Patterns in the primes

James Maynard

Mathematical Institute, Oxford

August 6, 2015

Primes: Why care?

Primes are the ‘atoms’ of the integers from the point of view of multiplication.

Primes: Why care?

Primes are the ‘atoms’ of the integers from the point of view of multiplication.

Theorem (Fundamental Theorem of Arithmetic)

Every integer $n > 1$ can be written as a product of primes $n = p_1 \times p_2 \times \cdots \times p_k$. Moreover, this is unique apart from rearranging the product.

Primes: Why care?

Primes are the 'atoms' of the integers from the point of view of multiplication.

Theorem (Fundamental Theorem of Arithmetic)

Every integer $n > 1$ can be written as a product of primes $n = p_1 \times p_2 \times \cdots \times p_k$. Moreover, this is unique apart from rearranging the product.

Example

$6 = 2 \times 3 = 3 \times 2.$
$1024 = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2.$

Primes: Why care?

Primes are the ‘atoms’ of the integers from the point of view of multiplication.

Theorem (Fundamental Theorem of Arithmetic)

Every integer $n > 1$ can be written as a product of primes $n = p_1 \times p_2 \times \cdots \times p_k$. Moreover, this is unique apart from rearranging the product.

Example

$6 = 2 \times 3 = 3 \times 2.$

$1024 = 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2.$

Mathematicians are **lazy**. This means we can simplify a problem about integers to one about primes.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ *with* $n > 2$ *apart from the obvious ones when* $xyz = 0$.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Proof.

1. Imagine $n = ab$ is composite, and there is a solution to $x^n + y^n = z^n$ for n.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Proof.

1. Imagine $n = ab$ is composite, and there is a solution to $x^n + y^n = z^n$ for n.
2. Then $(x^b)^a + (y^b)^a = (z^b)^a$, so (x^b, y^b, z^b) is a solution for a.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Proof.

1. Imagine $n = ab$ is composite, and there is a solution to $x^n + y^n = z^n$ for n.
2. Then $(x^b)^a + (y^b)^a = (z^b)^a$, so (x^b, y^b, z^b) is a solution for a.
3. So there is a solution for every prime factor of n.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Proof.

1. Imagine $n = ab$ is composite, and there is a solution to $x^n + y^n = z^n$ for n.
2. Then $(x^b)^a + (y^b)^a = (z^b)^a$, so (x^b, y^b, z^b) is a solution for a.
3. So there is a solution for every prime factor of n.
4. Contradiction to FLT for primes! Unless $n = 2^k$.

Fermat's last theorem

Theorem (Example: Fermat's last theorem)

There are no integer solutions to $x^n + y^n = z^n$ with $n > 2$ apart from the obvious ones when $xyz = 0$.

Claim

It is enough to prove FLT with n prime

Proof.

1. Imagine $n = ab$ is composite, and there is a solution to $x^n + y^n = z^n$ for n.
2. Then $(x^b)^a + (y^b)^a = (z^b)^a$, so (x^b, y^b, z^b) is a solution for a.
3. So there is a solution for every prime factor of n.
4. Contradiction to FLT for primes! Unless $n = 2^k$.
5. Fermat: There are no solutions for $n = 4$. □

Counting primes

Question

How many primes are there?

Counting primes

Question

How many primes are there?

Theorem

There are infinitely many primes.

Counting primes

Question

How many primes are there?

Theorem

There are infinitely many primes.

Gauss: "Around x the primes should occur with density $1/\ln x$".

Theorem (Prime Number Theorem)

$$\pi(x) = \#\{primes \ \leq x\} \approx \int_2^x \frac{dt}{\ln t}.$$

Counting primes

Question

How many primes are there?

Theorem

There are infinitely many primes.

Gauss: ”Around x the primes should occur with density $1/\ln x$”.

Theorem (Prime Number Theorem)

$$\pi(x) = \#\{primes\ \le x\} \approx \int_2^x \frac{dt}{\ln t}.$$

$$\pi(10^{10}) = 455,052,511, \qquad \int_2^{10^{10}} \frac{dt}{\ln t} = 455,055,613.8...$$

Difference 3102.8... (< 0.0007%).

