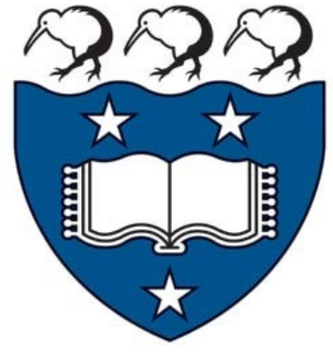
Prime numbers, factorisation, and algorithms

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ISA Bologna,
Sept 26, 2023

Andre Nies,
The University of Auckland




Munich
Slovakia
Košice
Bratislava
Budapest
Zurich
Austria
Vaduz
Hungary
Cluj-Napoca
Switzerland
Geneva
Lyon
Slovenia
Venice
Zagreb
Milan
Turin
Croatia
Genoa
Bologna
Bosnia and Herzegovina
Serbia
Marseille
Monaco
Italy
Sarajevo
Montenegro
Kosovo
Rome
Macedonia (FYROM)
Bari
Tirana
Naples
Albania
Tyrrhenian Sea
Greece
Athens
Αθήνα
Palermo
Catania
Algiers
الجزائر
Tunis
تونس
Malta
Sfax
صفاقس
Mediterranean Sea
Tunisia
Tripoli

Auckland

View of Auckland from Maungawhau/ Mt Eden

What is a prime number?

The list of prime numbers is 2,3,5,7,11,13,17,19,…

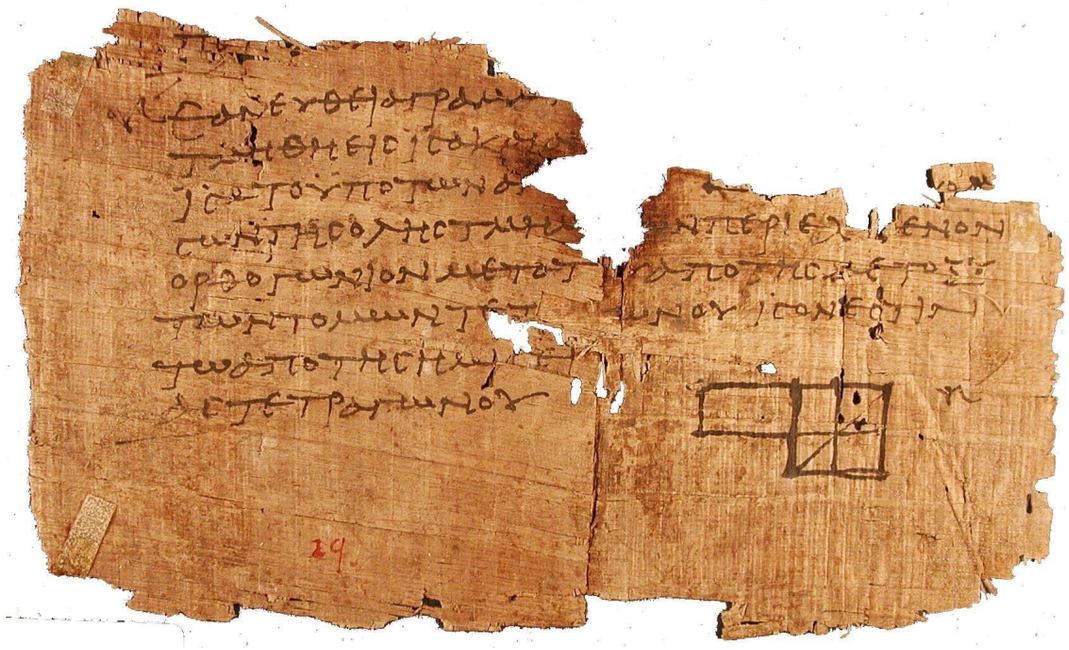
Definition.
A whole number n with n >1 is called prime if it cannot be written as a product of two smaller whole numbers.

For instance,

- 13 is prime
- 15 is not, because $15 = 3 \cdot 5$.

