Example 100. How can you check whether a huge randomly selected number N is prime?

Solution. Compute $2^{N-1} \pmod{N}$ using binary exponentiation. If this is $\not\equiv 1 \pmod{N}$, then N is not a prime. Otherwise, N is a prime or 2 is a Fermat liar modulo N (but the latter is exceedingly unlikely for a huge randomly selected number N; the bonus challenge below indicates that this is almost as unlikely as randomly running into a factor of N).

Comment. There is nothing special about 2 here (you could also choose 3 or any other generic residue).

How many primes are there?

Theorem 101. (Euclid) There are infinitely many primes.

Proof. Assume (for contradiction) there are only finitely many primes: $p_1, p_2, \ldots, p_n$.
Consider the number $N = p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$.
None of the p_i divide N (because division of N by any p_i leaves remainder 1).
Thus any prime dividing N is not on our list. Contradiction.

Just being silly. Similarly, there are infinitely many composite numbers.

Indeed, assume (for contradiction) there are only finitely many composites: $m_1, m_2, \ldots, m_n$.
Consider the number $N = m_1 \cdot m_2 \cdot \ldots \cdot m_n$ (don't add 1).
N is not on our list. Contradiction.

Historical note. This is not necessarily a proof by contradiction, and Euclid (300BC) himself didn't state it as such. Instead, one can think of it as a constructive machinery of producing more primes, starting from any finite collection of primes. □

The following famous and deep result quantifies the infinitude of primes.

Theorem 102. (prime number theorem) Let $\pi(x)$ be the number of primes $\leqslant x$. Then

$$\lim_{x \to \infty} \frac{\pi(x)}{x / \ln(x)} = 1.$$

In other words: $\boxed{\text{Up to } x \text{, there are roughly } x/\ln(x) \text{ many primes.}}$

Examples.

Proportion of primes up to 10^6: $\frac{78,498}{10^6} = 7.85\%$ vs the estimate $\frac{1}{\ln(10^6)} = \frac{1}{6\ln(10)} = 7.24\%$

Proportion of primes up to 10^{12}: $\frac{37,607,912,018}{10^{12}} = 3.76\%$ vs the estimate $\frac{1}{\ln(10^{12})} = \frac{1}{12\ln(10)} = 3.62\%$

An example of huge relevance for crypto.

By the PNT, the proportion of primes up to 2^{2048} is about $\frac{1}{\ln(2^{2048})} = \frac{1}{2048 \cdot \ln(2)} = 0.0704\%$.

That means, roughly, 1 in 1500 numbers of this magnitude are prime. That means we (i.e. our computer) can efficiently generate large random primes by just repeatedly generating large random numbers and discarding those that are not prime.

Comment. Here, $\ln(x)$ is the logarithm with base e. Isn't it wonderful how Euler's number $e \approx 2.71828$ is sneaking up on the primes?

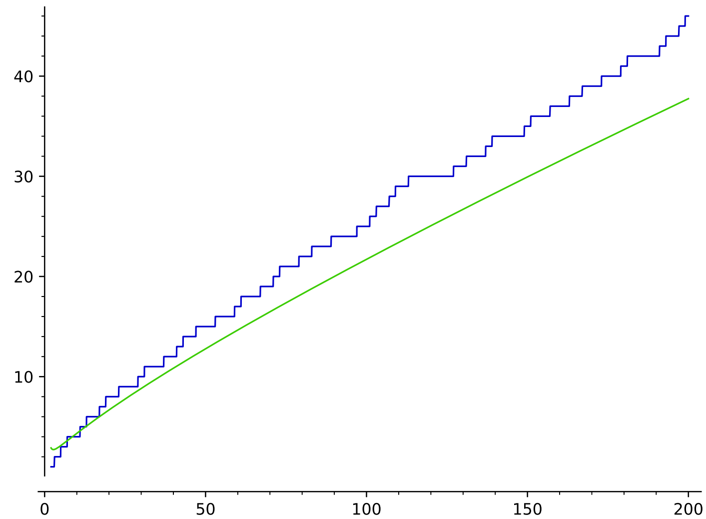
Historical comment. Despite progress by Chebyshev (who succeeded in 1852 in showing that the quotient in the above limit is bounded, for large x, by constants close to 1), the PNT was not proved until 1896 by Hadamard and, independently, de la Vallée Poussin, who both used new ideas due to Riemann.

Example 103. Playing with the prime number theorem in Sage:

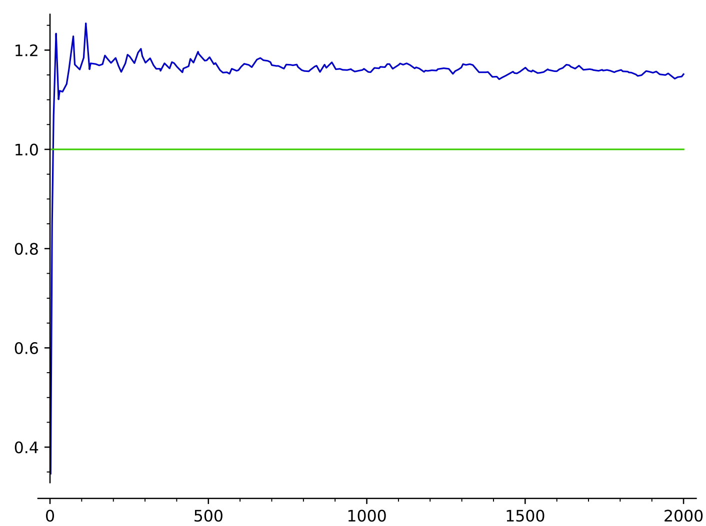
```
Sage] prime_pi(10)
```

4

```
Sage] plot([prime_pi(x), x/ln(x)], 2, 200)
```




```
Sage] plot([prime_pi(x)/(x/ln(x)), 1], 2, 2000)
```



Comment. As the final plot suggests, the quotient of $\pi(x)$ and $x / \ln(x)$ indeed approaches 1 from above. This is slightly stronger than the PNT, which only claims that the quotient approaches 1.

In particular, as the previous plot suggests, for large x, $x / \ln(x)$ is always an underestimate for $\pi(x)$ (though looking at a plot like this can be very misleading).

Armin Straub
straub@southalabama.edu