Analytic Number Theory.

§1. Tchebychev Estimates.

1. Definition: Let p_n denote the nth prime in the sequence of ascending order, $2, 3, 5, 7, \ldots$

Euclid proved that the sequence is infinite: we have $p_1 \cdots p_n + 1$ divisible by some prime other than $p_1, \ldots, p_n$.

Definition: Let $\pi(x)$ be the number of primes $\leq x$, where x is any positive real number, usually regarded as large.

We introduce the Tchebychev Functions: $\theta(x) = \sum_{p \leq x} \log p$, $\psi(x) = \sum_{p} \sum_{\substack{m: \\ p^m \leq x}} \log p$

Note that if $n \in \mathbb{N}$, then $\psi(n) = \log u(n)$, where $u(n) = \mathrm{lcm}(1, \ldots, n)$. Note also that $\psi(x) = \theta(x) + O(\sqrt{x}(\log x)^2)$, for we have $\psi(x) = \sum_{m \leq \log x / \log p} \theta(x^{1/m})$, and clearly $\theta(x^{1/m}) \leq \theta(x^{1/2}) \leq \sum_{n \leq x^{1/2}} \log n$, and this latter sum is at most $x^{1/2} \cdot \log x$.
We have also $\theta(x) \leq \pi(x) . \log x$, and $\theta(x) \geq \sum_{\sqrt{x} < p \leq x} \log p \geq \frac{1}{2} \log x . (\pi(x) - \pi(\sqrt{x}))$.
In fact, we have $\pi(x) \sim \theta(x)/\log x \sim \psi(x)/\log x$ as $x \to \infty$. (See Hardy & Wright p. 345)

Closely related to the Tchebychev function $\psi(x)$ is the von Mangoldt function $\Lambda(n)$. It is defined as $\log p$ if $n = p^j$ (p prime), and 0 otherwise. Then we have $\psi(x) = \sum_{n \leq x} \Lambda(n)$, and $\sum_{m | n} \Lambda(m) = \log n$.

2. Main Theorem.
Tchebychev proved in 1852 that $\pi(x)$ is bounded above and below by $x/\log x$, up to specific constants. More precisely, he proved:

Theorem 1: $\exists\, a, b > 0$ such that $a < \pi(x) \log x / x < b \quad \forall\, x \geq 2$.

(In Vinogradov's notation, we have $x/\log x \ll \pi(x) \ll x/\log x$, where by $y \ll z$ we mean $y < cz$ for some constant c, and similarly for $y \gg z$).
By the observations above, Theorem 1 implies $x \ll \theta(x) \ll x$, $x \ll \psi(x) \ll x$.

As a further corollary we obtain $n \log n \ll p_n \ll n \log n$. For we have $\pi(p_n) = n$, and so, by Theorem 1, $p_n / \log p_n \ll n \ll p_n / \log p_n$, hence $n \log p_n \ll p_n \ll n \log p_n$. Now $p_n \gg n$ and since $\log p_n < \sqrt{p_n}$ we get $\sqrt{p_n} \ll n$ and so $p_n \ll n^2$. Thus $\log n \ll \log p_n \ll \log n$, and the corollary follows.
Note also from $x \ll \theta(x) \ll x$ we see that there is always a prime between n and cn for some $c > 0$. This can be refined to give a prime p with $n \leq p < 2n$ for any integer $n > 1$. This is Bertrand's Postulate (for a proof, see Chandrasekharan, "Introduction to Analytic Number Theory").

Proof of Theorem 1: For any integer $n>1$, the binomial coefficient $\binom{2n}{n}$ is the largest term in the expansion $(1+1)^{2n}$. Hence $2^{2n}/(2n+1) \leq \binom{2n}{n} \leq 2^{2n}$. We now determine exponent to which any prime p divides $\binom{2n}{n}$. Now, p divides $n!$ with exponent $\sum_{j=1}^{\infty} [n/p^j]$ (since exactly $[n/p^j]$ of the numbers $1, \ldots, n$ are divisible by p^j), and so the exponent to which p divides $\binom{2n}{n}$ is $\sum_{j=1}^{\infty} \{[2n/p^j] - 2[n/p^j]\}$. $-(*)$
Now, each term in this sum is zero or 1, according as the fractional part of n/p^j is or is not less than $1/2$, and the number of non-zero terms is at most $\log(2n)/\log p$. Thus an upper bound for $\binom{2n}{n}$ is $\prod_{p \leq 2n} p^{\log(2n)/\log p} = (2n)^{\pi(2n)}$.
Further, the value of the sum $(*)$ is 1 when $n < p \leq 2n$, whence a lower bound for $\binom{2n}{n}$ is $\prod_{n<p\leq 2n} p \geqslant n^{\pi(2n)-\pi(n)}$.
On comparing the bounds, we obtain $(2n)^{\pi(2n)} \geqslant 2^{2n}/(2n+1)$ and $n^{\pi(2n)-\pi(n)} \leq 2^{2n}$.
The first inequality gives $\pi(2n) \geqslant \{2n\log 2 - \log(2n+1)\}/\log 2n$, whence $\pi(x) \gg x/\log x$.
The second inequality gives $\pi(2n) - \pi(n) \leq 2n \log 2/\log n$, whence $\pi(2x) - \pi(x) \ll x/\log x$.
Finally, on noting that $\pi(x) = \sum_{j=1}^{\infty} \{\pi(x/2^{j-1}) - \pi(x/2^j)\}$, and if $2^j < \sqrt{x}$ then $\pi(x/2^{j-1}) - \pi(x/2^j) \ll x/(2^j \log x)$, while if $2^j > \sqrt{x}$, then $\pi(x/2^{j-1}) - \pi(x/2^j) \ll x/2^{j-1} \ll \sqrt{x}$.
Since also the number of non-zero terms is $\ll \log x/\log 2$, we get
$\pi(x) \ll (x/\log x) \cdot \sum_{j=1}^{\infty} 1/2^{j-1} + \sqrt{x}\cdot(\log x/\log 2) \ll x/\log x$, as required.

Tchebychev himself obtained the estimates $a = 0.92129\cdots$, $b = 1.1055\cdots$ (for $x \geqslant 30$), for the constants in Theorem 1. Sylvester obtained (1892) the values $a = 0.95695\ldots$ and $b = 1.04423$. These were further improved by Rosser & Schoenfeld (1962). In particular, they showed that one can take $a=1$ for $x \geqslant 17$.

3. Mertens' Theorems.

Based on Tchebychev's results, Mertens obtained estimates for various sums and products involving primes. In particular, we have:

Theorem 2: $\sum_{p\leq x} \frac{\log p}{p} = \log x + O(1)$

Proof: We have (by comparison with the integral) $\sum_{n\leq x} \log n = x\log x + O(x)$.
The expression on the left is $\sum_{p\leq x}\sum_{j=1}^{\infty} \log p . [x/p^j]$ (Note that $[[x]/p^j] = [x/p^j]$).
The contribution from $j=1$ is $\sum_{p\leq x} \log p [x/p] = x \sum_{p\leq x} \frac{\log p}{p} + O(\pi(x)\log x)$.
Now by Theorem 1, the error term here is $O(x)$. The remaining terms contribute $\sum_{p\leq x}\sum_{j=2}^{\infty} \log p . [x/p^j] \leq x \sum_{n\leq x} \log n . \sum_{j=2}^{\infty} \frac{1}{n^j} \leq x \sum_{n\leq x} \log n/n(n-1)$. This latter sum converges as $n\to\infty$ and so the whole expression is $O(x)$. On combining our estimates, we get Theorem 2.