My favourite formula

Instead of counting primes with weight 1, it is easier to compensate for the density by counting with weight $\ln p$.

My favourite formula

Instead of counting primes with weight 1, it is easier to compensate for the density by counting with weight $\ln p$.

Theorem (Riemann's explicit formula)

If x is not an integer, then

$$\sum_{\substack{n,p \\ p^n < x}} \ln p = x - \sum_{\substack{\rho \\ \zeta(\rho)=0}} \frac{x^\rho}{\rho} - \ln(2\pi).$$

Here $\zeta(s)$ is the 'Riemann zeta function', and the sum is over zeros of the zeta function.

My favourite formula

Instead of counting primes with weight 1, it is easier to compensate for the density by counting with weight $\ln p$.

Theorem (Riemann's explicit formula)

If x is not an integer, then

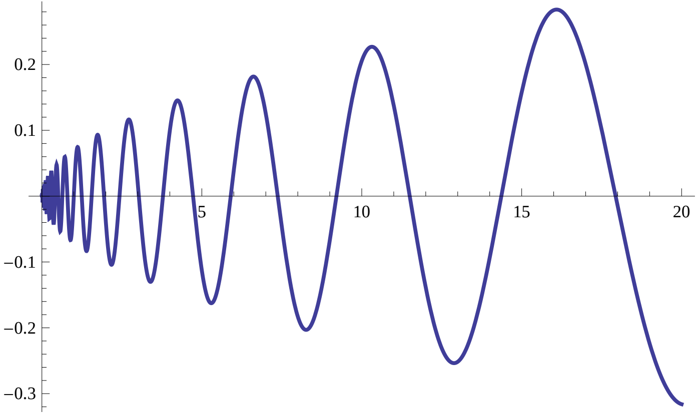
$$\sum_{\substack{n,p \\ p^n<x}} \ln p = x - \sum_{\substack{\rho \\ \zeta(\rho)=0}} \frac{x^\rho}{\rho} - \ln(2\pi).$$

Here $\zeta(s)$ is the 'Riemann zeta function', and the sum is over zeros of the zeta function.

Riemann's formula

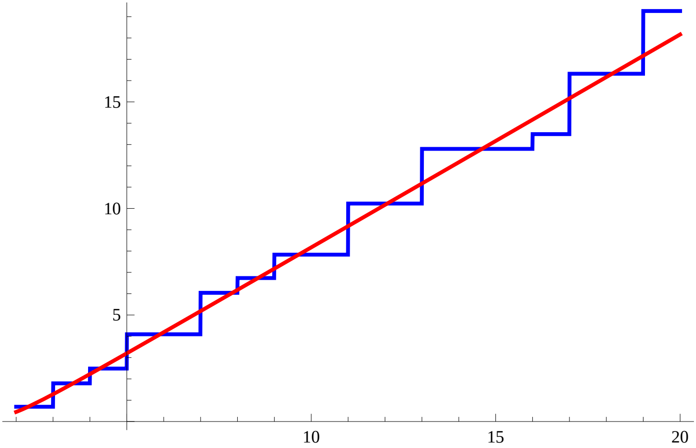
Therefore the zeros tell us exactly where the primes are!

Our step function is a sum of 'waves':

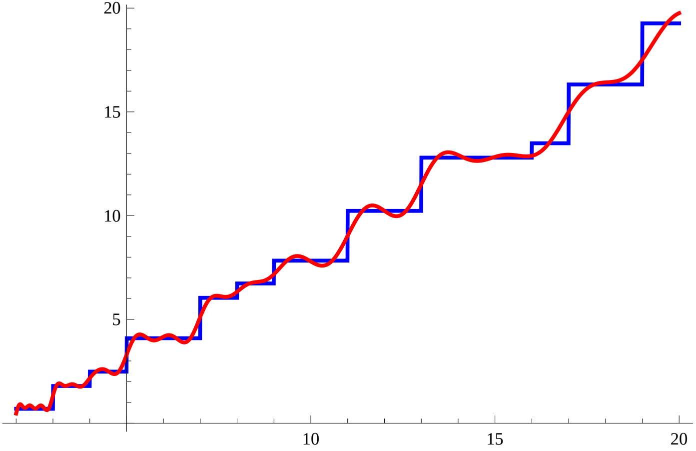


'The music of the primes'

