Mr G's Little Book on

The Prime Number Theorem

including Goodhand's conjecture

Preface

When I first started to investigate this around 1996 there was barely a commercial internet much less anything mathematical on it. What little information I had on the prime counting function $\pi(n)$ *only extended to about* 10^6.

Nowadays – 2020 – there is a wealth of information available, about 10% readable but boring and the remaining 90% unintelligible.

This booklet is a concise and comprehensible summary of the salient information surrounding the prime number theorem and includes my own conjecture to improve significantly on $n / \ln(n)$ *but still only using an elementary function.*

Definition of a Prime

Primes are the building blocks of all numbers, the individual numbers that multiply together to produce any given number n.

eg $60 = 2 \times 2 \times 3 \times 5$

Historically "1" has been considered a prime number.

From about the start of the 20th century "1" has been excluded and treated as a special case. To compound confusion "2" is also considered somehow "an exception".

I therefore adopt "Goodhand's" definition of a prime.

"A prime number has exactly two divisors" and consequently there is now no need to treat "1" and "2" separately.

The Prime Counting Function p(n)

$\pi(n)$ represents the number of primes up to an including n.

eg $\pi(10) = 4$ (ie 2, 3, 5, 7)

The objective is to determine an expression that might approximate $\pi(n)$ without actually having to determine all the primes up to n.

In short what is the density of primes up to any given value of n and what is the approximate value of the n^{th} prime?

The Prime Number Theorem

Gauss used to be considered the greatest mathematician ever. He added up the number 1 to 100 in seconds on his first day at school and never tired in later life telling people the story. He has more recently been dislodged by Riemann but both men "cut their teeth" on the same problem – deducing an approximation for $\pi(n)$.

Around 1892, when Gauss was 15, he had a book of log tables that gave primes probably up to about 10000. So he quickly counted up to note $\pi(10) = 4$, $\pi(100) = 25$, $\pi(1000) = 168$ and $\pi(10\,000) = 1229$.

He then determined the prime densities approximated to $1/\ln(n)$ and presumably constructed the following table

n	$\pi(n)$	density	$1/\ln(n)$	error
10	4	40%	0.434	-9%
100	25	25%	0.217	+13%
1000	168	16.8%	0.145	+14%
10000	1229	12.29%	0.109	+12%

and could probably see thereafter that the percentage error kept dropping.

This can be reformulated in three ways.

Using ~ to mean "tends to"

n^{th} prime $\sim n \ln(n)$

$\pi(n) \sim n / \ln(n)$

and critically $\pi(n) / (n / \ln(n)) \to 1$ as $n \to \infty$

This is the prime number theorem.

Developments of $n / \ln(n)$

The ratio $n / \ln(n)$ gives values that are not quite large enough so Legendre proposed around 1797 that a bit of tinkering might fix the problem.

He postulated that

$\pi(n) = 1 / (A \ln(n) + B)$

A was soon determined to be "1" but the value of B was termed Legendre's Constant. Legendre gave a value of

$B \approx 1.08366 \ldots$

Now with reference to the enclosed table it can be seen how Legendre came upon this number. Values are given in the 13^{th} column. The least error occurs at

$n \approx 100\,000$

which is probably the highest value for which he knew $\pi(n)$. Thereafter the error using B peaks at about $n = 10^{14}$ which was well beyond the calculating power of the late 18^{th} century.

Chebyshev proved in 1849 that if the limit exists then B = 1.

ie $\pi(n) = 1 / (\ln(n) + 1)$

Even with present computing power $\pi(n)$ is only known exactly to $n = 10^{27}$.

Legendre's formula still beats Chebyshev so it was an impressive stab at the problem. Presumably at much higher values of n, Chebyshev will improve on Legendre.

The Logarithmic Integral li (n)

Around 1811 Gauss was acknowledging he was working on the li (n) function where

$\mathrm{li}(n) = \int_{n=0}^{\infty} dt / t$ such that

$\pi(n) \sim \mathrm{li}(n)$

and was a big improvement on

$\pi(n) \sim n / \ln(n)$

This is clear with reference to the enclosed table. Although the absolute error of $\mathrm{li}(n) - \pi(n)$ <u>appears</u> to grow without limit its relative error approaches zero. *(see Skewes number)*

On the enclosed table I used the infinite series to the first 20 terms to calculate

$\mathrm{li}(n) = \gamma + \ln(\ln(n)) + (\ln n)/1.1! + (\ln n)^2/2.2! + (\ln n)^3/3.3! \ldots$

for the first 4 rows of the 7^{th} column.

γ is the Euler- Mascheroni constant.

Thereafter the series converges too slowly to be of use and I actually derived li (n) from the offset integral.

Then from 10^{18} I actually entered

li (n) – π(n) directly and deduced li (n) from that value.

The limitation of Excel to 15 significant figures and floating point decimal also presented problems.

All this admittedly was a bit *"cart before the horse"* but the principle had been demonstrated.

The Logarithmic Integral Li (n)

There is a minor technical problem with li (n) which has a singularity at

n ≈1.451.

To avoid this Cauchy proposed the offset logarithmic interval

$n=_2\int^{\infty} {}^{dt}/_{t}$

Thus Li (n) = li (n) – li (2).

