ILLUSTRATING THE PRIME NUMBER THEOREM VIA MATHEMATICA

Werner Horn

Abstract. This paper gives an introduction to the Prime Number Theorem via the use of a computer algebra system. Special consideration is given to the fact that the results might mislead the unsuspecting person to expect an asymptotically linear relationship. The article is accessible to anyone who had two semesters of calculus.

1. Introduction. Two intriguing questions prompted the investigation below. The first one is a mathematical question, namely, if one picks a 10–digit integer at random, what is the probability that this integer is a prime number? This question can be answered by just counting the number of primes with 10 or fewer digits and dividing the result by 10^{10}. While this is an arduous process, a computer algebra system such as *Mathematica* will certainly help us do this. Before such luxuries, mathematicians were searching for quick approximations for the answer to this and related questions. These approximations are collectively known as the Prime Number Theorem, which was conjectured in the last decade of the 18th century.

The second even more intriguing question is, how could Gauss and Legendre guess the correct approximate answer to the first question based on the empirical evidence available at this time? Or, loosely speaking, can we read what was on the mind of the teenager Gauss 200 years ago? While it is impossible to know the correct answer to these questions we will give some indications how Gauss might have arrived at his result.

Unfortunately, a rigorous treatment of the Prime Number Theorem is reserved to advanced students of mathematics. However, with the use of computer algebra systems the content (not the proof) of this theorem is accessible to anyone who had two semesters of calculus. The computer algebra system will therefore allow us to get a taste of the beauty of higher mathematics. Furthermore, it will show the interested student a use of calculus which is beyond the usual examples in every book. We will use *Mathematica* to illustrate the Prime Number Theorem and to introduce several special functions. The example was used in a class where mathematics students were introduced to the use of computer algebra systems. Most of the students had just finished a second semester of calculus before taking this class. The class consisted of ten projects where computer algebra systems were used. The Prime Number Theorem was one of the last projects. The project at least raised the awareness of students to this subject.

2. The Function π(x) and the Statement of the Prime Number Theorem. Most often the Prime Number Theorem is stated in terms of the function $\pi(x)$, which is defined for positive real numbers and is simply the number of primes which are less than or equal to x. So, for example, $\pi(3.5) = 2$, since there are two primes, 2 and 3, which are less than or equal to 3.5. $\pi(4.9) = 2$ and $\pi(5) = 3$, i.e. the function jumps at 5, or at any prime number for that matter; and $\pi(10^{10})/10^{10}$ is the explicit answer to the first of the questions of the introduction. In general, $\pi(x)$ is an upper–semi–continuous step function, which has a jump of magnitude one at each prime number. Another commonly used function is the function $P(n)$, which denotes the nth prime number, so $P(1) = 2$ and $P(4) = 7$, etc. The domain of $P(n)$ is the positive integers.

$P(n)$ and $\pi(x)$ are closely related; indeed one has $\pi(P(n)) = n$ and if p is a prime number one also has $P(\pi(p)) = p$ (i.e. these functions are inverses of each other when restricted to the primes).

The Prime Number Theorem gives asymptotic expressions for the functions $\pi(x)$ and $P(n)$. We say that $f(x) \approx g(x)$ (f is asymptotic to g), if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$. Using this we can state the different forms of the Prime Number Theorem.

1. Gauss' Statement:

$$\pi(x) \approx \int_2^x \frac{1}{\ln t}\, dt \ .$$

2. Legendre's Statement:

$$\pi(x) \approx \frac{x}{\ln x - 1.08366} \ .$$

3. Modern Statement:

$$\pi(x) \approx \frac{x}{\ln x} \ .$$

4. Statement for $P(n)$: $\qquad P(n) \approx n \ln n.$

The fourth statement says that the nth prime number is approximately of size $n \ln n$ for large n, e.g. the 10,000th prime is approximately as big as 10,000 ln 10,000 = 92,103.4 (actually the 10,000th prime is 104,729).

It is a worthwhile exercise in the use of l'Hospital's Rule to prove the equivalence of the first three statements. To show that the last statement is equivalent to the first three is a little more difficult.