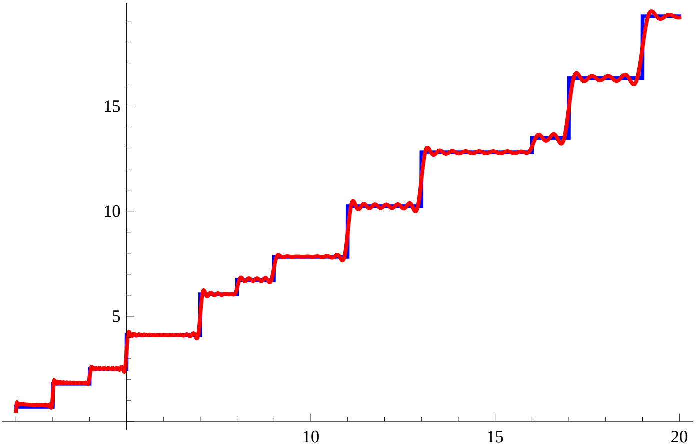
No Zeros



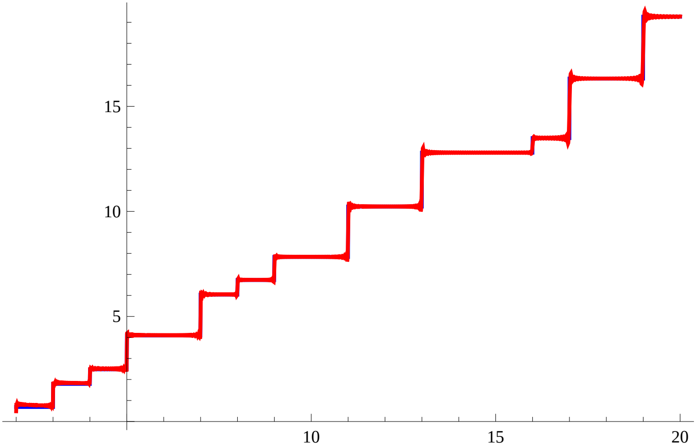
10 Zeros



100 Zeros



1000 Zeros



Riemann's Hypothesis

The size of x^{ρ} is $x^{\Re(\rho)}$.

Conjecture (Riemann's Hypothesis, $1, 000, 000)

All the non-trivial zeros of $\zeta(s)$ have real part 1/2.

Riemann's Hypothesis

The size of x^ρ is $x^{\Re(\rho)}$.

Conjecture (Riemann's Hypothesis, \$1,000,000)

All the non-trivial zeros of $\zeta(s)$ have real part 1/2.

This means all the terms x^ρ have size $\sqrt{x}$, which is much smaller than x.

Corollary

Assume RH. Then for all $x > 2$

$$\left|\pi(x) - \int_2^x \frac{dt}{\ln t}\right| < 4\sqrt{x}\ln x$$

This would completely explain why $\int_2^x dt/\ln t$ is such a good approximation! This explains the large-scale structure!

It isn't just me who's excited

(Hilbert)

"If I were to awaken after having slept for a thousand years, my first question would be: Has the Riemann hypothesis been proven?"

(Montgomery)

"So if you could be the Devil and offer a mathematician to sell his soul for the proof of one theorem - what theorem would most mathematicians ask for? I think it would be the Riemann Hypothesis."

Small-scale distribution

What about primes on a small scale - the gaps between them?

Small-scale distribution

What about primes on a small scale - the gaps between them?

Theorem (Prime Number Theorem)

$$\#\{primes \leq x\} \approx \int_2^x \frac{dt}{\ln t} \approx \frac{x}{\ln x}.$$

Corollary

The average size gap $p_{n+1} - p_n$ amongst primes $p_n \leq x$ is $\approx \ln x$

Small-scale distribution

What about primes on a small scale - the gaps between them?