As a corollary, we obtain:

Theorem 3: $\sum_{p\leq x} \frac{1}{p} = \log\log x + c + O(1/\log x)$, where c is an absolute constant.

Note that this gives another proof of the existence of infinitely many primes.

Proof of Theorem 3: Let $s(x) = \sum_{p\le x} \frac{\log p}{p}$. Then by theorem 2 we have $s(x) = \log x + t(x)$, where $t(x) = O(1)$. Now, by partial summation, we have
$\sum_{p\le x} \frac{1}{p} = \sum_{2\le n\le x} \frac{s(n)-s(n-1)}{\log n} = \sum_{n=2}^{[x]-1} s(n)\left\{\frac{1}{\log n} - \frac{1}{\log(n+1)}\right\} + s(x)\left\{\frac{1}{\log[x]}\right\}$, using that $s(1)=0$ and $s(x) = s([x])$.

The expression on the right is: $\int_2^x \frac{s(u)\,du}{u(\log u)^2} + \frac{s(x)}{\log x}$, since the integral can be written as $\sum_{n=2}^{[x]-1} \int_n^{n+1} \frac{s(u)\,du}{u(\log u)^2} + s(x)\int_{[x]}^x \frac{du}{u(\log u)^2}$ and we have $\int_n^{n+1} \frac{du}{u(\log u)^2} = \frac{1}{\log n} - \frac{1}{\log(n+1)}$.

Now, putting $s(u) = \log u + t(u)$, we get:

$$\int_2^x \frac{du}{u\log u} + \int_2^x \frac{t(u)\,du}{u(\log u)^2} + 1 + \frac{t(x)}{\log x} = \log\log x + c + O(1/\log x),$$ where

$$c = 1 - \log\log 2 + \int_2^\infty \frac{t(u)\,du}{u(\log u)^2}.$$

Here we are using the fact that $\int_x^\infty \frac{t(u)\,du}{u(\log u)^2} = O\left(\int_x^\infty \frac{du}{u(\log u)^2}\right) = O(1/\log x)$.

Note that to get a heuristic argument furnishing the main term in a sum over primes as above we can substitute $n\log n$ for p_n. In this case we get $\sum_{p\le x}\frac{1}{p}$ being approximately $\sum_{n\le x} \frac{1}{n\log n}$, ie $\int_2^x \frac{du}{u\log u} \cong \log\log x$.
Similarly, $\sum_{p\le x} \frac{\log p}{p}$ is approximately $\sum_{n\le x}\frac{1}{n}$, ie $\log x$.

Theorem 4: We have $\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1} = C.\log x + O(1)$, for some $C>0$.

Proof: We have $\left(1-\frac{1}{p}\right)^{-1} = e^{-\log(1-1/p)} = e^{\left(\frac{1}{p}+\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots\right)}$. Hence $\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1} = \exp\left\{\sum_{p\le x}\frac{1}{p} + \sum_{p\le x}\left(\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots\right)\right\}$

Now, $\sum_p \left(\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots\right)$ is convergent and $\sum_{p>x}\left(\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots\right) \le \sum_{n>x}\frac{1}{n(n-1)} = O(1/x)$.

So $\sum_{p\le x}\left(\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots\right) = C' + O(1/x)$.

From Theorem 3 we have $\sum_{p\le x}\frac{1}{p} = \log\log x + c + O(1/\log x)$.

Hence $\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1} = \exp\{\log\log x + c + C' + O(1/\log x)\} = e^{c+C'}.\log x . e^{O(1/\log x)}$
$= C.\log x.(1+O(1/\log x)) = C\log x + O(1)$, with $C = e^{c+C'}$.

The argument gives $C = e^\gamma$ where γ is Euler's constant, and $c = \gamma + \sum_p \left(\log\left(1-\frac{1}{p}\right) - \frac{1}{p}\right)$.

Exercise: Prove that $\prod_{p\le x}\left(1+\frac{1}{p}\right) \sim \frac{6e^\gamma\log x}{\pi^2}$ as $x\to\infty$.

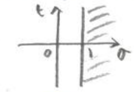
It is a simple deduction from Theorem 3 that if $\pi(x)\log x/x$ tends to a limit l as $x\to\infty$ then $l=1$. This follows from $\sum_{p\le x}\frac{1}{p} = \frac{\pi(x)}{x} + \int_2^x \frac{\pi(u)}{u^2}\,du$, which is established by partial summation.

This was the position until 1896, when Hadamard and de la Valleé Poussin independently proved the Prime Number Theorem, that is, $\pi(x) \sim x/\log x$ as $x\to\infty$. The proof utilised the Riemann Zeta function, $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$, introduced by Riemann in 1861. Euler had already studied the function for real values of s, but it was Riemann who first systematically investigated the function for complex s, and this initiated analytic number theory as we know it today. An "elementary" proof of the Prime Number Theorem was discovered by Erdős and Selberg independently in 1948. Though long sought, this has had less impact than originally envisaged; in particular, the error terms obtained by the "elementary" method are weaker than those obtainable from the Riemann Zeta function.

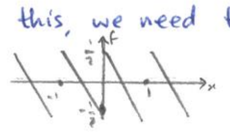
§ 2. The Riemann Zeta Function.

1. Analytic Continuation.

The Riemann Zeta Function is defined, for complex numbers $s = \sigma + it$ with $\sigma > 1$, by $\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}$. ($n^s = e^{s \log n}$ with $\log n$ real). This is a valid definition since $|n^s| = n^\sigma$ and so the series converges absolutely for $\sigma > 1$ (and indeed uniformly for $\sigma > 1 + \delta > 1$.



Riemann showed that $\zeta(s)$ can be analytically continued to the left of the line $\sigma = 1$. For this, we need the "saw-tooth" function, defined by $f(x) = [x] - x + \frac{1}{2}$, for all real x.



We have $|f(x)| \leq \frac{1}{2}$ for all x, and $F(x) = \int_1^x f(u)\,du$ is also bounded, for all x.

Further, $f(x)$ can be expressed as a simple Fourier series, $f(x) = \sum_{n=1}^{\infty} \frac{\sin(2n\pi x)}{n\pi}$

Now we show that $\zeta(s)$ can be analytically continued to the region $\sigma > 0$, by
$\zeta(s) = s \int_1^\infty \frac{f(x)}{x^{s+1}}\,dx + \frac{1}{s-1} + \frac{1}{2}$. with $x^{s+1} = e^{(s+1)\log x}$ where $\log x$ is real.
The integral converges for $\sigma > 0$, and indeed uniformly for $\sigma > \delta > 0$; since $|x^{s+1}| = e^{(1+\sigma)\log x}$.