While the names of C. F. Gauss and A. M. Legendre are both connected to this theorem, neither actually proved it. Both developed their conjectures in the last decade of the 18th century (when Gauss was still a teenager), based on the empirical evidence for $x < 1,000,000$. This conjecture was driving the development of mathematics, especially complex analysis, for most of the 19th century. The theorem was finally proven independently by J. Hadamard and C. J. de la Vallee–Poussin in the last decade of the 19th century. Both proofs relied on earlier works by Chebychev and Riemann. In 1949 A. Selberg and P. Erdős came up with a new (elementary) proof of the Prime Number Theorem. Here, elementary means not requiring complex analysis. The interested reader can find a proof of the theorem in [4] and additional information in [1, 2, 5, 6].

The right hand side of Gauss' statement is closely related to another special function, the logarithmic integral function $li(x)$ defined by:

$$li(x) = \int_0^x \frac{dt}{\ln t} \ .$$

Using this the Prime Number Theorem can be stated as

$$\pi(x) \approx li(x) - li(2) \ .$$

3. The use of Mathematica. All the necessary functions are available as special functions in *Mathematica*. The expressions

`PrimePi[x], Prime[k], LogIntegral[x],`

will produce $\pi(x)$, $P(k)$, and $li(x)$, respectively.

However, *Mathematica* will actually use some asymptotic algorithms to evaluate $\pi(x)$ and $P(k)$ for large values of x and k [7]. These asymptotic algorithms are somewhat more sophisticated than just the Prime Number Theorem, however, the results for very large values of x should still be considered with care. But since we only want to use the data produced by *Mathematica* as an illustration we should not be too concerned with this fact. To achieve some peace of mind we compared the results of

`PrimePi[x]`

for some values of $x \leq 100,000,000$ with the tabulated exact values. *Mathematica* did produce the exact answer in the cases we checked. One of the advantages of using *Mathematica* is that we can represent the results graphically which should give us an edge over the tables which were available to Gauss and Legendre. To obtain the graph of the function $P(n)$ one uses the command

```
ListPlot[Table[Prime[n], {n,200}]]
```

which first tabulates the first 200 primes and then represents them in a coordinate system. The result of this is shown in Figure 1. However, this method of plotting functions, which is defined on the integers becomes very time and memory consuming if we want to allow large domains, i.e. when 200 is replaced by 1,000,000 or any other huge number.

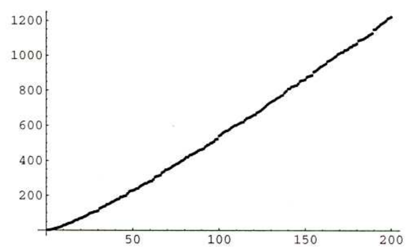

Figure 1: The first 200 primes; the graph of $P(n)$ for $n \leq 200$.

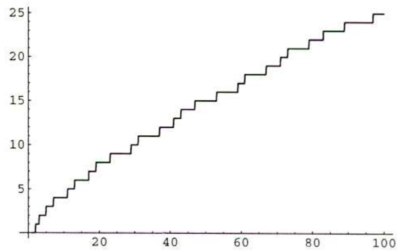

Figure 2: The graph of $\pi(x)$ for $1 \leq x \leq 100$.

To get a quicker graphical representation we extend the function $P(n)$ to all real numbers, which are greater or equal to one via $F(x) = P([x])$, where the bracket denotes the usual greatest integer function. In *Mathematica* one writes

```
F[x_]:=Prime[Floor[x]];
```

For functions like $F(x)$, $\pi(x)$, or $li(x)$, which are defined on the positive real numbers, the usual `Plot` command of *Mathematica* can be applied. This command samples the function values over the interval of plotting and interpolates the function between the sampling points. The function is only evaluated at a few hundred points. If we used `ListPlot` for $n \leq 1,000,000$, $P(n)$ would be evaluated at 1,000,000 points. The interpolation process will also produce a smoother graph than one would obtain via `ListPlot`. The command `Plot` was used to obtain the graph of $\pi(x)$ for $1 \leq x \leq 100$ shown in Figure 2. However, since we are only interested in the large scale behavior, the smoother graph contains all the necessary information.