Theorem (Prime Number Theorem)

$$\#\{primes \leq x\} \approx \int_2^x \frac{dt}{\ln t} \approx \frac{x}{\ln x}.$$

Corollary

The average size gap $p_{n+1} - p_n$ amongst primes $p_n \leq x$ is $\approx \ln x$

Proof.

$$\begin{aligned} \text{Average gap} &= \frac{\sum_{p_n<x}(p_{n+1} - p_n)}{\#\{p_n \leq x\}} \\ &= \frac{p_N - 2}{\pi(x)} \approx \frac{x}{x/\ln x} \\ &\approx \ln x. \end{aligned}$$

□

Small gaps between primes

Question

Are prime gaps always this big?

- $(2, 3)$ is the only pair of primes which differ by 1.

Small gaps between primes

Question

Are prime gaps always this big?

- $(2,3)$ is the only pair of primes which differ by 1.
  (One of n and $n+1$ is a multiple of 2 for every integer n).

Small gaps between primes

Question

Are prime gaps always this big?

- $(2, 3)$ is the only pair of primes which differ by 1.
  (One of n and $n + 1$ is a multiple of 2 for every integer n).
- There are lots of pairs of primes which differ by 2:
  $(3, 5)$, $(5, 7)$, $(11, 13)$, $\dots$, $(1031, 1033)$, $\dots$,
  $(1000037, 1000039)$, $\dots$, $(1000000007, 1000000009)$, $\dots$

Question

Are prime gaps always this big?

- $(2, 3)$ is the only pair of primes which differ by 1.
  (One of n and $n + 1$ is a multiple of 2 for every integer n).
- There are lots of pairs of primes which differ by 2:
  $(3, 5)$, $(5, 7)$, $(11, 13)$, $\dots$, $(1031, 1033)$, $\dots$,
  $(1000037, 1000039)$, $\dots$, $(1000000007, 1000000009)$, $\dots$

Conjecture (Twin prime conjecture)

There are infinitely many pairs of primes (p, p') which differ by 2.

This is one of the oldest problems in mathematics, and is very much open!

How many Twin primes are there?

- If we randomly picked a number n of size x, then the probability n is prime is about $1/\ln x$.

How many Twin primes are there?

- If we randomly picked a number n of size x, then the probability n is prime is about $1/\ln x$.
- If we randomly picked a number n of size x, then the probability $n+2$ is prime is about $1/\ln x$.

How many Twin primes are there?

- If we randomly picked a number n of size x, then the probability n is prime is about $1/\ln x$.
- If we randomly picked a number n of size x, then the probability $n+2$ is prime is about $1/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $1/(\ln x)^2$.

How many Twin primes are there?

- If we randomly picked a number n of size x, then the probability n is prime is about $1/\ln x$.
- If we randomly picked a number n of size x, then the probability $n+2$ is prime is about $1/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $1/(\ln x)^2$.

Guess

$$\#\{\textit{twin primes } \leq x\} \approx \int_2^x \frac{dt}{(\ln t)^2}$$

How many Twin primes are there?

- If we randomly picked a number n of size x, then the probability n is prime is about $1/\ln x$.
- If we randomly picked a number n of size x, then the probability $n+2$ is prime is about $1/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $1/(\ln x)^2$.

Guess

$$\#\{\textit{twin primes} \ \le x\} \approx \int_2^x \frac{dt}{(\ln t)^2}$$

But this can't be right, as n and $n+1$ can't both be prime!

$$\#\{\text{twin primes} \ \le 10^8\} = 440312, \qquad \int_2^{10^8} \frac{dt}{(\ln t)^2} = 333530.2...$$

Difference 106781.8... (about **24.2**%)

Second Attempt

Lets use the fact primes > 2 are odd.

- If we randomly picked an odd number n of size x, then the probability n is prime is about $2/\ln x$.

Second Attempt

Lets use the fact primes > 2 are odd.

- If we randomly picked an odd number n of size x, then the probability n is prime is about $2/\ln x$.
- If we randomly picked an odd number n of size x, then the probability $n + 2$ is prime is about $2/\ln x$.