To establish the equation for $\zeta(s)$ we observe that $\int_1^\infty \frac{f(x)}{x^{s+1}}\,dx = \sum_{n=1}^{\infty} \int_n^{n+1} \frac{f(x)}{x^{s+1}}\,dx = \sum_{n=1}^{\infty} \int_n^{n+1} \frac{n - x + \frac{1}{2}}{x^{s+1}}\,dx$.
This integral is: $\left[\frac{1}{(s-1)x^{s-1}} - \frac{n + \frac{1}{2}}{s x^s} \right]_n^{n+1} = \left(\frac{1}{s} - \frac{1}{s-1} \right)\left(\frac{1}{n^{s-1}} - \frac{1}{(n+1)^{s-1}} \right) + \frac{1}{2s}\left(\frac{1}{n^s} + \frac{1}{(n+1)^s} \right)$

On summing over n and noting that the series on the right converge for $\sigma > 1$, we obtain $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ $(\sigma > 1)$, by rearrangement.
The integral $\int_1^\infty \frac{f(x)}{x^{s+1}}\,dx$ converges not only for $\sigma > 0$ but in fact for $\sigma > -1$, and uniformly for $\sigma > -1 + \delta > -1$, for we have $\int_X^Y \frac{f(x)}{x^{s+1}}\,dx = \left[\frac{F(x)}{x^{s+1}} \right]_X^Y + (s+1)\int_X^Y \frac{F(x)}{x^{s+2}}\,dx$, and, as already noted, $F(x)$ is bounded for all x.
Now, if $\sigma < 0$ we have $s\int_0^1 \frac{f(x)}{x^{s+1}}\,dx = s\int_0^1 \frac{-x + \frac{1}{2}}{x^{s+1}}\,dx = \frac{1}{s-1} + \frac{1}{2}$.
Thus, $\zeta(s) = s\int_0^\infty \frac{f(x)}{x^{s+1}}\,dx$ $(-1 < \sigma < 0)$.

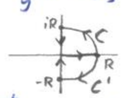
2. The Functional Equation.

We have the Fourier Series $f(x) = \sum_{n=1}^{\infty} \frac{\sin(2\pi n x)}{n\pi}$, for all x. On substituting into the expression above for $\zeta(s)$ $(-1 < \sigma < 0)$ and interchanging summation and integration (as is justified by, eg, dominated convergence), we get $\zeta(s) = \frac{s}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \int_0^\infty \frac{\sin(2\pi n x)}{x^{s+1}}\,dx$ $(-1 < \sigma < 0)$
$= \frac{s}{\pi} \sum_{n=1}^{\infty} \frac{(2n\pi)^s}{n} \int_0^\infty \frac{\sin y}{y^{s+1}}\,dy$
Now we have the integral, $\int_0^\infty \frac{\sin y}{y^{s+1}}\,dy = -\sin\left(\frac{s\pi}{2}\right)\Gamma(-s)$, whence $\zeta(s) = s.2^s.\pi^{s-1}.\sin\left(\frac{s\pi}{2}\right).(-\Gamma(-s)).\sum_{n=1}^{\infty} \frac{1}{n^{1-s}}$
On using $z\Gamma(z) = \Gamma(z+1)$ with $z = -s$, this gives: $\zeta(s) = 2^s.\pi^{s-1}.\sin\left(\frac{s\pi}{2}\right).\Gamma(1-s)\zeta(1-s)$. $[-1 < \sigma < 0)$.

Now, the right hand side is regular for $s = \sigma + it$ with $\sigma < 0$. Hence, the equation serves to give the analytic continuation for $\zeta(s)$ to the region $\sigma \leq -1$, and it is then valid throughout the complex plane. The equation can be simplified by defining $\Xi(s) = \pi^{-s/2}.\Gamma\left(\frac{s}{2}\right)\zeta(s)$ and using $\Gamma\left(\frac{1}{2}(1-s)\right)/\Gamma\left(\frac{1}{2}s\right) = \pi^{1/2}.2^s.\sin\left(\frac{s\pi}{2}\right).\Gamma(1-s)$.
We get $\Xi(s) = \Xi(1-s)$, which is the Functional equation in terms of Ξ.

Note that one can also take $\xi(s) = s(s-1)\Xi(s)$, and we again obtain $\xi(s) = \xi(1-s)$.
Now, by the definition of $\zeta(s)$, we see that it is a regular function throughout the complex plane, except for a simple pole at $s=1$, with residue 1. Hence Ξ has a simple pole at $s=1$, and so ξ is an entire function.
So Ξ and ξ have symmetry about the line $\sigma = 1/2$.

Note: to get the value of $\int_0^\infty \frac{\sin y}{y^{s+1}}\,dy =: I$, we use $I = \frac{1}{2i}\left\{\int_0^\infty \frac{e^{iy}}{y^{s+1}}\,dy - \int_0^\infty \frac{e^{-iy}}{y^{s+1}}\,dy\right\}$
and $\int_C \frac{e^{iz}}{z^{s+1}}\,dz = \int_0^R \frac{e^{iy}}{y^{s+1}}\,dy + i\int_0^R \frac{e^{-t}}{(it)^{s+1}}\,dt + \int_0^{\pi/2} \frac{e^{iRe^{i\theta}}}{(Re^{i\theta})^{s+1}}\cdot iRe^{i\theta}\,d\theta$,



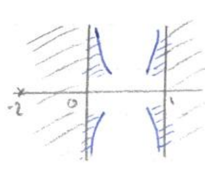
and similarly for $\int_{-C'} \frac{e^{-iz}}{z^{s+1}}\,dz$.
By Jordan's Lemma, the arc integral is $o(1)$ as $R\to\infty$, whence $I = -\frac{1}{2}\left\{\int_0^\infty \frac{e^{-t}}{(it)^{s+1}}\,dt + \int \frac{e^{-t}}{(-it)^{s+1}}\,dt\right\}$
$= -\frac{1}{2}\left\{\frac{1}{i^s} - \frac{1}{(-i)^s}\right\}\Gamma(-s) = -\sin\left(\frac{\pi s}{2}\right)\Gamma(-s)$.

3. The Euler Product.

This establishes the basic connection between the zeta-function and the primes. Namely, $\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$ $(\sigma > 1)$. To prove this, let N be a large positive integer. Then $\prod_{p\le N}\left(1-\frac{1}{p^s}\right)^{-1} = \prod_{p\le N}\left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) = \sum_m m^{-s}$, where the sum is over all positive integers m divisible by only primes $p \le N$. Further, we have, $\left|\sum_m m^{-s} - \sum_{n\le N} n^{-s}\right| \le \sum_{n>N} n^{-\sigma} \to 0$ as $N\to\infty$, whence result.
As an immediate consequence we obtain $\zeta(s) \ne 0$ for $\sigma > 1$. This leads to the basis question of the positions of the zeroes of $\zeta(s)$.

From the functional equation, $\zeta(s) = 2^s\cdot\pi^{s-1}\cdot\sin(\frac{1}{2}s\pi)\Gamma(1-s)\zeta(1-s)$, and the result $\zeta(s)\ne 0$ for $\sigma > 1$ we see that $\zeta(s) \ne 0$ for $\sigma < 0$ except at the zeroes of $\sin(\frac{1}{2}s\pi)$, that is, at $s = -2, -4, -6, \dots$
These are termed the trivial zeroes of $\zeta(s)$. We have also $\zeta(0) \ne 0$ since $\zeta(s)$ has a pole at $s=1$ with residue 1, and this "cancels" the zero of $\sin(\frac{1}{2}\pi s)$. There remains the question of the zeroes of $\zeta(s)$ for $0 \le \sigma \le 1$. This region is called the critical strip. The famous "Riemann Hypothesis" asserts that the only zeros of $\zeta(s)$ other than the trivial zeroes all lie on the lie $\sigma = 1/2$. This is unproved but it is known that at least the first several million zeros of $\zeta(s)$ in the critical strip do indeed lie on $\sigma = 1/2$. Also, by results of Selberg, Hardy and others, it is known that a "positive proportion" of the zeroes in the critical strip do lie on the line $\sigma = 1/2$. However, the best that has been established as regards zero-free regions is $\zeta(s) \ne 0$ for $\sigma > 1 - c/\{(\log t)^{2/3}(\log\log t)^{1/3}\}$ $(t \ge t_0)$. (Vinogradov & Korotov, 1958).