4. The Large Scale Behavior of $\pi(\mathbf{x})$ and $\mathbf{P}(\mathbf{n})$. In this section we want to represent the behavior of $\pi(x)$ and $P(n)$ for very large values of x and n, respectively. In particular, we will try to follow Gauss and Legendre and guess the asymptotic behavior of these functions. We will make use of our technological advantage over the late 18th century and use more data than Gauss or Legendre could ever dream. Figure 3 shows the graph of $\pi(x)$ for $1 \leq x \leq 1,000,000,000$ (1 billion!). Figure 4 shows $P(n)$ for $1 \leq n \leq 50,000,000$ (the 50 millionth prime is around 1 billion).

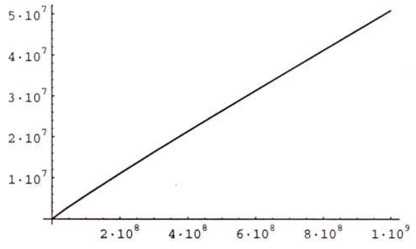

Figure 3: The graph of $\pi(x)$ for $x \leq 1.0 \times 10^9$.

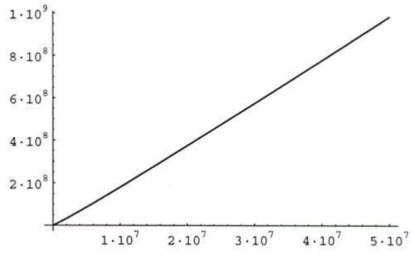

Figure 4: The graph of $P(n)$ for $n \leq 5.0 \times 10^7$.

Looking at Figures 3 and 4 the untrained eye will detect straight lines. To study the large scale behavior closer, we list 10 values of x and $\pi(x)$ with x between 10,000,000 and 100,000,000 in Table 1.

Gauss himself introduced a method of finding the line of best fit for a given set of data. Applying the method of linear regression to the values of x and $\pi(x)$ in Table 1, we obtain that the line $y = ax + b$ with $a = 0.05654$ and $b = 149,975.67$ is the line of best fit for these data. A good fit of a line is indicated by a coefficient of correlation ρ which is close to 1. The coefficient of correlation of the above data is $\rho = .99988$! This will convince most students that $\pi(x) \approx ax + b$, i.e. the function $\pi(x)$ is approximately linear for large values of x. This would also mean that the density of primes is approximately constant for large numbers, i.e. there is approximately the same number of primes in every interval of equal length with large lower bounds. For example, the interval $[2 \times 10^7, 2.5 \times 10^7]$ should contain approximately the same number of primes as the interval $[1.0 \times 10^{10}, 1.0 \times 10^{10} + 5.0 \times 10^6]$. Of course, this is not correct but an extremely tempting conjecture. A similar investigation for $P(n)$ will yield similar results.

We close this section by graphically representing two statements of the Prime Number Theorem. First we consider the modern statement. To plot this we enter `Plot[PrimePi[x]/(x/Log[x]), {x,10,100000000}]`

x	$\pi(x)$	x	$\pi(x)$
10,000,000	664,579	60,000,000	3,562,115
20,000,000	1,270,607	70,000,000	4,118,064
30,000,000	1,857,859	80,000,000	4,669,382
40,000,000	2,433,654	90,000,000	5,216,954
50,000,000	3,001,134	100,000,000	5,761,455

Table 1. Some values of x and $\pi(x)$.

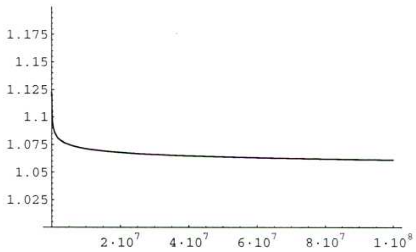

Figure 5: The graph of $\pi(x)\ln(x)/x$ for $x \leq 1.0 \times 10^8$.

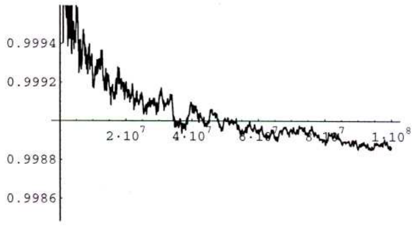

Figure 6: The graph of $\pi(x)(\ln x - 1.08366)/x$ for $x \leq 1.0 \times 10^8$.

