term R, we can see that it is desirable to integrate over a large interval $(\omega-\eta, \omega+\eta)$ to reduce the error in (2.6), and include as many zeta zeros in the exponential sum of (2.5) to reduce the second term in R. Finally, (2.8) disappears if the Riemann Hypothesis is true. Given the current computational power and the number of zeros that have been verified to lie on the critical line, the error term is negligible and the computation reduces down to a simple sum. To give a concrete example, the calculation done by Bays and Hudson gives, for

$$A = 10^7, \quad \alpha = 10^{10}, \quad \eta = .002,$$
$$T = \gamma_1,000,000 = 600269.677\ldots, \quad \omega = 77.95209,$$

the following bounds on $|R|$:

$$(2.6) < .00418985 + 5.94\times 10^{-50} + 3.64\times 10^{-8689} + 1.03\times 10^{-8682} < .00418986$$
$$(2.7) < 1.53\times 10^{-9},$$
$$(2.8) < 1.20\times 10^{-2001}.$$

Since the sum $\sum_{0<|\gamma|\leq T} \frac{e^{i\gamma w}}{\rho} e^{-\gamma^2/2\alpha}$ under these constraints is $1.012762\ldots$ with error smaller than .0078, this value establishes the existence of a crossover in the interval $(1.398201\times 10^{316}, 1.398244\times 10^{316})$.

Lehman's method would have been impractical if the Riemann Hypothesis had not been verified to such a large number of zeros, or if large-scale computations were not possible. The breakthrough was a real triumph of modern computational mathematics. Unfortunately, the constraints imposed on the parameters A, ω, α, and η meant that there would be gaps in the intervals where this theorem cannot be applied. An effort to relax or clarify the constraints (2.2), and (2.3) should prove useful in further investigation of Skewes' number.

The proof.
We let $\mathrm{li}(x)$ denote the principal value of the logarithmic integral[10]. We recall that

$$J(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) \dots , \tag{2.9}$$

The Riemann-von Mangoldt explicit formula states

$$J(x) = \mathrm{li}(x) - \sum_{\rho} \mathrm{li}(x^{\rho}) + \int_x^{\infty} \frac{du}{(u^2-1)u \log u} - \log 2 \quad (x > 1)\,. \tag{2.10}$$

Then we can use two estimates on $\pi(x)$:

$$\begin{aligned} \pi(x) - \frac{x}{\log x} &< \frac{3/2}{(\log x)^2}, \text{ and} \\ \pi(x) &< \frac{2x}{\log x} \qquad (x > 1). \end{aligned}$$

Define

$$\begin{aligned} \vartheta_1(x) &= \frac{\pi(x) - \frac{x}{\log x}}{\frac{3/2}{(\log x)^2}}, \text{ and} \\ \vartheta_2(x) &= \frac{\pi(x)}{\frac{2x}{\log x}}\,. \end{aligned}$$

Then both $\vartheta_1(x), \vartheta_2(x) < 1$, and we can conveniently express $\pi(x)$ thus:

$$\begin{aligned} \pi(x) &= \frac{x}{\log x} + \frac{\vartheta_1(x)3/2x}{(\log x)^2}, \text{ and} \\ \pi(x) &= \frac{\vartheta_2(x)2x}{\log x}\,. \end{aligned} \tag{2.11}$$

Since $\pi(x^{1/k}) = 0$ when $x^{1/k} < 2$, this tells us that there are at most $\frac{\log x}{\log 2}$ terms in $J(x)$, so

$$\begin{aligned} J(x) &< \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) \left(\frac{\log x}{\log 2}\right), \text{ and} \\ J(x) &= \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\vartheta_3(x)\pi(x^{1/3}) \left(\frac{\log x}{\log 2}\right), \end{aligned}$$

where $\vartheta_3(x)$ is defined in the same way as $\vartheta_1(x)$ and $\vartheta_2(x)$ and does not exceed 1. Combining (2.9) and (2.10), we have

$$\begin{aligned} \pi(x) - \mathrm{li}(x) = &- \sum_{\rho} \mathrm{li}(x^{\rho}) + \int_x^{\infty} \frac{du}{(u^2-1)u \log u} - \log 2 \\ &- \frac{1}{2}\pi(x^{1/2}) - \frac{1}{3}\vartheta_3(x)\pi(x^{1/3}) \left(\frac{\log x}{\log 2}\right)\,. \end{aligned}$$

[10] see Introduction, page 8.

Substituting the first expression in (2.11) for $\pi(x^{1/2})$ and the second expression for $\pi(x^{1/3})$, we get

$$\pi(x) - \mathrm{li}(x) = -\sum_\rho \mathrm{li}(x^\rho) + \int_x^\infty \frac{du}{(u^2-1)u\log u} - \log 2$$
$$- \left(\frac{x^{1/2}}{\log x} + 3\frac{\vartheta_1(x^{1/2})x^{1/2}}{(\log x)^2}\right) - \vartheta_3(x)\left(\frac{\vartheta_2(x^{1/3})2x^{1/3}}{\log x}\right)\left(\frac{\log x}{\log 2}\right).$$

The occurrence of multiple estimating functions is not a serious problem. Indeed, for Lehman's purpose, $\vartheta_1(x^{1/2}), \vartheta_2(x^{1/3})$, and $\vartheta_3(x)$ can be absorbed into a single estimate $\vartheta(x)$ with modulus < 1, since, by above,

$$3\frac{\vartheta_1(x^{1/2})x^{1/2}}{(\log x)^2} + \frac{2\vartheta_3(x)\vartheta_2(x^{1/3})x^{1/3}}{\log x}\left(\frac{\log x}{\log 2}\right) < \frac{3x^{1/2}}{(\log x)^2} + \frac{2x^{1/3}}{\log 2}, \text{ and thus}$$
$$3\frac{\vartheta_1(x^{1/2})x^{1/2}}{(\log x)^2} + \frac{\vartheta_3(x)\vartheta_2(x^{1/3})x^{1/3}}{\log x}\left(\frac{\log x}{\log 2}\right) = \vartheta(x)\left(\frac{3x^{1/2}}{(\log x)^2} + \frac{2x^{1/3}}{\log 2}\right).$$