So we have:



This improved the basic result going back to Hadamard and de la Valleé Poussin, that $\zeta(s) \ne 0$ for $\sigma > 1 - c/\log|t|$ $(t \ge t_0)$.

In particular, we have:

Theorem 5: $\zeta(s) \ne 0$ on the line $\sigma = 1$.

This is enough to prove the prime number theorem.

Proof of Theorem 5: Let $s = \sigma + it$, $1 < \sigma \leq 2$. By the Euler product, we have

$\log \zeta(s) = -\sum_p \log\left(1 - \frac{1}{p^s}\right) = \sum_p \sum_{m=1}^{\infty} \frac{1}{m} p^{-ms}$, and since $\operatorname{Re} \log \zeta(s) = \log|\zeta(s)|$, we obtain

$\log|\zeta(s)| = \sum_p \sum_{m=1}^{\infty} \frac{1}{m} p^{-m\sigma}. \cos(mt\log p)$

Now, the identity $3 + 4\cos\theta + \cos 2\theta = 2(1+\cos\theta)^2 \geq 0$ gives

$3\log|\zeta(\sigma)| + 4\log|\zeta(\sigma+it)| + \log|\zeta(\sigma+2it)| \geq 0$, with $\theta = mt\log p$, whence

$|\zeta(\sigma)|^3 . |\zeta(\sigma+it)|^4 . |\zeta(\sigma+2it)| \geq 1.$ $- (*)$

Now, $\zeta(s)$ has a simple pole at $s=1$ and so $|\zeta(\sigma)| \ll (\sigma-1)^{-1}$, where the implied constant depends only on t. Further, $\zeta(s)$ is regular at $\sigma + 2it$, and thus $|\zeta(\sigma+2it)| \ll 1$.

Now, if $\zeta(s)$ possessed a zero at $s = 1+it$, then $|\zeta(s)| \ll (\sigma - 1)$. This gives, from $(*)$, $\sigma - 1 \gg 1$, and thus we obtain a contradiction if σ is taken sufficiently close (depending on t) to the line $\sigma = 1$.

4. Approximate Functional Equation.

We shall prove that, in the critical strip (actually for $\sigma > 0$), we have

$$\zeta(s) = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} + O(|t|.N^{-\sigma}) \quad - (*)$$

This will suffice to prove the Prime Number Theorem, but there are stronger results. In particular, (see Ivic, p. 21), we have $\zeta(s) = \sum_{n\leq x} \frac{1}{n^s} + \frac{x^{1-s}}{s-1} + O(x^{-\sigma})$ with $0 < \sigma_0 \leq \sigma \leq 2$, $x \geq |t|/\pi$, and with the constant in the O-term depending on σ_0.

The most famous "approximate functional equation" is due to Hardy and Littlewood (1921), and states $\zeta(s) = \sum_{n\leq x} \frac{1}{n^s} + \chi(s) \sum_{n \leq y} \frac{1}{n^{1-s}} + O(x^{-\sigma}) + O(t^{\frac{1}{2}-\sigma} y^{\sigma - 1})$ for $0 \leq \sigma \leq 1$, $x, y, t > c > 0$, and $2\pi xy = t$, where $\chi(s) = 2^s . \pi^{s-1} . \sin(\frac{1}{2}s\pi).\Gamma(1-s)$.

We now establish $(*)$. By dividing the interval $[1, N]$ into unit intervals and integrating by parts, we obtain: $s\int_1^N \frac{f(x)}{x^{s+1}}\,dx + \frac{1}{s-1} + \frac{1}{2} = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - \frac{1}{2}N^{-s}$, and so

$$\zeta(s) = \sum_{n=1}^{N} \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - \frac{1}{2}N^{-s} + s\int_N^{\infty} \frac{f(x)}{x^{s+1}}\,dx$$

Now, $\frac{1}{2}N^{-s} = O(N^{-\sigma})$, and $\int_N^{\infty} \frac{f(x)}{x^{s+1}}\,dx = O\left(\int_N^{\infty} \frac{dx}{x^{\sigma+1}}\right)$, and $s = O(|t|)$, so the desired equation follows.

5. The Logarithmic Derivative.

This is defined as $\frac{d}{ds}(\log \zeta(s)) = \frac{\zeta'(s)}{\zeta(s)}$. Now, for $\sigma > 1$, we have $\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$, where $\Lambda(n)$ is the von Mangoldt function. We recall that $\psi(x) = \sum_{n\leq x} \Lambda(n)$. These equations enable us to prove:

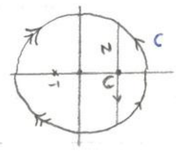
Theorem 6: $$\int_0^x \psi(u)\,du = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\zeta'(s)}{\zeta(s)} . \frac{x^{s+1}}{s(s+1)}\,ds, \quad x > 0,\ c > 1.$$

There is a more direct connection between $\psi(x)$ and the Riemann Zeta Function, namely $\psi(n) = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{(n+1/2)^s}{s} . \frac{\zeta'(s)}{\zeta(s)}\,ds$ $(n \in \mathbb{Z}, c > 1)$, and this could be used in place of Theorem 6 to establish the Prime-Number Theorem, but the convergence on the right is weaker than that in Theorem 6, and this gives additional problems.

Proof of Theorem 6: We shall show if $\eta(s)$ is the Dirichlet Series $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$, and if $\eta(s)$ converges absolutely for $\sigma > b > 0$ and if $S(x) = \sum_{n \le x} a_n$, then $\int_0^x S(u)\,du = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\eta(s)\, x^{s+1}}{s(s+1)}\, ds$ $(x > 0,\ c > b)$.
To get theorem 6, we take $\eta(s) = \frac{-\zeta'(s)}{\zeta(s)}$, so that $a_n = \Lambda(n)$, $b = 1$.
We need the following:

Lemma: $\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)}\, ds = \max(0, x-1)$, $c > 0$.

Proof: Consider the integral $I = \frac{1}{2\pi i} \int_C \frac{x^{s+1}}{s(s+1)}\, ds$, where C is the contour given by:
C is the right-hand part of the circle centre O and radius $\sqrt{c^2+N^2}$ if $x \le 1$, and left-hand part of it if $x > 1$, together with line from $c-iN$ to $c+iN$.



Cauchy's Theorem $\Rightarrow$ $I = \max(0, x-1)$.
Now, the integral over the circular part is at most $\frac{2\pi x^{c+1}}{N-1}$ in both cases, and this tends to 0 as $N \to \infty$. This proves the lemma.