Triskaidekaphobia:
fear of the number 13
Primonumerophobia:
fear of prime numbers

Euclid’s Elements (ca 300 BC)



A piece of the Elements, found among the [Oxyrhynchus papyri]

Translation into Arabic, 1274

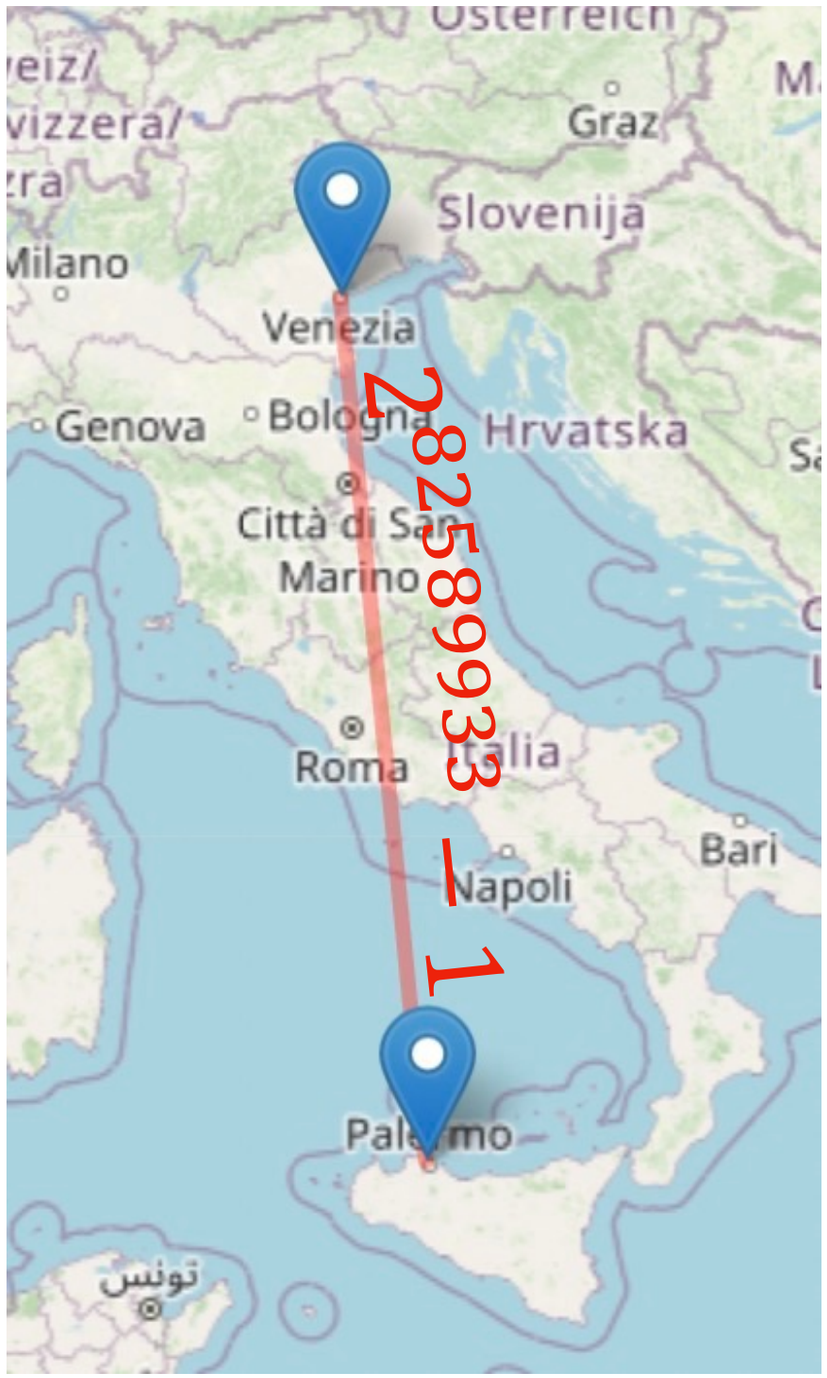
The list of primes never ends

Euclid (ca. 300 BC) proved a result in his geometric language, which amounts to saying that there are infinitely many prime numbers.

The proof is by contradiction:

- If there are only finitely many, form their product; call it N.
- Let k be the least number >1 that divides N+1.
- k is prime because it was chosen least.
- k can’t be any of the known primes, for they all divide N. Contradiction!

The largest known prime number



The largest known prime number is

$$2^{82589933} - 1$$

Written in decimal in 10pt font (three digits per centimetre), it is about 800km long!