This will produce the graph shown in Figure 5. It certainly seems that the function in Figure 5 is converging and that its limit is greater or equal to 1. However, the conclusion that the limit is equal to 1 seems to be impossible from that graph. Neither Gauss nor Legendre considered the limit of

$$\frac{\pi(x)\ln x}{x}$$

which is shown in Figure 5. Legendre considered the limit of

$$\frac{\pi(x)(\ln x - 1.08366)}{x}.$$

It is easy to see that the constant 1.08366 will lower the graph and show quicker convergence. This is shown in Figure 6.

5. Gauss' Statement of the Prime Number Theorem. In this section we investigate the statement

$$\pi(x) \approx \int_2^x \frac{1}{\ln t}\,dt,$$

which is due to the teenage Gauss. The sequence of commands

```
M:=LogIntegral[2];
Plot[PrimePi[x]/(LogIntegral[x]-M), {x,10,100000000}]
```

will produce the graph of

$$\frac{\pi(x)}{\int_2^x \frac{1}{\ln t}\,dt}$$

shown in Figure 7.

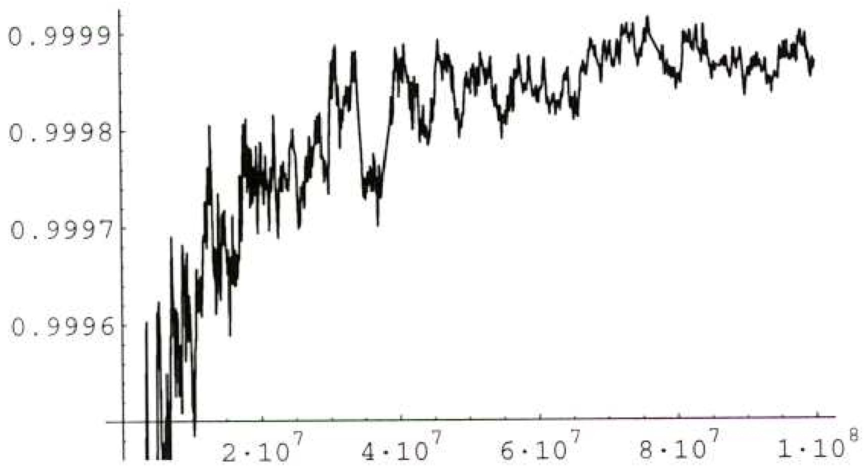

Figure 7. The graph of $\pi(x)/\int_2^x \frac{dt}{\ln t}$ for $x \leq 1.0 \times 10^8$.

The graph is oscillating quite rapidly, however, when considering the scale of the vertical axis, one sees that this expression is converging quicker than either of the previous expressions. The oscillations are the same as the ones in Figure 6. They are due to $\pi(x)$ which is in both graphs. These oscillations are also present in Figure 5, but not visible due to the scale. An interesting question connected with this is how Gauss could have come up with such a "weird" expression. The logarithmic integral is usually not on the top of one's mind even if one looks at prime numbers. However, maybe Gauss did not even look at the function $\pi(x)$ at all, but at rates of change. In his letter to Encke [3] (in German) or [6] (translated) he writes, "I soon observed that for all oscillations this frequency is on the average near an inverse ratio to the logarithm, so that the number of all primes under a given n is determinable approximately by integral

$$\int \frac{dn}{\log n},$$

where we mean the natural logarithm." He says in this sentence that he investigated the frequencies of primes, which would be the derivative of $\pi(x)$, if it were differentiable.

n	$\frac{P(10{,}000(n+1))-P(10{,}000n)}{10{,}000}$	$\ln P(10,000n)$	$\ln P(10,000(n+1))$
1	12.0008	11.5591	12.3227
2	12.564	12.3227	12.7668
3	12.9532	12.7668	13.0814
4	13.2044	13.0814	13.3244
5	13.482	13.3244	13.5235
6	13.5604	13.5235	13.6904
7	13.8002	13.6904	13.8357
8	13.9144	13.8357	13.9635
9	14.0186	13.9635	14.0777
10	14.134	14.0777	14.1809

Table 2. Difference quotients for $P(n)$.