Back to Theorem 6: The lemma gives $\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\eta(s) \cdot x^{s+1}}{s(s+1)}\, ds = \sum_{n=1}^{\infty} \frac{a_n}{2\pi i} \cdot \int_{c-i\infty}^{c+i\infty} \frac{n(x/n)^{s+1}}{s(s+1)}\, ds = \sum_{n=1}^{\infty} n a_n \cdot \max(0, \frac{x}{n} - 1)$
$= \sum_{n=1}^{\infty} a_n \cdot \max(0, x-n) = \sum_{n=1}^{\infty} (S(n) - S(n-1)) \cdot \max(0, x-n) = \sum_{n \le [x]} S(n) + S([x])(x-[x])$, and this last expression is $\int_0^x S(u)\,du$, which gives Theorem 6.

Our aim is to prove the Prime Number Theorem, and for this we need an estimate for $\zeta'(s)/\zeta(s)$ for $\sigma \ge 1$. We know this exists for $\sigma = 1$ by Theorem 5, and it is easy to verify that the argument there gives a bound for $\zeta(s)$ from below of the form $(\log|t|)^{-\lambda}$ for some $\lambda > 0$. In fact, it is easy to get $|\zeta(s)| \gg (\log|t|)^{-8}$, with the constant absolute. Further from $(*)$ in §2.4, we get $|\zeta(s)| \ll \log|t|$, and from this we obtain $|\zeta'(s)| \ll (\log|t|)^2$. Hence we get $|\zeta'(s)/\zeta(s)| \ll (\log|t|)^{10}$ for $\sigma \ge 1$.
The estimate for $|\zeta(s)|$ from $(*)$ follows since for $\sigma > 1 - \frac{1}{\log|t|}$, have
$\left|\sum_{n=1}^{N} \frac{1}{n^s}\right| \ll \sum_{n=1}^{N} \frac{1}{n^\rho} \ll \frac{N^{1-\rho}}{|1-\rho|} \ll \log|t|$, where $\rho = 1 - \frac{1}{\log|t|}$.
Then, the estimate for $|\zeta'(s)| \ll (\log|t|)^2$ follows from $\zeta'(s) = \frac{1}{2\pi i} \int_C \frac{\zeta(s+z)}{z^2}\, dz$, where C is the circle centre O, radius $1/\log|t|$.

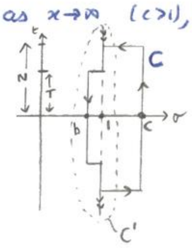
6. The Prime-Number Theorem

We shall prove that $\pi(x) \sim x/\log x$ as $x \to \infty$. We shall establish this theorem by the following lemma.

Lemma: If $\int_0^x \psi(u)\,du \sim \frac{1}{2}x^2$ as $x \to \infty$ then $\pi(x) \sim x/\log x$.

Proof: We have $\psi(x)$ increasing for $x > 0$ and so for any $h > 0$,
$$\frac{1}{h}\left\{\int_0^x \psi(u)\,du - \int_0^{x-h} \psi(u)\,du\right\} < \psi(x) < \frac{1}{h}\left\{\int_0^{x+h} \psi(u)\,du - \int_0^x \psi(u)\,du\right\}$$
Now if $\int_0^x \psi(u)\,du \sim \frac{1}{2}x^2$ we get: $x - \frac{1}{2}h + o(x^2/h) < \psi(x) < x + \frac{1}{2}h + o(x^2/h)$.
Taking $h = \delta x$ $(\delta > 0)$, we have $(1-\delta)x < \psi(x) < (1+\delta)x$ for x sufficiently large, ie $\psi(x) \sim x$ as $x \to \infty$.
Finally we have $\psi(x) = \pi(x)\log x + o(x)$, since:
$\psi(x) \le \sum_{p \le x} [\log x/\log p] \log p \le \pi(x)\log x$, and $\psi(x) \ge \sum_{p \le x} \log p \ge \sum_{\frac{x}{(\log x)^2} \le p \le x} \log p \ge (\log x - 2\log\log x)(\pi(x) - x/(\log x)^2)$.
This gives $\psi(x) \sim \pi(x)\log x$ as $x \to \infty$.
Hence from $\psi(x) \sim x$ we have $\pi(x) \sim x/\log x$.

Now, by the Lemma of §5 it suffices to prove that $\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Phi(s)\,ds \sim -\frac{1}{2}x^2$ as $x \to \infty$ $(c>1)$, where $\Phi(s) = \frac{\zeta'(s)}{\zeta(s)} \cdot \frac{x^{s+1}}{s(s+1)}$. We use contour integration. Let C be the contour:



Here, N is a large positive number $(N \to \infty)$. T is also chosen large.
Then, $\zeta(s)$ has only a finite number of zeroes in any bounded region (true of any analytic function), so $b<1$ can be chosen so that $\zeta'(s)/\zeta(s)$ is regular within and on C except for a simple pole at $s=1$, with residue -1.
By Cauchy's Theorem, we have thus $\frac{1}{2\pi i}\int_C \Phi(s)\,ds = -\frac{1}{2}x^2$
On taking components of C and allowing N to tend to ∞, we see that it suffices to show that $\int_{C'} \Phi(s)\,ds = o(x^2)$ as $x \to \infty$, where C' is the contour shown above.

Now, by §5, we have $|\Phi(1+it)| \ll (\log T)^{10}.(x/T)^2$ for $|t| \geq T$, whence $\int_{C''} \Phi(s)\,ds = o(x^2)$ for T large enough, where C'' is the part of C' on the line $\sigma = 1$. Thus if C''' is the remaining part of C' (over indentation), we get $x^{-2}\int_{C'''} \Phi(s)\,ds \ll \left\{\int_{-T}^{T} x^{b-1}\,dt + 2\int_b^1 x^{\sigma-1}\,d\sigma\right\}$, where the implied constant depends on T and b, but not on x.
The expression in braces is: $2Tx^{b-1} + 2(1-x^{b-1})/\log x \to 0$ as $x \to \infty$, hence $\int_{C'''} \Phi(s)\,ds = o(x^2)$.

As already mentioned, Erdös and Selberg gave an "elementary proof" of the Prime Number Theorem in 1948. This depends on Selberg's formula, $\psi(x)\log x + \sum_{n \leq x} \Lambda(n)\psi(x/n) = 2x\log x + O(x)$
This is equivalent to: $\log x . \sum_{p \leq x} \log p + \sum_{p p' \leq x} \log p \log p' = 2x\log x + O(x)$
To prove the formula, one uses a double applications of the Möbius result:
$F(x)\log x + \sum_{n \leq x} \Lambda(n).F(x/n) = \sum_{d \leq x} \mu(d).G(x/d)$, where $G(x) = \sum_{m \leq x} F(x/m \log x)$,
first with $F(x) = \psi(x)$, and secondly with $F(x) = x - \gamma - 1$, where γ = Euler's constant.

7. Refinements to the Prime Number Theorem.

Gauss conjectured in 1849 that a good approximation to $\pi(x)$ was given by $Li(x) = \int_2^x \frac{dt}{\log t}$, so that $\frac{\pi(x)}{li(x)} \to 1$ as $x \to \infty$.
On integrating by parts, we have, $li(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \dots + \frac{m!\,x}{(\log x)^{m+1}} \cdot (1 + o(x))$,
and so the conjecture (as thus interpreted) gives more than the prime number theorem.
(Agreement very good: 1985, Lagarius, Miller and Odlyzko showed $\pi(x) = 1,075,292,778,753,150$ when $x = 4 \times 10^{16}$, and $li(x) = 1,075,292,728, \dots, \dots$)
See also Ribenbaum.

