2.1 The Prime Number Theorem

Let $\pi(x)$ be the number of primes $p \leq x$. Somewhat more generally let $\pi_{a,b}(x)$ be the number of primes $p \leq x$ of the form $p = ak + b$ (in other words: congruent to b modulo a). The prime number theorem states the asymptotic relation ()

$$\pi_{a,b}(x) \sim \frac{1}{\varphi(a)} \cdot \frac{x}{\ln(x)}$$

provided a and b are coprime. The special case $a = 1$, $b = 0$, is:

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

There are many theoretical and empirical results concerning the quality of this approximation. An instance is a formula by Rosser and Schoenfeld:

$$\frac{x}{\ln(x)} \cdot \left(1 + \frac{1}{2\ln(x)}\right) < \pi(x) < \frac{x}{\ln(x)} \cdot \left(1 + \frac{3}{2\ln(x)}\right) \quad \text{for } x \geq 59.$$

The prime number theorem helps for answering the following questions (albeit not completely exactly):

How many prime numbers $< 2^k$ do exist?

Answer: $\pi(2^k)$, that is about

$$\frac{2^k}{k \cdot \ln(2)},$$

at least (for $k \geq 6$)

$$\frac{2^k}{k \cdot \ln(2)} \cdot \left(1 + \frac{1}{2k\ln(2)}\right).$$

For $k = 128$ this number is about $3.8 \cdot 10^{36}$, for $k = 256$, about $6.5 \cdot 10^{74}$.

How many k-bit primes do exist?

Answer: $\pi(2^k) - \pi(2^{k-1})$, that is about

$$\frac{2^k}{k \cdot \ln(2)} - \frac{2^{k-1}}{(k-1) \cdot \ln(2)} = \frac{2^{k-1}}{\ln(2)} \cdot \frac{k-2}{k(k-1)} \approx \frac{1}{2} \cdot \pi(2^k)\,.$$

For $k = 128$ this amounts to about $1.9 \cdot 10^{36}$, for $k = 256$, to about $3.2 \cdot 10^{74}$. In other words, a randomly chosen k-bit integer is prime with probability

$$\frac{\pi(2^k) - \pi(2^{k-1})}{2^{k-1}} \approx \frac{\pi(2^k)}{2^k} \approx \frac{1}{k \cdot \ln(2)} \approx \frac{1.44}{k}\,.$$

For $k = 256$ this is about 0.0056.

The inequality

$$\pi(2^k) - \pi(2^{k-1}) > 0.71867 \cdot \frac{2^k}{k} \qquad \text{for } k \geq 21.$$

gives a reliable lower bound.

In any case the number of primes of size relevant for RSA is huge and makes an exhaustion attack completely obsolete.

Special Primes

Often cryptologists want their primes to have special properties:

Definition A **special prime** (or **safe prime**) is a prime of the form $p = 2p' + 1$ where p' is an odd prime (then p' is also called a GERMAIN prime).

Remark Let p be special. Then $p \equiv 3 \pmod 4$, for $p = 2p' + 1 \equiv 2 \cdot a + 1$ where $a = 1$ or 3.

Definition A **superspecial prime** is a prime of the form $p = 2p' + 1$ where $p' = 2p'' + 1$ is a special prime.

Examples The two smallest superspecial primes are $p = 23$ (with $p' = 11$, $p'' = 5$) and $q = 47$ (with $q' = 23$, $q'' = 11$).

Are there enough primes to fulfill these special or superspecial requests?

Frankly speaking, there is no exact answer. However we can give (unproven!) fairly exact estimates for these numbers:

- As we saw, a (positive) k-bit integer is prime with probability $\frac{\alpha}{k}$ where $\alpha \approx 1.44$.
- If $p = 2p' + 1$ is special, then p' is a $k/2$-bit integer, and is prime (heuristically, but in fact unknown) with probability $\frac{2\alpha}{k}$.
- Thus we estimate that a random k-bit integer is a special prime with probability $\frac{\alpha}{k} \cdot \frac{2\alpha}{k} = \frac{2\alpha^2}{k^2}$, and we expect that $\frac{\alpha^2}{k^2} \cdot 2^k$ of the 2^{k-1} k-bit integers are special primes (assuming that the "events" p prime and $(p-1)/2$ prime are independent).
- Moreover $p'' = (p' - 1)/2$ is a $k/4$-bit integer, hence prime with probability $\frac{4\alpha}{k}$.

- This makes up for a probability of

$$\frac{\alpha}{k} \cdot \frac{2\alpha}{k} \cdot \frac{4\alpha}{k} = \frac{8\alpha^3}{k^3}$$

  for a k-bit integer to be a superspecial prime.

- By this consideration—although we have no mathematical proof for it—we expect that

$$\frac{\alpha^3}{k^3} \cdot 2^{k+2}$$

  of the 2^{k-1} k-bit integers are superspecial primes.

- For $k = 256 = 2^8$ (and $\alpha^2 \approx 2$, $\alpha^3 \approx 3$) we may hope for

$$\begin{array}{rcll} 2 \cdot 2^{256} \cdot 2^{-16} & \approx & 3.5 \cdot 10^{72} & \text{special primes,} \\ 3 \cdot 2^{258} \cdot 2^{-24} & \approx & 8.3 \cdot 10^{70} & \text{superspecial primes.} \end{array}$$

Extensions

Let p_n be the n-th prime, thus $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, ... Let $\vartheta(x)$ be the sum of the logarithms of the primes $\leq x$,

$$\vartheta(x) = \sum_{p \leq x,\, p \text{ prime}} \ln(p).$$

Then we have the asymyptotic formulas

$$\begin{array}{rcl} p_n & \sim & n \cdot \ln(n), \\ \vartheta(x) & \sim & x, \end{array}$$

and the error bounds due to Rosser/Schoenfeld:
(1)

$$n \cdot \left(\ln(n) + \ln\ln(n) - \frac{3}{2}\right) < p_n < n \cdot \left(\ln(n) + \ln\ln(n) - \frac{1}{2}\right) \quad \text{for } n \geq 20,$$

(2)
$$x \cdot \left(1 - \frac{1}{\ln(x)}\right) < \vartheta(x) < x \cdot \left(1 - \frac{1}{2\ln(x)}\right) \quad \text{for } n \geq 41.$$

For a proof of the prime number theorem see any textbook on analytic number theory, for example

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