Though the difference is miniscule Li (n) is also a better approximation to π(n). From row 5, I directly calculated Li (n) from the expression

Li (n) = ${}^{n}/_{\ln(n)}$ { 1 + ${}^{1!}/_{\ln(1)}$+ ${}^{2!}/_{\ln(2)}$ + ${}^{3!}/_{\ln(3)}$ … } *to 20 terms.*

The Riemann Hypothesis

The unsolved problem in Mathematics is RH which states

"the Riemann zeta function has its zeros only at the negative even integers and complex numbers with real part ½".

The zeta function is

$\xi(s) = {}_{n=1}\Sigma^{\infty} {}^{1}/_{n\hat{}s} = {}^{1}/_{1\hat{}s} + {}^{1}/_{2\hat{}s} + {}^{1}/_{3\hat{}s} \ldots$

This relates to the prime number theorem because Euler recast this function in terms of primes.

$\xi(s) = {}_{p\ prime}\Pi\ {}^{1}/_{1-p\hat{}-s}$

$= ({}^{1}/_{1-2\hat{}-s}) \cdot ({}^{1}/_{1-3\hat{}-s}) \cdot ({}^{1}/_{1-5\hat{}-s}) \ldots$

using a process similar to unique factorisation.

The zeta function thus directly connects to primes. Next we extend the function to s complex. This means that through "complex analysis" information on the properties of the function are determined by the zero values.

The upshot is the proof of this hypothesis sets boundary conditions on the maximum absolute value of π(n) – Li (n) in terms of n.

That is as accessible as I can make it.

Skewes Number

Li (n) appears consistently to over-estimate $\pi(n)$. However in 1933 Skewes by assuming the truth of the Riemann Hypothesis proved that

$$\pi(n) < \text{Li}(n)$$

is violated for some value

$$n < 10^{10^{10^{34}}}.$$

That is the value Li (n) eventually drops below $\pi(n)$ but at an incomprehensibly high number.

It has been also been shown that Li (n) crosses $\pi(n)$ an infinite number of times. Therefore it is pointless trying to tweak Li (n) and it just remains to determine how the error of Li (n) in determining $\pi(n)$ behaves.

From the nineteen fifties onwards the Guinness Book of Records would give Skewes number as the highest number derived naturally from a mathematical calculation. That original estimate of Skewes has now been reduced to around 10^{316} still assuming the truth of RH. The value reduces as more zeros are identified on the critical line.

As there are only 10^{79} protons in the Universe no computer will ever be able to calculate this exactly

Goodhand's Conjecture

The problem with li (n) and Li (n) is that they cannot be expressed by an elementary expression.

However on constructing a table of functional values I noticed that

The error in the approximation of

$$n / \ln(n) \text{ to } \pi(n)$$

is approximately equal to $\pi(n) / n$

This can be seen by comparing column 3 with column 5

I therefore equated these two expressions, rearranged to a quadratic and solved to produce the elementary expression

$$\pi(n) \approx \tfrac{1}{2}n \{ 1 - \sqrt{(1 - 4/\ln(n))} \}$$

It can be seen that this expression considerably improves on Legendre, Cauchy and consequently Gauss's original determination of the prime number theorem but disappointingly not Li (n).

Despite an extensive internet search and queries on Maths websites I can find no reference to this so I claim this conjecture as mine.

Further Concepts

Prime numbers at first sight seem to occur almost at random. Like Bristol buses you wait ages for one and then you'll get two in quick succession.
Euclid showed that the number of primes is infinite.
Then there is the concept of twin primes. However it is only conjectured that the number of twin primes is infinite. This remains one of the most elusive unsolved problems in number theory.
However even though both may be "infinite" there is still a difference in "magnitude".
The summation of the inverse of primes is also infinite but the summation of the inverse of twin primes is approximately 1.90216054
This is known as Brun's constant.
Both the concept and the calculation of this does defy comprehension. *(although I can sort of see how it could come about if the spacing between twins gets larger than the next decimal place)*

Final Comments

As previously stated primes are the building blocks of all numbers. The Fundamental Theorem of Arithmetic states that any number's decomposition into primes is unique.
So if the behaviour of primes can be securely nailed down then a massive raft of unproved conjectures become proven.
In fact it is now often assumed at the beginning of a partial proof *"given the truth of RH then …".*
The smart money is that RH is definitely true. However that still leaves open the question of whether it is provable.
Truth and provability are definitely not the same. There exist theorems which can be assumed either to be true or false strictly within the axioms of mathematics and the totality will still be consistent but not provable.

Also available in this series is

- On My TI Calculator what's the difference between Sx and σx?

  It's not what I thought for the first 40 years of the scientific calculator
- Beyond Pascal – Multinomials and Dice Throwing

  How a lower set exercise in dice throwing led to the discovery of multinomials
- Conditional Probability and Bayes Theorem

  An investigation into the pitfalls of medical screening
- Hypercomplex Numbers

  Instead of making $i^2 = -1$ as in complex numbers what if we just make $i^2 = 1$
- Propositional Calculus

  Sherlock Holmes was the great inductive detective but not infallible
- The Harmonic Triangle

  How investigating harmonic triangles led to the discovery of a universal series summation formula