Second Attempt

Lets use the fact primes > 2 are odd.

- If we randomly picked an odd number n of size x, then the probability n is prime is about $2/\ln x$.
- If we randomly picked an odd number n of size x, then the probability $n+2$ is prime is about $2/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $4/(\ln x)^2$.

Second Attempt

Lets use the fact primes > 2 are odd.

- If we randomly picked an odd number n of size x, then the probability n is prime is about $2/\ln x$.
- If we randomly picked an odd number n of size x, then the probability $n+2$ is prime is about $2/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $4/(\ln x)^2$.

Guess (Second attempt)

$$\#\{twin\ primes\ \leq x\} \approx 2\int_2^x \frac{dt}{(\ln t)^2}$$

Worse!

Third Attempt

If n and $n + 2$ are prime, n must be 2 more than a multiple of 3, and so 1 less than a multiple of 6.

- If we randomly picked a number n of the form $6k - 1$ of size x, then the probability n is prime is about $3/\ln x$.

Third Attempt

If n and $n+2$ are prime, n must be 2 more than a multiple of 3, and so 1 less than a multiple of 6.

- If we randomly picked a number n of the form $6k-1$ of size x, then the probability n is prime is about $3/\ln x$.
- The probability $n+2$ is prime is also about $3/\ln x$.

Third Attempt

If n and $n+2$ are prime, n must be 2 more than a multiple of 3, and so 1 less than a multiple of 6.

- If we randomly picked a number n of the form $6k-1$ of size x, then the probability n is prime is about $3/\ln x$.
- The probability $n+2$ is prime is also about $3/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $9/(\ln x)^2$.

Third Attempt

If n and $n+2$ are prime, n must be 2 more than a multiple of 3, and so 1 less than a multiple of 6.

- If we randomly picked a number n of the form $6k-1$ of size x, then the probability n is prime is about $3/\ln x$.
- The probability $n+2$ is prime is also about $3/\ln x$.
- If these were independent events, then the probability n and $n+2$ are both prime would be $9/(\ln x)^2$.

Guess (Third attempt)

$$\#\{\textit{twin primes} \leq x\} \approx \frac{3}{2}\int_2^x \frac{dt}{(\ln t)^2}$$

Error $\approx$ **13%**. Better!

Lets try to correct for **all** primes $p > 2$.

- There are $p - 2$ possible remainders for n after dividing by p if n and $n + 2$ are prime.

Lets try to correct for **all** primes $p > 2$.

- There are $p - 2$ possible remainders for n after dividing by p if n and $n + 2$ are prime.
- So the probability than neither n nor $n + 2$ are a multiple of p is $(p - 2)/p$.

Lets try to correct for **all** primes $p > 2$.

- There are $p-2$ possible remainders for n after dividing by p if n and $n+2$ are prime.
- So the probability than neither n nor $n+2$ are a multiple of p is $(p-2)/p$.
- If n and $n+2$ were 'independent', then the probability neither were a multiple of p is $(p-1)/p \times (p-1)/p$.

Lets try to correct for **all** primes $p > 2$.

- There are $p-2$ possible remainders for n after dividing by p if n and $n+2$ are prime.
- So the probability than neither n nor $n+2$ are a multiple of p is $(p-2)/p$.
- If n and $n+2$ were 'independent', then the probability neither were a multiple of p is $(p-1)/p \times (p-1)/p$.
- So we were off by a factor $\frac{p(p-2)}{(p-1)^2}$.

Guess (Fourth attempt)

$$\#\{\textit{twin primes } \leq x\} \approx 2C_2 \int_2^x \frac{dt}{(\ln t)^2}$$

with $C_2 = \prod_{p>2} p(p-2)/(p-1)^2$.

$\#\{\text{twin primes } \leq 10^8\} = 440312, \qquad 2C_2 \int_2^{10^8} \frac{dt}{(\ln t)^2} = 440367.8...$

Difference 55.8... (this is $<$ **0.2**%). Success!

Other patterns

We can look at more than just gaps of size 2.