We have,

$$\int_x^\infty \frac{du}{(u^2-1)u\log u} < 2\int_x^\infty \frac{du}{u^3} = \frac{1}{x^2} < \log 2 \quad (x > e), \text{ thus}$$
$$\int_x^\infty \frac{du}{(u^2-1)u\log u} - \log 2 < 0 < x^{1/3}\log 2 \quad (x \geq e).$$

Finally,

$$\pi(x) - \mathrm{li}(x) \leq -\sum_\rho \mathrm{li}(x^\rho) - \frac{x^{1/2}}{\log x} + \frac{3x^{1/2}}{(\log x)^2} + \frac{2x^{1/3}}{\log 2} + x^{1/3}\log 2$$
$$\leq -\sum_\rho \mathrm{li}(x^\rho) - \frac{x^{1/2}}{\log x} + \frac{3x^{1/2}}{(\log x)^2} + x^{1/3}\left(\frac{2}{\log 2} + \log 2\right),$$

and since $\left(\frac{2}{\log 2} + \log 2\right) = 3.578\ldots < 4$, we have

$$\pi(x) - \mathrm{li}(x) = -\sum_\rho \mathrm{li}(x^\rho) - \frac{x^{1/2}}{\log x} + \vartheta'(x)\left(\frac{3x^{1/2}}{(\log x)^2} + 4x^{1/3}\right),$$

where $\vartheta'(x)$ is the final estimate derived from the inequality above. The problem with using this formula directly is that it isn't clear how to truncate the series $\sum_\rho \mathrm{li}(x^\rho)$ such that only finitely many zeros are needed in the computation. To overcome this problem Lehman integrates this term against the Gaussian kernel. For the ease of inspecting numerical data he has also scaled the difference so that one would only need to look for values greater than 1 in the resulting exponential sum $\sum_{0<|\omega|\leq T} \frac{e^{i\gamma\omega}}{\rho} e^{-\gamma^2/2\alpha}$.

We will begin with the simpler terms in the difference. Recall that $K(y) =$

$\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha y^2/2}$, and we have

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\{\pi(e^u)-\mathrm{li}(e^u)\}du$$
$$= \int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left[-\frac{e^{u/2}}{u}-\sum_\rho \mathrm{li}(e^{u\rho})+\vartheta'(x)\left(3\frac{e^{u/2}}{u^2}+4e^{u/3}\right)\right]du$$
$$= \int_{\omega-\eta}^{\omega+\eta} -K(u-\omega)du \qquad \text{(S1)}$$
$$+ \int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left(-\sum_\rho \mathrm{li}(e^{u\rho})\right)du \qquad \text{(S2)}$$
$$+ \int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left[\vartheta'(x)\left(3\frac{e^{u/2}}{u^2}+4e^{u/3}\right)\right]du\,. \qquad \text{(S3)}$$

The estimate for the integral (S1) depends on the interval of integration, (S2) provides the oscillatory exponential sum, and (S3) contributes to the error term. Consider (S1), we can write

$$\int_{\omega-\eta}^{\omega+\eta} -K(u-\omega)\,du$$
$$= -\int_{-\infty}^{\infty} K(u-\omega)\,du + \int_{-\infty}^{\omega-\eta} K(u-\omega)\,du + \int_{\omega-\eta}^{\infty} K(u-w)\,du\,.$$

Since $K(u-\omega)$ is the probability density function for the normal distribution centered at ω, $\int_{-\infty}^{\infty} K(u-\omega)du = 1$. With a change of variable we can write

$$\int_{\omega+\eta}^{\infty} K(u-\omega)\,du = \int_{\eta}^{\infty} K(y)\,dy$$
$$\int_{-\infty}^{\omega-\eta} K(u-\omega)\,du = \int_{-\infty}^{-\eta} K(y)\,dy\,.$$

Since $K(y)$ is symmetric about the y-axis, these two sums are the same, and we only need to estimate one of them. Note

$$\frac{d}{dy}\left[\frac{e^{-\alpha y^2/2}}{y}\right] = -\frac{e^{-\alpha y^2/2}}{y^2} - \alpha e^{-\alpha y^2/2}\,.$$

Thus

$$\int_{\eta}^{\infty} K(y)\,dy = \int_{\eta}^{\infty}\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha y^2/2}\,dy$$
$$< -\sqrt{\frac{1}{2\pi\alpha}}\int_{\eta}^{\infty}\frac{d}{dy}\left[\frac{e^{-\alpha y^2/2}}{y}\right]dy$$
$$= -\sqrt{\frac{1}{2\pi\alpha}}\left[\frac{e^{-\alpha y^2/2}}{y}\right]_{\eta}^{\infty}$$
$$= \frac{e^{-\alpha\eta^2/2}}{\sqrt{2\pi\alpha}\eta}\,.$$

Since $\alpha > 0$, and we have

$$(\text{S1}) < -1 + 2\frac{e^{-\alpha\eta^2/2}}{\sqrt{2\pi\alpha}\eta}\,.$$

Also, the integral (S3) becomes

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)\vartheta'(x)\left(\frac{3}{u} + 4ue^{-u/6}\right)\,du\,,$$

and

$$|S3| \le \frac{3}{\omega-\eta} + 4(\omega+\eta)e^{-(\omega-\eta)/6}\,.$$

The most difficult term to estimate is (S2), and we will start by interchanging integration and summation:

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left(-\sum_{\rho}\text{li}(e^{u\rho})\right)\,du$$
$$= -\sum_{\rho}\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\,\text{li}(e^{u\rho})\,du.$$