To further investigate this subject consider the expression

$$\frac{P(m) - P(n)}{m - n},$$

which is a difference quotient for the function $P(n)$. It signifies the average rate of change for this function between n and m. In the terminology of prime numbers it is the average difference of two consecutive primes between the nth and the mth prime. This expression can be easily computed from a list of prime numbers, however, we will use *Mathematica* to save some time. The command line

```
Table [N[(Prime[10000*(n+1)]-Prime[10000*n])/10000,{n,10}]
```

produces a list containing the values of this expression for $m - n = 10,000$ and $n = 10,000,\ 20,000,\ \cdots\ 100,000$. The results of this computation are shown in the second column of Table 2.

One sees that this difference quotient is slowly increasing with n. In the third and fourth column of Table 2, we listed $\ln(10,000n)$ and $\ln(10,000(n+1))$ for comparison. We see that for the few examples in the table we have

$$\ln P(n) \leq \frac{P(m) - P(n)}{m - n} \leq \ln P(m).$$

If m and n are large and not too far apart, $\ln P(m)$ and $\ln P(n)$ are rather close. The table shows this even for $n \leq 100,000$. One can imagine the conjecture that the average difference of two consecutive primes $P(n+1) - P(n)$ is approximately equal to the $\ln P(n)$. The word average is important here since the actual differences can vary a lot.

Now, if we use the inverse function relationship between $P(n)$ and $\pi(x)$ we get

$$\ln x \leq \frac{y - x}{\pi(y) - \pi(x)} \leq \ln y,$$

or reciprocally

$$\frac{1}{\ln y} \leq \frac{\pi(y) - \pi(x)}{y - x} \leq \frac{1}{\ln x}$$

i.e. the rate of change for $\pi(x)$ is approximately equal to $\frac{1}{\ln x}$. From this "differential" statement of the Prime Number Theory it is an easy step to the integral statement

$$\pi(x) \approx \int_2^x \frac{1}{\ln t}\, dt.$$

6. Conclusion. The above exposition shows the difficulty of finding a good asymptotic expression for a function even with a large amount of data. The graph of $\pi(x)$ for $2 \le x \le 1,000,000,000$ would mislead most of us to a linear relationship. However, if one considers a "differential" statement, i.e. a look at the local behavior of $\pi(x)$ the correct asymptotic behavior of $\pi(x)$ is more plausible. Neither Gauss nor Legendre had the means of plotting $\pi(x)$ for an interval containing hundreds of thousands of primes, all they had was a table, which was probably viewed one page at a time. Instead of looking at the entire set they would only see the "local" behavior of a few hundred primes. The difference between the first prime and the last prime on one page would certainly increase as the page number increases and show a similar behavior as in Table 2. The lack of modern methods like *Mathematica* may actually have been a blessing in disguise, because they never saw Figure 3 and was not tempted by the straight line.

Acknowledgement. I would like to thank my colleagues for some interesting discussions on the subject and some valuable hints in the use of *Mathematica*.

References

1. T. M. Apostol, *Introduction to Analytic Number Theory*, Springer Verlag, New York, 1976.

2. C. K. Caldwell, "How Many Primes are There?" Univ. of Tennessee–Martin 1 (1998), <http://www.utm.edu/research/primes/howmany.shtml>.

3. C. F. Gauss, "Letter to Eucke," *Werke Bd. 2,* Göttingen, 1863.

4. D. J. Newman, "Simple Analytic Proof of the Prime Number Theorem," *American Mathematical Monthly*, 87 (1980), 693–696.

5. J. O'Connor and E. F. Robertson, "MacTutor History of Mathematics Archive," Univ. of St. Andrews 1(1998), <http://www-groups.dcs.stand.ac.uk/ ∼history>.

6. J. Pintz, "On Legendre's Prime Number Theorem," *American Mathematical Monthly* 87 (1980), 733–735.

7. S. Wolfram, *The MATHEMATICA Book*, third ed., Mathematica Version 3, Wolfram Media/Cambridge University Press, Champaign, IL, 1996.

Werner Horn
Department of Mathematics
California State University, Northridge
Northridge, CA 91330–8313
email: werner.horn@csun.edu