The conjecture was proved by de la Vallée Poussion in 1899, and in fact he showed that $\pi(x) = li(x) + O(xe^{-c\sqrt{\log x}})$, where c and the implied constant for O-term are absolute.
Error term was improved to $O(x\exp(-c(\log x)^{3/5}/(\log\log x)^{1/5}))$, by Vinogradov and Korokov in 1958 as a consequence of their zero-free region for $\zeta(s)$, and this is the best result to date.
However, if the Riemann Hypothesis (R.H.) is true, then we have $\pi(x) = li(x) + O(x^{1/2}.\log x)$.

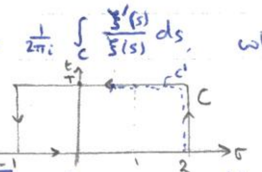
On numerical evidence, it was thought that $\pi(x) < li(x)$ for all x. This was disproved by Littlewood in 1914. He showed that $\pi(x) = li(x) + \Omega_{\pm}(x^{1/2}.\log\log\log x/\log x)$, where

$f(x) = g(x) + \Omega_{\pm}(w(x))$ means $f(x) - g(x) \gg w(x)$ and $\ll -w(x)$, for an infinite sequence of values of x tending to ∞, where the implied constants are positive and absolute.
Littlewood's original proof involved two cases, according as R.H. is true or not, and in view of the indirect character, it did not give an x_0 such that $\pi(x) > li(x)$ for some $x < x_0$. However, in 1955 Skewes succeeded in calculating an x_0, namely $10^{10^{10^{10^3}}}$, later termed "Skewes number." This has now been improved to a single exponential, but still we cannot give a specific x for which $\pi(x) > li(x)$.
Proofs of refinements in the prime-number theorem involve an explicit formula for $\zeta'(s)/\zeta(s)$, namely
$\zeta'(s)/\zeta(s) = C - \frac{1}{s-1} - \frac{1}{2}\frac{\Gamma'(\frac{1}{2}s+1)}{\Gamma(\frac{1}{2}s+1)} + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right)$, where ρ runs through all the zeroes in the critical strip, and C is a constant. The Γ-term can be expressed as $\frac{1}{2}\gamma + \sum_{n=1}^{\infty}\left(\frac{1}{s+2n} - \frac{1}{2n}\right)$, and so represents the contribution of the trivial zeroes of $\zeta(s)$.

8. The Riemann-von Mangoldt Formula.

In Riemann's original memoir, he conjectured that $N(T)$, the number of zeroes of $\zeta(s)$ in the critical strip such that $0 < t \leq T$ satisfies: $N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + O(\log T)$.
This was proved by von Mangoldt in 1905.
Hardy later showed that $N_0(T) \gg T$, where $N_0(T)$ is the number of zeroes of $\zeta(s)$ on the line $\sigma = \frac{1}{2}$, and Selberg improved this to $N_0(T) \gg T\log T$, so that a positive proportion of the zeroes of $\zeta(s)$ are on the critical line. The implied constant has been given as $1/3$ by Levinson; for the record, see Ribenbaum. "At least a third of the zeroes of $\zeta(s)$ lie on the critical line."

For the proof of the Riemann-von-Mangoldt formula, we use $N(T) = \frac{1}{2\pi i}\int_C \frac{\xi'(s)}{\xi(s)}\,ds$, where $\xi(s) = s(s-1)\pi^{-\frac{1}{2}s}\Gamma(\frac{1}{2}s)\zeta(s)$, where C is the contour shown:



By symmetry, $\xi(s) = \xi(1-s)$ (functional equation), and $\xi(1-s) = \overline{\xi(1-\bar{s})}$, whence we obtain
$N(T) = \frac{1}{\pi}[\xi(s)]_{C'}$, where $[\xi(s)]_{C'}$ denotes the variation of the amplitude of $\xi(s)$ on the part C' of C, from $(2,0)$ to $(\frac{1}{2}, iT)$. We get $[s(s-1)]_{C'} = \pi$, $[\pi^{-\frac{1}{2}s}]_{C'} = -\frac{1}{2}T\log\pi$, $[\Gamma(\frac{1}{2}s)]_{C'} = \frac{T}{2}\log\frac{T}{2} - \frac{\pi}{8} - \frac{T}{2} + O(\frac{1}{T})$, $[\zeta(s)]_{C'} = O(\log T)$.

As applications of the Riemann-von Mangoldt formula we have:
(i) $\gamma_n \sim \frac{2\pi n}{\log n}$ as $n \to \infty$
(ii) $\gamma_{n+1} - \gamma_n \gg 1/\log n$ for an infinite sequence of values of n,
where $\beta_n + i\gamma_n$ $(n = 1, 2, \ldots)$ denote the zeroes of $\zeta(s)$ in ascending order of γ_n.
These follow from $N(\gamma_n) = n$ and $N(\gamma_n) \sim \frac{\gamma_n}{2\pi}\log\gamma_n$.

In recent times, the function $N(\sigma, T)$ has been introduced. It is defined as the number of zeroes of $\zeta(s)$, $\beta + i\gamma$, with $\beta > \sigma$ and $0 < \gamma \leq T$. If RH. is true, then $N(\sigma, T) = 0$ for $\sigma > 1/2$.
The quasi-Riemann hypothesis asserts that $N(\sigma, t) = 0$ for some $\sigma < 1$. We have $N(T) = N(0, T)$, and so by the Riemann-von Mangoldt formula, we obtain $N(\sigma, T) \ll T\log T$.
Lindelöf has conjectured that $\zeta(\frac{1}{2} + it) \ll t^{\varepsilon}$ for any $\varepsilon > 0$. This would give $N(\sigma, T) \ll T^{2(1-\sigma)}$, and this is called the density hypothesis.
Main results here are due to Ingham: $N(\sigma, T) \ll 2^{K(1-\sigma)}(\log T)^{K'}$, with $K = \frac{3}{2-\sigma}$, Huxley, Montgomery, etc...

Ingham established a direct connection with the problem of estimating the difference between consecutive primes. Namely, if the estimate above holds for $N(\sigma, T)$, then $p_{n+1} - p_n \ll p_n^{(1-1/k)+\varepsilon}$. Thus Ingham's result gave $p_{n+1} - p_n \ll p^{5/8+\varepsilon}$, and $\frac{5}{8}$ has been improved to $\frac{7}{12}$ (Huxley), $\frac{11}{20}$ (Heath-Brown). Prior to Ingham, Hoheisel had given $1 - \frac{1}{3300}$, Heilbronn $1 - \frac{1}{250}$, Chudakov $3/4$

§3. Sieve Methods.

1. Eratosthenes.

He observed that if one deletes from the set $1, 2, \ldots, n$ first all multiples of 2, then all multiples of 3, and so on up to $\sqrt{n}$, then apart from 1, precisely the primes between $\sqrt{n}$ and n remain. This can be expressed by the formula: $\pi(n) - \pi(\sqrt{n}) + 1 = \sum_{d|P} \mu(d)[n/d]$, where $P = \prod_{p \le \sqrt{n}} p$.