This is justified since $\sum_\rho \text{li}(e^{u\rho})$ is a convergent series in the finite interval $\omega-\eta < u < \omega+\eta$. Recall

$$\text{li}(e^{u\rho}) := \int_{-\infty+iu\gamma}^{u\beta+iu\gamma} \frac{e^z}{z}\,dz\,.$$

With $z = \rho u - t$ we have

$$\text{li}(e^{u\rho}) = \int_0^\infty \frac{e^{\rho u}\cdot e^{-t}}{\rho u - t}\,dt\,.$$

Choosing $w = (\rho u - t)^{-1}, dv = e^{-t}$, integration by parts gives,

$$\begin{aligned}\int_0^\infty \frac{e^{\rho u}\cdot e^{-t}}{\rho u - t}\,dt &= e^{\rho u}\left[\frac{1}{\rho u} - \int_0^\infty \frac{-e^{-t}}{(\rho u - t)^2}\right] \\ &< e^{\rho u}\left[\frac{1}{\rho u} + \int_0^\infty \frac{e^{-t}}{(\gamma u)^2}\,dt\right] \quad \text{and so,} \\ \text{li}(e^{u\rho}) &= \frac{e^{\rho u}}{\rho u} + \frac{\vartheta''(\rho u)e^{\rho u}}{(\gamma u)^2},\end{aligned}$$

where $\vartheta''(\rho u)$ is a function of ρ and u defined by the inequality. We can plug this into the formula for (S2). Take A to be a number for which the Riemann

Hypothesis holds for zeros where $|\gamma| < A$, then

$$\begin{aligned}
&\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left(-\sum_{\rho}\operatorname{li}(e^{u\rho})\right)du \\
&= -\sum_{\rho}\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)ue^{-u/2}\left[\frac{e^{\rho u}}{\rho u} + \vartheta''(\rho u)\frac{e^{\rho u}}{(\gamma u)^2}\right]du \\
&= -\sum_{0<|\gamma|\leq A}\frac{1}{\rho}\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)e^{i\gamma u}du \qquad \text{(E1)} \\
&- \sum_{0<|\gamma|\leq A}\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)\vartheta''(\rho u)\frac{e^{i\gamma u}}{\gamma^2 u}\,du \qquad \text{(E2)} \\
&- \sum_{|\gamma|>A}\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)\operatorname{li}(e^{u\rho})\,du\,. \qquad \text{(E3)}
\end{aligned}$$

To improve the speed of the calculation we would like to use as few of the zeta zeros as we can while tolerating a certain amount of error. This means that we will have to find some way to estimate the terms that are dropped from the computation. In (E1), we are dealing with the integral

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)e^{i\gamma u}\,du\,.$$

With a change of variable taking $y = u - \omega$, the integral becomes

$$e^{i\gamma\omega}\int_{-\eta}^{\eta} K(y)e^{i\gamma y}\,dy = e^{i\gamma\omega}\int_{-\eta}^{\eta}\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha y^2/2}e^{i\gamma y}\,dy\,.$$

Since

$$-\frac{\gamma^2}{2\alpha} - \frac{\alpha(y-i\gamma/\alpha)^2}{2} = -\frac{\alpha y^2}{2} + i\gamma y,$$

we can rewrite the integral as

$$e^{-\gamma^2/2\alpha+i\gamma\omega}\sqrt{\frac{\alpha}{2\pi}}\int_{-\eta}^{\eta} e^{-\alpha(y-i\gamma/\alpha)^2/2}\,dy\,.$$

Take $(y-i\gamma/\alpha)^2 = t^2/\alpha$, change of variable gives us

$$= e^{-\gamma^2/2\alpha+i\gamma\omega}\sqrt{\frac{\alpha}{2\pi}}\int_{\sqrt{\alpha}(-\eta-i\gamma/\alpha)}^{\sqrt{\alpha}(\eta-i\gamma/\alpha)} e^{-t^2/2}\,\frac{dt}{\sqrt{\alpha}}\,.$$

The standard normal distribution is given with the probability density function

$$\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\,.$$

This integrates to 1 over $(-\infty, \infty)$. Now we write

$$\int_{\omega-\eta}^{\omega+\eta} K(u-\omega)e^{i\gamma u}\,du$$
$$= e^{-\gamma^2/2\alpha+i\gamma\omega}\sqrt{\frac{\alpha}{2\pi}}\int_{-\eta}^{\eta} e^{-\alpha(y-i\gamma/\alpha)^2/2}\,dy$$
$$= e^{-\gamma^2/2\alpha+i\gamma\omega}\sqrt{\frac{\alpha}{2\pi}}$$
$$\times\left(\int_{-\infty}^{\infty} e^{-\alpha(y-i\gamma/\alpha)^2/2}\,dy - \int_{\eta}^{\infty} e^{-\alpha(y-i\gamma/\alpha)^2/2}\,dy - \int_{-\infty}^{-\eta} e^{-\alpha(y-i\gamma/\alpha)^2/2}\right)\,dy\,.$$

By the change of variable above we have

$$= e^{-\gamma^2/2\alpha+i\gamma\omega} - e^{i\gamma\omega}\left(\int_{\eta}^{\infty} K(y)e^{i\gamma y}\,dy + \int_{-\infty}^{-\eta} K(y)e^{i\gamma y}\,dy\right)\,.$$

Now (E1) becomes

$$-\sum_{0\le|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + \sum_{0\le|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}\left(\int_{\eta}^{\infty} K(y)e^{i\gamma y}dy + \int_{-\infty}^{-\eta} K(y)e^{i\gamma y}dy\right)\,,$$

where

$$\left|\sum_{0<|\gamma|<A}\frac{e^{i\gamma\omega}}{\rho}\left(\int_{\eta}^{\infty} K(y)e^{i\gamma y}dy + \int_{-\infty}^{-\eta} K(y)e^{i\gamma y}dy\right)\right| \le 2\sum_{0<|\gamma|<A}\frac{1}{\rho}\left|\int_{\eta}^{\infty} K(y)e^{i\gamma y}dy\right|.$$