Written in binary, this is a sequence of 1s about 2400 km long:

1111111111111 . . . 11111111111111

Mersenne primes

Padre Marin Mersenne, 1588-1639

Prime numbers of the form $2^p - 1$ are called Mersenne primes: 3,7, 31, 127, … It is unknown whether there are infinitely many!

Padre Mersenne already knew:

if $2^p - 1$ is prime then p must be prime.

E.g., 15= $2^4 - 1$ and 63= $2^6 - 1$ aren't prime.

The converse fails: for instance,
2047= $2^{11} - 1$ is not prime even though 11 is.

The primeness of $2^{82589933} - 1$ was established by GIMPS (Great Internet Mersenne Prime Search) in January 2019, using the Lucas-Lehmer test.

Why are primes so interesting?

Theoretical reasons:

- Very simple concept leading to super-deep conjectures and results
- they are the building blocks of numbers: every number is a unique product of primes (up to order).

Practical reasons:

- Use in cryptography, in particular RSA public key system
- Playground for inventing fast algorithms, also quantum.

The rest of the talk will follow this outline.

Some old conjectures, and progress on them

Prime twin conjecture: there are infinitely many primes p such that p+2 is also a prime. Such as 101 and 103, both are prime.

Progress on this (Maynard, Tao, Wang, polymath, 2019): there are infinitely many primes p such there is another prime strictly between $p \; and \; p + 247$.

Goldbach conjecture (1742): Every even number $n \geq 4$ is the sum of two prime numbers. For instance, 100= 47+53.

Progress (Helfgott, arXiv 2014): Each odd number $n \geq 9$ is the sum of three odd prime numbers. The Goldbach conjecture would imply this because then $n - 3$ is sum of two primes.

Arithmetic progressions in the primes

Ben Green and Terrence Tao (Annals of Mathematics, 2008) studied primes using methods of structure versus randomness.
In this way they showed that the primes contain arbitrarily long arithmetic progressions. Here is an example of such a progression:

5,11,17,23,29.

That is, starting from 5, for four times one can add 6 and obtain a new prime.

The longest known arithmetical progression has length 27. It was found in 2019 by Rob Gahan and PrimeGrid, and starts with 224,584,605,939,537,911.

Writing a number as a product of primes

Prime numbers are the “building blocks” of natural numbers.

Fundamental theorem of Arithmetic (Euclid)

Every number $n \geq 2$ can be written as a product of primes, which is unique (disregarding the order).

Examples:

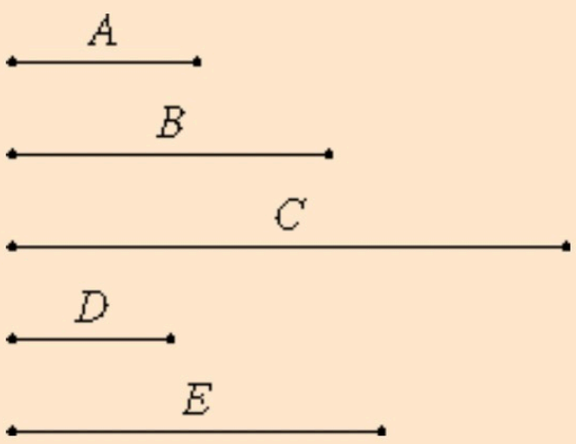
$$999 = 3 \cdot 3 \cdot 3 \cdot 37$$
$$1000 = 2 \cdot 2 \cdot 2 \cdot 5 \cdot 5 \cdot 5,$$
$$1001 = 7 \cdot 11 \cdot 13$$

For the uniqueness, Euclid proved in Book VII, Prop. 30: if a prime divides a product of two numbers, it divides one of the two numbers.

If two numbers, multiplied by one another make some number, and any prime number measures the product, then it also measures one of the original numbers.

Let the two numbers A and B multiplied by one another make C, and let any prime number D measure C.

I say that D measures one of the numbers A or B.



Let it not measure A.

Now D is prime, therefore A and D are relatively prime. VII.29

Let as many units be in E as the times that D measures C.

Since then D measures C according to the units in E, therefore D multiplied by E makes C. VII.Def.15

Further, A multiplied by B also makes C, therefore the product of D and E equals the product of A and B.

Therefore D is to A as B is to E. VII.19

But D and A are relatively prime, relatively prime numbers are also least, and the least measure the numbers which have the same ratio the same number of times, the greater the greater and the less the less, that is, the antecedent the antecedent and the consequent the consequent, therefore D measures B. VII.21 VII.20

Similarly we can also show that, if D does not measure B, then it measures A. Therefore D measures one of the numbers A or B.