In fact, the left-hand side is: $\sum_{\substack{m=1 \\ (m,P)=1}}^{n} 1 = \sum_{m=1}^{n} \sum_{d|(m,P)} \mu(d) = \sum_{d|P} \mu(d) \sum_{\substack{m=1 \\ d|m}}^{n} 1 = \sum_{d|P} \mu(d)[n/d]$

The above formula cannot be used however to give a non-trivial estimate for $\pi(n)$, for if one replaces $[n/d]$ by n/d then one gets: $\pi(n) - \pi(\sqrt{n}) + 1 = n \sum_{d|P} \mu(d)/d + O(\sum_{d|P} |\mu(d)|)$
$= n . \prod_{p \le \sqrt{n}} (1 - \frac{1}{p}) + O(d(P))$, where $d(P) = \#$ numbers dividing P.
Now, $\prod_{p \le \sqrt{n}} (1 - \frac{1}{p})^{-1} \ge \sum_{m \le \sqrt{n}} 1/m \ge \log \sqrt{n}$. (In fact, by Theorem 4 in §1, we know that $\prod_{p \le \sqrt{n}} (1 - p^{-1})^{-1} = C \log \sqrt{n} + O(1)$). Whence, the first term on the right hand side is $\le \frac{2n}{\log n}$
The error term is $O(2^{\pi(\sqrt{n})})$ and since $\pi(\sqrt{n}) \sim 2\sqrt{n}/\log n$, this is of larger order of magnitude than the main term.

However, if one sieves only up to $\frac{1}{2}\log_2 n$, ie $P = \prod_{p \le \frac{1}{2}\log_2 n} p$ then the main term is $\le n / \log(\frac{1}{2}\log_2 n) \le (1+\varepsilon)n/\log\log n$, and the error term is $O(2^{\frac{1}{2}\log_2 n}) = O(\sqrt{n})$.
Hence we have, for any $\varepsilon > 0$ and all sufficiently large n, $\pi(n) - \pi(\sqrt{n}) \le (1+\varepsilon)n/\log\log n$.

2. Selberg

We shall sieve the set $1, \ldots, N$. Let M be an integer with $1 \le M \le N$. (In Eratosthenes, $M = \sqrt{N}$, in refined version, $M = \frac{1}{2}(\log_2 N)$). Put $P = \prod_{p \le M} p$. Let $f(n)$ be any function on the integers, and taking non-negative integer values. (In Eratosthenes, $f(n) = n$). We wish to estimate S, the number of $f(n)$ $(1 \le n \le N)$ not divisible by a prime $p \le M$. We have
$S = \sum_{\substack{n=1 \\ (f(n),P)=1}}^{N} 1 = \sum_{n=1}^{N} \sum_{d|(f(n),P)} \mu(d) = \sum_{d|P} \mu(d) S(d)$, where $S(d)$ is the number of $f(n)$ divisible by d.

Selberg observed that if $\lambda(n)$ is any real function with $\lambda(1) = 1$, then for any integer k $\left(\sum_{d|k} \lambda(d)\right)^2 \ge \sum_{d|k} \mu(d)$. Now we can write $\left(\sum_{d|k} \lambda(d)\right)^2 = \sum_{d|k} \rho(d)$, where $\rho(d) = \sum_{\substack{d_1|P \\ d = \operatorname{lcm}\{d_1, d_2\}}} \sum_{d_2|P} \lambda(d_1)\lambda(d_2)$.

Here, we are assuming that $k | P$.
Hence we have $\sum_{d|k} \mu(d) \le \sum_{d|k} \rho(d)$, and so $S \le \sum_{n=1}^{N} \sum_{d|(f(n),P)} \rho(d) = \sum_{d|P} \rho(d) S(d)$.

This is the crux of Selberg's upper bound sieve.

We know that if $f(n_1)\equiv f(n_2) \pmod d$ when $n_1 \equiv n_2 \pmod d$, as we shall assume, then $S(d)$ is "about" $N\nu(d)$, where $\nu(d)d$ is the number of $f(n)$ $(1\le n\le d)$ divisible by d. (Thus when $f(n)=n$ we have $\nu(d)={}^1/d$). Hence we put $S(d)=N\nu(d)+R(d)$, which gives

$$S \le N\sum_{d|P}\nu(d)\rho(d) + \sum_{d|P}|\rho(d).R(d)|$$

Now, we can select the values of λ (subject to $\lambda(1)=1$) so that the first sum is minimal. Since ν is multiplicative, we have $\sum_{d|P}\rho(d)\nu(d) = \sum_{d_1|P}\sum_{d_2|P}\lambda(d_1)\lambda(d_2)\cdot\frac{\nu(d_1)\nu(d_2)}{\nu([d_1,d_2])}$

Thus the problem is essentially that of minimalising a quadratic form.

We write: $g(n)=\sum_{d|n}\mu(n/d)/\nu(d)$, (ie $g(n)=\frac{1}{\nu(n)}\cdot\prod_{p|n}(1-\nu(p))$), and $h(\delta)=\sum_{\substack{d|P\\ \delta|d}}\lambda(d)\nu(d)$.

Noting that $\frac{1}{\nu(n)}=\sum_{d|n}g(d)$, we get $\sum_{d|P}\rho(d)\nu(d)=\sum_{\delta|P}g(\delta)h(\delta)^2$

(ie, we have "diagonalised" the quadratic form).

By Möbius inversion, we have $\lambda(d)\nu(d)=\sum_{\substack{\delta|P\\ d|\delta}}\mu(\delta/d)h(\delta)$, whence the condition $\lambda(1)=1$ gives $\sum_{\delta|P}\mu(\delta)h(\delta)=1$.

We now take $h(\delta)=0$ for $\delta>M$ and $h(\delta)=\frac{\mu(\delta)}{g(\delta)}\cdot\frac{1}{Q}$ for $\delta\le M$, where $Q=\sum_{\substack{\delta|P\\ \delta\le M}}{}^1/g(\delta)$.

With these definitions, we have $\sum_{\delta|P}g(\delta)h(\delta)^2={}^1/Q$, and this is the minimal value. (Indeed, LHS is $\sum_{\substack{\delta|P\\ \delta\le M}}\frac{1}{g(\delta)}(g(\delta)h(\delta)-\mu(\delta)/Q)^2+{}^1/Q$ – we are "completing the square")

Hence we conclude that $S\le {}^N/Q+\sum_{d|P}|\rho(d)R(d)|$.

It remains to estimate the error term. By the definition of ρ we have $\sum_{d|P}|\rho(d)R(d)|\le\sum_{d_1|P}\sum_{d_2|P}|\lambda_1(d_1)\lambda_2(d_2)R(\{d_1,d_2\})|$ and $R(\{d_1,d_2\})\le d_1d_2\nu(d_1)\nu(d_2)$.

Hence, $\sum_{d|P}|\rho(d)R(d)|\le(\sum_{d|P}|d\lambda(d)\nu(d)|)^2$.

But $\sum_{d|P}|d\lambda(d)\nu(d)|\le\sum_{d|P}d.\sum_{\substack{\delta|P\\ \delta\le M\\ d|\delta}}\left(\frac{1}{Q(g(\delta))}\right)\le\sum_{\substack{d|P\\ d\le M}}{}^d/g(d)\le M\sum_{d|P}{}^1/g(d)$.