The constant 2 comes from taking the absolute value of the integral

$$\left|\int_{\eta}^{\infty} K(y)e^{i\gamma y}\,dy\right| \le \int_{\eta}^{\infty}|K(y)||e^{i\gamma y}|\,dy = \int_{\eta}^{\infty}|K(y)|\,dy = \int_{-\infty}^{-\eta}|K(y)||e^{i\gamma y}|\,dy\,.$$

Integrating by parts with $w = k(y)$ and $dv = e^{i\gamma y}$ gives

$$\int_{\eta}^{\infty} K(y)e^{i\gamma y}\,dy = \left[K(y)\frac{e^{i\gamma y}}{i\gamma}\right]_{\eta}^{\infty} - \int_{\eta}^{\infty} K'(y)\frac{e^{i\gamma y}}{i\gamma}\,dy\,.$$

Note that in the first term, $|e^{i\gamma y}|$ is bounded by 1, thus since $K(y)$ decreases to 0 at infinity, the first term becomes $K(\eta)\frac{e^{i\gamma\eta}}{i\gamma}$. Note that this is simply

$$\int_{\eta}^{\infty} K'(y)\frac{e^{i\gamma\eta}}{i\gamma}\,dy\,,$$

so we can go back to the integral above and get

$$\int_{\eta}^{\infty} K(y)e^{i\gamma y}\,dy = \int_{\eta}^{\infty} K'(y)\frac{(e^{i\gamma\eta}-e^{i\gamma y})}{i\gamma}\,dy\,.$$

Furthermore,

$$\left|\frac{(e^{i\gamma\eta}-e^{i\gamma y})}{i\gamma}\right| \le \frac{2}{\gamma}\,,$$

thus

$$\left|\int_\eta^\infty K(y)e^{i\gamma y}\,dy\right| \le \frac{2}{\gamma}\int_\eta^\infty |K'(y)|\,dy = \frac{2}{\gamma}K(\eta) = \frac{2}{\gamma}\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha\eta^2/2}\,.$$

Based on the sequence of estimates:

$$\begin{aligned}(\text{E1}) &= -\sum_{0\le|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + 2\vartheta\sum_{0\le|\gamma|<A}\frac{1}{\rho}\left|\int_\eta^\infty K(y)e^{i\gamma y}\,dy\right| \\ &= -\sum_{0\le|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + 4\vartheta\sum_{0\le|\gamma|<A}\frac{1}{\gamma^2}\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha\eta^2/2}\,. \qquad (2.12)\end{aligned}$$

Previously, the estimates used in the calculations merely depended on the knowledge of the Gaussian kernel $K(y)$ at infinity. Now we will split the sum involving the zeta zeros into two parts based on the following lemmas. We can make the terms involving $\sum_\gamma \frac{1}{\gamma^2}$ more precise. Using Lemma 0.1 and the inequality $(2\pi)^{-1/2} < 0.4$,

$$\begin{aligned}(\text{E1}) &= -\sum_{0<|\gamma|<A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + 8\vartheta\sum_{0<\gamma<A}\frac{1}{\gamma^2}\sqrt{\frac{\alpha}{2\pi}}e^{-\alpha\eta^2/2} \\ &\le -\sum_{0<|\gamma|<A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + (8\cdot 0.025\cdot 0.4)\sqrt{\alpha}e^{-\alpha\eta^2/2} \\ &= -\sum_{0<|\gamma|<A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha} + .08\sqrt{\alpha}e^{-\alpha\eta^2/2}\,.\end{aligned}$$

Now we split the exponential sum so that we can enter any number of zeros we desire:

$$\left|\sum_{T<|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha}\right| \le 2\sum_{T<\gamma<\infty}\frac{e^{-\gamma^2/2\alpha}}{\gamma}$$

as we write the roots in conjugate pairs, and

$$\begin{aligned}&\le 2\int_T^\infty \frac{e^{-t^2/2\alpha}}{t}\,dN(t) \\ &\le \int_T^\infty \frac{e^{-t^2/2\alpha}}{\pi t}\log\frac{t}{2\pi}\,dt + 2\int_T^\infty \frac{e^{-t^2/2\alpha}}{t}Q'(t)\,dt\,.\end{aligned}$$

Choosing $dv = Q'(t)$ and $u = \frac{e^{-t^2/2\alpha}}{t}$,

$$\le \int_T^\infty \frac{e^{-t^2/2\alpha}}{\pi t}\log\frac{t}{2\pi}\,dt + 2\left[\left(Q(t)\frac{e^{-t^2/2\alpha}}{t}\Big|_T^\infty\right) + \int_T^\infty Q(t)\left(\frac{e^{-t^2/2\alpha}}{t^2} + \frac{e^{-t^2/2\alpha}}{\alpha}\right)dt\right].$$

We use the inequality $Q(t) < 2\log t$,

$$\begin{aligned}&< \int_T^\infty \frac{e^{-t^2/2\alpha}}{\pi t}\log\frac{t}{2\pi}\,dt + \frac{4\log T e^{-T/2\alpha}}{T} + 4\frac{\log T}{T}\int_T^\infty t\cdot\frac{e^{-t^2/2\alpha}}{\alpha}\,dt + \int_T^\infty \frac{e^{-t^2/2\alpha}}{t^2}\,dt \\ &= \int_T^\infty \frac{e^{-t^2/2\alpha}}{\pi t}\log\frac{t}{2\pi}\,dt + \frac{8\log T e^{-T/2\alpha}}{T} + \int_T^\infty \frac{e^{-t^2/2\alpha}}{t^2}\,dt\,.\end{aligned}$$