Conjecture (De Polignac)

For every positive integer h, there are infinitely many pairs of primes which differ by $2h$.

Again, we guess the number less than x is roughly $C_h x/(\ln x)^2$ for some constant C_h.

Other patterns

We can look at more than just gaps of size 2.

Conjecture (De Polignac)

For every positive integer h, there are infinitely many pairs of primes which differ by $2h$.

Again, we guess the number less than x is roughly $C_h x/(\ln x)^2$ for some constant C_h.

Theorem

This is true for at least one h in $\{1, \ldots, 123\}$.

Other patterns

We can look at more than just gaps of size 2.

Conjecture (De Polignac)

For every positive integer h, there are infinitely many pairs of primes which differ by $2h$.

Again, we guess the number less than x is roughly $C_h x/(\ln x)^2$ for some constant C_h.

Theorem

This is true for at least one h in $\{1, \ldots, 123\}$.

In particular:

Theorem

There are infinitely many pairs (p_1, p_2) of primes such that $|p_1 - p_2| \leq 246$.

Other patterns II

We can also look for **triples** of primes $n, n+h_1, n+h_2$ for some fixed shifts h_1, h_2

1. If $h_1 = 2$, $h_2 = 4$ then $(3, 5, 7)$ is the only triple. (one of $n, n+2, n+4$ must be a multiple of 3)
2. If $h_1 = 2$, $h_2 = 6$ then there are many such triples.

Conjecture

There are infinitely many n such that n, $n+h_1, \dots, n+h_k$ are prime if there isn't an obvious reason why they can't be.

'Obvious reason' means one is always a multiple of some prime for all n.

Theorem

There exists $h_1, \dots, h_k$ such that n, $n+h_1, \dots, n+h_k$ are all primes for infinitely many n.

Optimistic extensions

If we assume a well-believed technical conjecture about primes in arithmetic progressions, then we can get close to the twin prime conjecture!

Theorem

Assume 'GEH'. Then there are infinitely many pairs (p_1, p_2) *of primes with* $|p_1 - p_2| \leq 6$.

Goldbach's conjecture

This conjecture also allows us to say something about another old conjecture

Conjecture (Goldbach's conjecture)

Every even number can be written as the sum of at most two primes.

Theorem

Assume 'GEH'. Then ***at least*** *one of the following is true:*

1. *There are infinitely many twin primes*
2. *For every large even integer N, one of N, $N+2$ or $N-2$ is the sum of two primes.*

Of course we expect both to be true!

Real-World Example

1. Alice wants to send Bob a Facebook message containing sensitive gossip.

Real-World Example

1. Alice wants to send Bob a Facebook message containing sensitive gossip.
2. This can be done securely if her laptop can find $N = pq$ which is hard to factor into primes.

Real-World Example

1. Alice wants to send Bob a Facebook message containing sensitive gossip.
2. This can be done securely if her laptop can find $N = pq$ which is hard to factor into primes.
3. If $p - 1$ has only small prime factors, then there is a way to factor N easily (Bad).

Real-World Example

1. Alice wants to send Bob a Facebook message containing sensitive gossip.
2. This can be done securely if her laptop can find $N = pq$ which is hard to factor into primes.
3. If $p - 1$ has only small prime factors, then there is a way to factor N easily (Bad).
4. On Wikipedia it had been suggested that one could choose p, q such that $(p - 1)/2$ and $(q - 1)/2$ are prime.

Real-World Example

1. Alice wants to send Bob a Facebook message containing sensitive gossip.
2. This can be done securely if her laptop can find $N = pq$ which is hard to factor into primes.
3. If $p - 1$ has only small prime factors, then there is a way to factor N easily (Bad).
4. On Wikipedia it had been suggested that one could choose p, q such that $(p - 1)/2$ and $(q - 1)/2$ are prime.
5. If there are only 10 (say) 1024-digit primes p such that $(p - 1)/2$ is prime, then this is a VERY bad idea! Bob would die before Alice finds one!

A slight generalization of our model predicts there are many such primes.

Long path to go

It is an exciting time for prime number theory!

Any questions?