Thus we obtain:

Theorem 7: $S\le {}^N/Q+M^2(\sum_{d|P}{}^1/g(d))^2$

3. Applications.

(i) Brun-Titchmarsh inequality. We recall that N is a large integer and $1\le M\le N$. Further, $P=\prod_{p\le M}p$, $Q=\sum_{\substack{\delta|P\\ \delta\le M}}{}^1/g(\delta)$, $g(n)=\frac{1}{\nu(n)}\prod_{p|n}(1-\nu(p))$. Theorem 7 asserts that $S\le {}^N/Q+M^2(\sum_{d|P}{}^1/g(d))^2$

Assume $M\le N^{\frac12-\varepsilon}$ where $\varepsilon>0$, given. We estimate $\pi(M+N)-\pi(M)$. The result is a special case of the Brun-Titchmarsh inequality. We take $f(n)=M+n$ (note the congruence condition is satisfied, and indeed this is so for any polynomial f). We have $S\ge\pi(M+N)-\pi(M)$

We need an estimate for Q from below. We have $Q=\sum_{\delta\le M}\mu^2(\delta)/g(\delta)$, and $\nu(d)={}^1/d$, whence $g(n)=\varphi(n)$, the Euler totient function.

Hence $Q\ge\sum_{m\le M}{}^1/m$ since $\frac{1}{\varphi(\delta)}=\prod_{p|\delta}(\frac1p+\frac{1}{p^2}+\cdots)$, so $Q\gg\log M$. It remains to estimate $\sum_{\delta|P}{}^1/g(d)$.

But $g(d)=\varphi(d)$ in this example and so $\sum_{\delta|P}{}^1/g(\delta)\le\prod_{p\le M}(1+\frac{1}{\varphi(p)})=\prod_{p\le M}(1+\frac{1}{p-1})$, and by Theorem 3 of chapter 1, $\prod_{p\le M}(1-\frac1p)^{-1}\ll\log M$. Hence we can conclude that $S\le\frac{N}{\log M}+cM^2(\log M)^2$. $(c>0)$

Thus, provided that $N^{\frac{1}{2}-\varepsilon} \ll M \ll N^{\frac{1}{2}-\varepsilon}$ we get $\pi(M+N) - \pi(M) \le (2+\varepsilon)N/\log N$.

(ii) Twin prime conjecture. This asserts that there are infinitely many n such that $p_{n+1} - p_n = 2$. We shall use Theorem 7 to give an upper bound for the number of twin primes $p, p+2$ with $p \le N$.
We take $f(n) = n(n+2)$. Then $\nu(p) = \frac{2}{p}$, $g(p) = \frac{p}{2} - 1$, so $\frac{1}{g(p)} = \frac{2}{p} + (2/p)^2 + \cdots$, and $Q \ge \sum_{m \le M} d(m)/m \ge \left(\sum_{m \le \sqrt{M}} 1/m\right)^2$.
Further, $\sum_{d|P} 1/g(d) \le \prod_{p \le M}\left(1 + \frac{1}{g(p)}\right) = \prod_{p \le M}\left(1 + \frac{2}{p-2}\right) = \prod_{p \le M}\left(1 - \frac{2}{p}\right)^{-1}$ and by Theorem 3 this gives $\sum_{d|P} \frac{1}{g(d)} \ll \log M$.
Thus, taking $M = N^{1/4}$ we find that the number of $p \le N$ with $p+2$ prime is $\ll N/(\log N)^2$

Hence the series $\sum_p \frac{1}{p}$ taken over all twin primes p converges (cf: $\sum_{\text{all } p} \frac{1}{p}$ divergent).

(iii) Goldbach Conjecture This asserts that even integer (>2) is the sum of two primes. The same argument as in (ii) above applied to the function $f(n) = n(N-n)$ gives the estimate $N/(\log N)^2$ for the number of primes p with $N-p$ also prime.

To some extent, the theory can be developed to give lower bounds for quantities analogous to S. Plainly if one could find non-trivial lower bounds for S itself in cases (ii) and (iii), this would amount to a solution of both the twin prime and Goldbach conjectures. This best results to date in this connection are due to Chen Jing-Run. He showed that the equation $2N = p + P_2$ is soluble for sufficiently large N, where p is a prime, and P_2 is a number with at most two prime factors. A similar result holds in connection with the twin prime conjecture, namely $\exists$ infinitely many (p, P_2). Before Chen, Rennyi had obtained $2N = p + P_r$, where P_r is a number with at most r prime factors ($r \le 9$) and $r = 3$ followed from results on the large sieve.
Also, there is a famous result of Vinogradov asserting that every sufficiently large odd integer N satisfies $N = p_1 + p_2 + p_3$ with p_i prime. The proof is an application of the Hardy-Littlewood method in additive number theory.

4. The Large Sieve.

This was introduced by Linnich in 1941 and developed by Rennyi subsequently. The inequalities obtained by Rennyi weregiven in optimal form by Roth in 1964. He proved:

Theorem 8: Let $\mathcal{A}$ denote a set of integers between 1 and N inclusive. Then for any positive integer Q we have: $\sum_{p \le Q} p \sum_{a=1}^{p} \left(Z(a,p) - Z/p\right)^2 \ll (Q^2 + N)Z$, where $Z(a,p)$ denotes the number of integers in $\mathcal{A}$ which are congruent to a mod p.

This sphere of ideas was applied by A.I. Vinogradov (a different one...) and E. Bombieri in 1965 to show that the Riemann Hypothesis is true "in an average sense". The basic result here is:

Theorem 9: Let $\psi(x; a, q) = \sum_{\substack{n \le x \\ n \equiv a \bmod q}} \Lambda(n)$. Then for any $A > 0$, $\exists$ $B > 0$ such that

$$\sum_{q \le x^{1/2}/(\log x)^B} \max_{y \le x} \max_{(a,q)=1} \left|\psi(y; a, q) - y/\varphi(q)\right| \le x/(\log x)^A.$$

The proof of the prime number theorem and its refinements gives: $\psi(x;a,q) = \frac{x}{\varphi(q)} + O(x e^{-c(\log x)^{1/2}})$, for fixed a,q, and indeed Siegel-Walfisz proved that this holds uniformly for $q \leq (\log x)^A$.
If the generalised Riemann Hypothesis holds, then $\psi(x;a,q) = x/\varphi(q) + O(x^{1/2}\log x)$.
Thus Theorem 9 establishes this conjecture in an average sense.
The Siegel-Walfisz conjecture depends on the theory of L-functions originated by Dirichlet in the last century. He proved that there exist infinitely many primes in the arithmetical progression $a, a+q, a+2q, \dots$, $(a,q)=1$. He introduced characters $\chi(n)$ such that $\chi(n) \equiv \chi(n')$ if $n \equiv n' \bmod q$ and $\sum_{\substack{n=1 \\ (n,q)=1}}^{q} \chi(n) = \varphi(q)$ if $\chi = \chi_0$, $=0$ if $\chi \neq \chi_0$, and $\sum_{\chi} \chi(n) = \begin{cases} \varphi(q) & \text{if } n \equiv 1 \bmod q \\ 0 & \text{if } n \not\equiv 1 \bmod q \end{cases}$

Then the L-function is defined as $L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$
χ_0 is the principal character: $\chi_0(n) = \begin{cases} 1 & \text{if } (n,q)=1 \\ 0 & \text{otherwise.} \end{cases}$

So, $L(s,\chi_0) = \zeta(s) \prod_{p|q} (1-p^{-s})$.

The theory of L-functions has been developed in an analogous way to the theory of the Riemann Zeta Function.