Note that the integrals in the expression above involve multiplying a monotone decreasing function with $e^{-t/2\alpha}$. We have

$$\begin{aligned}
&\left|\sum_{T<|\gamma|\le A}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha}\right| \\
&\le \left(-\alpha e^{-t^2/2\alpha}\Big|_T^{\infty}\right)\frac{\log T/2\pi}{\pi T^2}+\frac{8\log T e^{-T^2/2\alpha}}{T}+\left(-\alpha e^{-t^2/2\alpha}\Big|_T^{\infty}\right)\frac{4}{T^3} \\
&= e^{-T^2/2\alpha}\left\{\frac{\alpha\log T/2\pi}{\pi T^2}+\frac{8\log T}{T}+\frac{4\alpha}{T^3}\right\}.
\end{aligned}$$

This completes our description of (E1). To summarize, we have

$$\text{(E1)}=\sum_{0<|\gamma|<T}\frac{e^{i\gamma\omega}}{\rho}e^{-\gamma^2/2\alpha}+R_{E1}\,,$$

where

$$|R_{E_1}|\le e^{-T^2/2\alpha}\left\{\frac{\alpha\log T/2\pi}{\pi T^2}+\frac{8\log T}{T}+\frac{4\alpha}{T^3}\right\}+0.08\sqrt{\alpha}e^{-\alpha\eta^2/2}\,.$$

To estimate (E2), note that

$$\begin{aligned}
|\text{(E2)}| &\le \sum_{0<|\gamma|\le A}\int_{\omega-\eta}^{\omega+\eta}\left|K(u-\omega)\vartheta''(\rho u)\frac{e^{i\gamma u}}{\gamma^2 u}\right|du \\
&\le \sum_{0<|\gamma|\le A}\frac{1}{\gamma^2}\frac{1}{\omega-\eta}\,,
\end{aligned}$$

and when we apply the inequality 0.1,

$$\le \frac{0.05}{\omega-\eta}\,.$$

The most laborious part of the proof comes up in estimating (E3), the sum involving zeros of the zeta function that does not lie on the critical line (if they exist). Lehman does not assume any properties of these zeros other than the fact that if they exist, the magnitude of their imaginary components needs to exceed 14, which is approximatly the value of the imaginary component of the first zeta zero above the real axis. The rest of the proof is a steadfast application of Cauchy's integral formula to the expression in the integral. The order to which the derivatives should be taken are carefully chosen to ensure that the estimate at the end depends only on η, A, and α. This gives rise to the conditions (2.2) and (2.3), which are hitherto unused.

We would like to estimate the sum

$$\text{(E3)}=-\sum_{|\gamma|\ge A}\int_{\omega-\eta}^{\omega+\eta}K(u-\omega)ue^{-u/2}\,\mathrm{li}(e^{\rho u})\,du\,.$$

Consider the function

$$f_\rho(s)=\rho s e^{-\rho s}\,\mathrm{li}(e^{\rho s})e^{-\alpha(s-\omega)^2/2}$$

in the sector $-\frac{\pi}{4} \leq \arg(s) \leq \frac{\pi}{4}$. Note also that $\frac{5}{12}\pi < |\arg(\rho)| < \frac{\pi}{2}$ for every zeta zero simply because all of them lie in an area bound by a line that goes through $(0,0)$ and $(\frac{1}{2}, 14)$, which lies at an angle of $\frac{5}{12}\pi$ to the x-axis. Combining these inequalities, we have

$$\frac{\pi}{6} < |\arg(\rho s)| < \frac{3}{4}\pi \, .$$

We recall that

$$\operatorname{li}(e^{\rho u}) = \int_0^\infty e^{\rho u} \frac{e^{-t}}{\rho u - t} \, dt \, .$$

Applying this to $f_\rho(s)$, we get

$$\begin{aligned} |f_\rho(s)| &= \left| \rho s e^{-\rho s} e^{-\alpha(s-\omega)^2/2} \left[\int_0^\infty e^{\rho s} \frac{e^{-t}}{\rho s - t} \, dt \right] \right| \\ &\leq \frac{|\rho s| |e^{-\alpha(s-\omega)^2/2}|}{|\operatorname{Im}(\rho s)|} \left[-e^{-t} \right]_0^\infty \\ &\leq 2|e^{-\alpha(s-\omega)^2/2}| \, . \end{aligned} \tag{2.13}$$

We expand (E3) into

$$\begin{aligned} (\mathrm{E3}) &= -\sum_{|\gamma| \geq A} \int_{\omega-\eta}^{\omega+\eta} \sqrt{\frac{\alpha}{2\pi}} e^{-\alpha(u-\omega)^2/2} u e^{-u/2} \operatorname{li}(e^{\rho u}) \, du \\ &= -\sqrt{\frac{\alpha}{2\pi}} \sum_{|\gamma| > A} \frac{1}{\rho} \int_{\omega-\eta}^{\omega+\eta} e^{(\rho - \frac{1}{2})u} f_\rho(u) \, du \, . \end{aligned}$$

If we evaluate the integral using integration by parts and choose $v = f_\rho(u), dw = e^{(\rho-\frac{1}{2})u}$, after N times we get

$$\begin{aligned} &(\mathrm{E3}) = \\ &-\sqrt{\frac{\alpha}{2\pi}} \sum_{\gamma \geq A} \frac{1}{\rho} \Bigg\{ \sum_{n=0}^{N-1} \frac{(-1)^n e^{(\rho-1/2)\omega}}{(\rho-1/2)^{n+1}} \left[e^{(\rho-1/2)\eta} f_\rho^{(n)}(\omega+\eta) - e^{-(\rho-1/2)\eta} f_\rho^{(n)}(\omega-\eta) \right] \\ &+ \frac{(-1)^N}{(\rho-1/2)^N} \int_{\omega-\eta}^{\omega+\eta} e^{(\rho-1/2)u} f_\rho^{(N)}(u) \, du \Bigg\} \, . \end{aligned} \tag{2.14}$$

In order to apply the Cauchy Integral formula we need only to find a path in the domain where $f_\rho(s)$ is analytic to evaluate the integral. Let C be a circle of radius $r < \omega/4$ centered at u. If we invoke condition (2.3), specifically $\eta < \omega/2$, we have that for any s on C,

$$\operatorname{Re}(s) \geq \omega - \eta - \frac{\omega}{4} > \omega - \frac{\omega}{2} - \frac{\omega}{4} = \frac{\omega}{4}$$

and $|\operatorname{Im}(s)| \leq \frac{\omega}{4}$. Thus C lies below the line $x = y$ and in the sector $|\arg(s)| \leq \frac{\pi}{4}$. By the previous discussion, $f_\rho(s)$ is analytic in a region enclosing C. Cauchy's integral formula states that

$$f_\rho^{(n)}(u) = \frac{n!}{2\pi i} \oint_C \frac{f_\rho(z)}{(z-u)^{n+1}} \, dz \, .$$

We can then bound the absolute value of $f_\rho^{(n)}(s)$ on C:

$$|f_\rho^{(n)}(u)| \le \frac{n!}{2\pi i}\oint_C \frac{\max_{s\in C} f_\rho(s)}{|(s-u)|^{n+1}}\,ds$$

writing $|s-u| = r$ and using (2.13), we have

$$\le \frac{2n!}{r^n}\max_{|s-u|=r} |e^{-\alpha(s-\omega)^2/2}|\,.$$

If $s = \sigma + it$, then on the circle $(\sigma - u)^2 + t^2 = r^2$,

$$\begin{aligned}|e^{-\alpha(s-\omega)^2/2}| &= e^{\alpha(t^2-(\alpha-\omega)^2/2}\\ &= e^{\alpha(r^2-(\sigma-u)^2-(\sigma-\omega)^2)/2} \le e^{\alpha r^2/2}\end{aligned}$$

If $N \le \alpha\omega^2/16$, we can discard r by choosing it to be $\sqrt{N/\alpha}$, which satisfies the constraint $r \le \omega/4$. We plug the inequality back into our estimate for $f_\rho^{(n)}(u)$.

$$\begin{aligned}|f_\rho^{(N)}(u)| &\le \frac{N!}{(N/\alpha)^{N/2}} e^{\alpha N/2\alpha}\cdot 2\\ &= 2N!N^{-N/2}\alpha^{N/2}e^{N/2}\,.\end{aligned}$$

To estimate $f_\rho^{(N)}$, we choose $r = \eta/2$ Since by condition (2.3), $\eta < \omega/2$, $r < \omega/4$, this allows us to apply Cauchy's integral formula again and get

$$\begin{aligned}|f_\rho^{(n)}(\omega\pm\eta)| &\le 2n!(\eta/2)^{-n}\max_{|s-(\omega\pm\eta)|=r}|e^{-\alpha(s-\omega)^2/2}|\\ &= 2n!(\eta/2)^{-n}e^{\alpha(r^2-(\sigma-(\omega\pm\eta))^2-(\sigma-\omega)^2)/2}\\ &\le 2n!(\eta/2)^{-n}e^{-\alpha\eta^2/2}\end{aligned}$$

after expanding the terms in the exponent of e. We will proceed with estimating the first sum of the expression (2.14) of (E3) using some simple inequalities:

$$|e^{(\rho-1/2)\omega+(\rho-1/2)\eta}| = e^{(\beta-1/2)(\omega+\eta)} \le e^{1/2(\omega+\eta)}\,.$$

since $0 < \beta < 1$. Similarly,

$$|e^{-(\rho-\frac{1}{2})(\omega-\eta)}| \le e^{\frac{1}{2}(\omega+\eta)},$$

and we have

$$\begin{aligned}&|e^{(\rho-\frac{1}{2})\omega}\{e^{(\rho-\frac{1}{2})\eta}f_\rho^{(n)}(\omega+\eta) - e^{-(\rho-\frac{1}{2})\eta}f_\rho^{(n)}(\omega-\eta)\}\\ &\le |e^{(\rho-\frac{1}{2})(\omega+\eta)}f_\rho^{(n)}(\omega+\eta)| + |e^{-(\rho+\frac{1}{2})(\omega-\eta)}f_\rho^{(n)}(\omega-\eta)\}|\\ &\le \left(e^{\frac{1}{2}(\omega+\eta)}\cdot 2n!(\eta/2)^{-n}e^{-\alpha\eta^2/8}\right)\cdot 2\\ &= 4n!e^{\frac{1}{2}(\omega+\eta)}\left(\frac{\eta}{2}\right)^{-n}e^{-\alpha\eta^2/8}\end{aligned}$$

by using our estimates on $|f_\rho^{(n)}(\omega \pm \eta)|$ above.
The first sum of (E3) becomes

$$\leq \left| -\sqrt{\frac{\alpha}{2\pi}} \sum_{|\gamma|\geq A} \frac{1}{\rho} \sum_{n=0}^{N-1} e^{\frac{1}{2}(\omega+\eta)} \frac{4n!}{(\eta/2)^\eta \gamma^{n+1}} e^{-\alpha\eta^2/8} \right|$$

$$\leq \sqrt{\frac{\alpha}{2\pi}} e^{\frac{1}{2}(\omega+\eta)} \sum_{|\gamma|>A} \frac{4e^{-\alpha\eta^2/8}}{\gamma^2} \sum_{n=0}^{N-1} \frac{n!}{(\eta/2)^n \gamma^n} \, .$$

For the second integral in (2.14), we have

$$\left| \frac{(-1)^N}{(\rho-1/2)^N} \right| \leq \frac{1}{\gamma^n} \, ,$$

and

$$|e^{(\rho-1/2)u}| \leq e^{(\omega+\eta)/2}$$

in the interval. With $|f_\rho^{(N)}(u)| \leq 2N! N^{-N/2} \alpha^{N/2} e^{N/2}$,

$$\left| \sqrt{\frac{\alpha}{2\pi}} \sum_{|\gamma|\geq A} \frac{1}{\rho} \frac{(-1)^N}{(\rho-1/2)^N} \int_{\omega-\eta}^{\omega+\eta} e^{(\rho-1/2)u} f_\rho^{(N)}(u) du \right|$$

$$\leq \sqrt{\frac{\alpha}{2\pi}} e^{\frac{1}{2}(\omega+\eta)} \frac{(-1)^N}{\rho(\rho-1/2)^N} 2N! \left(\frac{\alpha e}{N}\right)^{N/2} (\omega+\eta-\omega+\eta)$$

$$\leq \sqrt{\frac{\alpha}{2\pi}} e^{\frac{1}{2}(\omega+\eta)} \frac{4\eta N!}{\gamma^{N+1}} \left(\frac{\alpha e}{N}\right)^{N/2} \, .$$

Now we have the final estimate for (E3):

$$|(\mathrm{E3})| \leq 2\sqrt{\frac{\alpha}{2\pi}} e^{(\omega+\eta)/2} \left[\sum_{\gamma>A} \frac{4e^{-\alpha\eta^2/8}}{\gamma^2} \sum_{n=0}^{N-1} \frac{n!}{(\eta/2)^n \gamma^n} + \frac{4\eta N!}{\gamma^{N+1}} \left(\frac{\alpha e}{N}\right)^{N/2} \right] .$$

Applying estimates such as Lemma 0.4, 0.5, and equation (0.1) for the sum over γ completes the proof. □

The implementation of this theorem consists of first choosing the parameters A, α, η, ω such that the conditions (2.2) and (2.3) are satisfied, and then deciding how many actual zeros should be entered into the calculation. Then the sum

$$-\sum_{0<|\gamma|\leq T} \frac{e^{i\gamma\omega}}{\rho} e^{-\gamma^2/2\alpha}$$

is computed over a wide range of ω (in Lehman's case, from 14 to 12000) to find heuristic values of x where the sum could be expected to exceed 1. After that, calculations with the estimates of the error term R is carried out to prove that $\pi(x) - \mathrm{li}(x)$ around this value. Lehman found 3 points where the value of the sum exceeds 0.96 with 1000 Riemann zeta zeros.

$$727.952, \quad 853.853, \quad 2682.977 \, .$$

Then he used the following parameters to prove that $\pi(x) > \mathrm{li}(x)$ around $e^{2682.977}$.

$$A = 17,000, \quad \alpha = 10^7, \quad \eta = 0.034,$$
$$T = \gamma_{12,000}, \quad \omega = 2682.9768$$

where the sum ≥ 1.00133. The error term R is smaller than 0.001268, thus

$$-1 - \sum_{0<|\gamma|<12,000} \frac{e^{i2682.9768\gamma}}{\rho} e^{-\gamma^2/(2\times 10^7)} + R$$
$$\geq -1 + 1.00133 - 0.001268 = 6.2 \times 10^{-5} > 0\,.$$

The number 6.2×10^{-5} is small, but this means that

$$\int_{2682.9428}^{2683.0108} K(u - 2682.9768) u e^{-u/2} \{\pi(e^u) - \mathrm{li}(e^u)\} du > 6.2 \times 10^{-5}\,.$$

Also

$$\int_{2682.9428}^{2683.0108} K(u - 2682.9768) u e^{-u/2} \{u^{-1} e^{u/2}\} du < 1$$
$$6.2 \times 10^{-5} \cdot \int_{2682.9428}^{2683.0108} K(u - 2682.9768) < 6.2 \times 10^{-5}\,,$$

since the probability function integrates to 1 over the entire real line. $K(u-\omega)$ and $ue^{-u/2}$ are both positive functions and so this implies that $\pi(e^u) - \mathrm{li}(e^u) > u^{-1}e^{u/2}$ for some $u \in (\omega - \eta, \omega + \eta)$. If we substitute $u = \omega - \eta = 2682.9428$, then we have that $\pi(e^u) - \mathrm{li}(e^u) > 10^{500}$.

The number obtained by Lehman is $\approx 10^{1165}$, this is an enormous improvement over Skewes' $10^{10^{10^{10^3}}}$. Why does Lehman's formula yield a much smaller number? His approach followed Littlewood's theorem, which does not assume anything about the distribution of the zeros beyond their number at a certain height. Dirichlet's lemma gives that for a large list of numbers any two of them will lie close to each other, but this does not account for how the zeros at a lower height may have more than two which lie closer to each other modulo 2π.

The number obtained by Lehman exposes the distance between Littlewood's model of the zeta zeros and the actual distribution of the zeros less than T.

3 Recent developments

After Lehman's breakthrough, the results on finding a lower upper bound for the Littlewood violation mainly consist of reapplying Lehman's theorem with little or no modification and increasing the number of zeros entered into the formula for higher accuracy. The results are summarized in the following table.

The most resent result by Chao and Plymen is notable not only for computing the lowest bound currently known, but also because it reduced error in Lehman's formula, where little had been done [2, p.1288]. Of computational methods I will refer the reader to [3] for a description of the formulas used to evaluate the zeros and the errors incurred from using approximations. Bay and Hudson in particular

	Number of zeros entered	Bound
Lehman	12,000	$\approx 1.65 \times 10^{1165}$
Te Riele	15,000	$\approx 6.69 \times 10^{370}$
Bay & Hudson	1,000,000	$\approx 1.398 \times 10^{316}$
Chao & Plymen	2,000,000	$\approx 1.397 \times 10^{316}$

Table 2: Upper bounds computed from 1966 to 2008

provides a graph of computed values of the sum and heuristic for where the first lower bound can be expected to occur [2, p.1291–1293].

There are many obstacles in refining Lehman's result—our imperfect knowledge of the influence of the 4 parameters, and the lack of more precise knowledge of the Riemann zeta zeros, not to mention our sole dependence to Riemann's formulas for the prime counting function. However, with recent results on better formulas for approximating $\pi(x)$, we can reduce the main source of error from Lehman's formula. The result depends on a more careful examination of the several $\vartheta's$ on page 27.

We absorb $\vartheta_3(x), \vartheta_2(x^{1/3})$ into one $\vartheta(x)$ and following similar arguments

$$\pi(x) - \mathrm{li}(x) = -\sum_{\rho} \mathrm{li}(x^{\rho}) - \frac{x^{1/2}}{\log x} + \vartheta_1(x^{1/2})\frac{3x^{1/2}}{(\log x)^2} + 4\vartheta(x^{1/2})x^{1/3}\,.$$

When we multiply this by the Gaussian kernel and $ue^{-u/2}$ and integrate over $(\omega - \eta, \omega + \eta)$, we see that the error term involving $\vartheta_1(x^{1/2}), \vartheta(x)$ become

$$\int_{\omega-\eta}^{\omega+\eta} K(u-w)ue^{-u/2}\left\{\vartheta_1(e^{u/2})\frac{3e^{u/2}}{(\log e^u)^2} + \vartheta(e^u)4(e^u)^{1/3}\right\}du$$
$$= \frac{3\vartheta_1(e^{u/2})}{\omega-\eta} + 4\vartheta(e^u)(\omega+\eta)e^{-(\omega-\eta)/6}\,.$$

Thus if we reduce $\vartheta_1(e^{u/2})$, we can reduce the major error term. We use the estimates on $\pi(x)$ by L. Panaitopol:

$$\pi(x) > \frac{x}{\log x - 1 + (\log x)^{-0.5}} \quad \text{for } x \geq 59\,, \text{ and}$$
$$\pi(x) < \frac{x}{\log x - 1 - (\log x)^{-0.5}} \quad \text{for } x > 6\,.$$

We get

$$\frac{x}{\log x - 1 + (\log x)^{-0.5}} < \pi(x) < \frac{x}{\log x - 1 - (\log x)^{-0.5}}\,.$$

By the inequality by Schoenfeld used to derive $\vartheta_1(e^{u/2})$, we get

$$\pi(x) = \frac{x}{\log x} + \frac{3/2x\vartheta_1(x)}{(\log x)^2} < \frac{x}{\log x - 1 - (\log x)^{-0.5}}$$

and

$$\pi(x) > \frac{x}{\log x}\,.$$

Denote $y = y(x) := (\log x)^{1/2}$. The first inequality of the two above will lead to an upper bound for ϑ_1:

$$0 < \vartheta_1(x) < \frac{2}{3} \cdot \frac{y^3 + y^2}{y^3 - y - 1} \quad \text{for all } x \geq 59\,.$$

The function

$$F(u) := \frac{y^3 + y^2}{y^3 - y - 1}$$

is a monotone decreasing function for $x \geq e$. Thus we can bound ϑ_1 by a very small number if x is chosen to be large enough. For example, if we choose 10^{59},

$$0 \leq \vartheta_1 < 0.71523279\,.$$

Finally, we have that the error term involving ϑ_1 is 2.14569. This is a substantial improvement over 3, even though it is unclear how this reduction has influenced the outcome.

Conclusion Where does one go from here? Littlewood's theorem has illustrated well enough the oscillatory nature of the difference $\pi(x) - \mathrm{li}(x)$. This proof was later made effective in computing a numerical value for the sign switch by Skewes. Lehman's theorem shows us that an upper bound for the sign switch can be efficiently computed. A possible direction for improvement would be to continue to reduce the error term R, or find a different representation of the prime counting function. There is of course always the option of including more zeros in Lehman's formula, but this will most likely not yield a great deal of improvement because the dominant term in the error comes from the range of the interval we wish to consider.

References

[1] R. J. Backlund, Über die Nullstellen der Riemannscher ζ-funktion, Acta Math. 41 (1918) 345–375.

[2] C. Bays and R.H. Hudson, A new bound for the smallest x with $\pi(x) > \mathrm{li}(x)$, Mathematics of Computation 69 (2000) 1285–1296.

[3] P. Borwein, S. Choi, B. Rooney and A. Weirathmueller, The Riemann Hypothesis: A Resource for the Afficionado and Virtuoso Alike , Springer, 2008.

[4] K. F. Chao and R.J. Plymen, A new bound for the smallest x with $\pi(x) > \mathrm{li}(x)$, http://arxiv.org/pdf/math/0509312

[5] J.B. Conrey, The Riemann Hypothesis, AMS Notices 50 (2003) 341–353, www.ams.org/notices/200303/fea-conrey-web.pdf.

[6] H. M. Edwards, Riemann's Zeta Function, Dover, 1974.

[7] A. Granville and G. Martin, Prime Number Races, American Math. Monthly, January (2006).

[8] T. Kotnik, The prime-counting function and its analytic approximations, Adv. Comput. Math. 29 (2008) 55–70.

[9] R. S. Lehman, On the difference $\pi(x) - \mathrm{li}(x)$, Acta Arith. 11 (1966) 397–410.

[10] J.E. Littlewood, Sur la distribution des nombres premiers, C.R. Acad. Sci. Paris 158 (1914) 1869–1872.

[11] H. L. Montgomery and R. C. Vaughan, Multiplicative Number Theory I. Classical Theory, Cambridge University Press, 2006.

[12] L. Panaitopol, Inequalities Concerning the Function $\pi(x)$: Applications, Acta Arith. 94 (2000) 373–381.

[13] B. Riemann, On the number of primes less than a given quantity. (Über die Anzahl der Primahlen unter einer gegebenen Grösse.), (1859), Translated into English by D. R. Wilkins (1998), `www.claymath.org/millennium/Riemann_Hypothesis/1859_manuscript/EZeta.pdf`

[14] S. Skewes, On the Difference $\pi(x) - \mathrm{li}(x)$ (II), Proc. London Math. Soc. 5 (1955) 48–70.

Email: rueyshan.lee@gmail.com