Therefore, *if two numbers, multiplied by one another make some number, and any prime number measures the product, then it also measures one of the original numbers.*

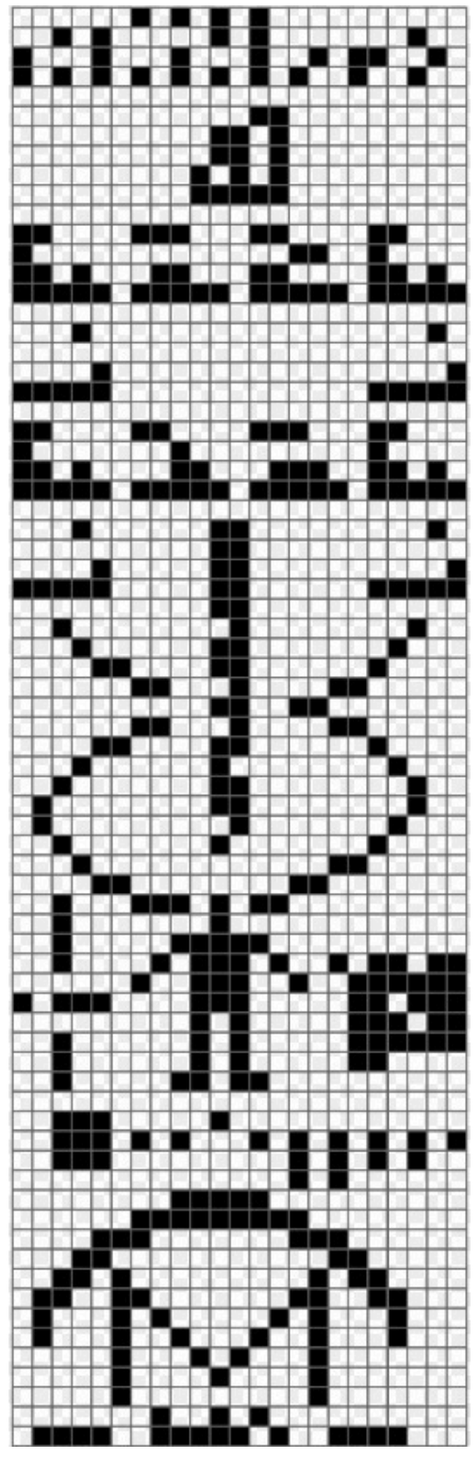
Q.E.D.

The 1974 Arecibo message

Humanity sent a radio message containing a sequence of 1679 bits to Messier 13, a globular star cluster 22000 light years away.

It started 00000101010101000…

Since the factorization of 1679 is $23 \cdot 73 = 1679$, the aliens living in M13 will be able to turn the sequence of bits into this picture:



Algorithms?

- Is there an algorithm to recognise whether a number N is prime?
- Is there an algorithm to find the factoring of N into prime numbers?

For instance, consider $N = 4321$.
Can you (not your phone) carry out these tasks for N?

Answer: 4321 is not prime. In fact,

$$4321 = 29 \cdot 149$$

RSA factoring challenges

During 1991-2007, RSA labs offered cash prizes for factoring particular numbers (that were not primes). RSA offered $75000 for factoring the following 270-digit number. Its prime factors are still unknown.

```
412023436986659543855531365332575948179811699
844327982845455626433876445565248426198098870
423161841879261420247188869492560931776375033
421130982397485150944909106910269861031862704
114880866970564902903653658674337317208131O4
105190864254793282601391257624033946373269391
```

Algorithmic questions (recall)

- Is there an algorithm to recognise whether a number N is prime?

- Is there an algorithm to find the factoring of N into prime numbers?

It may come a surprise that the first question has an affirmative answer, while the second is open!

The questions ask for a feasible algorithm that comes up with an answer. (Not one that takes 1000s of years.)

Algorithms for primeness

- The ``trial division" algorithm attempts to divide N by $2,3,5,7,11, \ldots$, all the way up to $\sqrt{N}$.

- If N is not prime then some factor has to be less than $\sqrt{N}$, so this certainly finds out if N is prime.

- But it is not feasible. If N has 200 decimal digits, in the worst case one needs to try about 10^{98} potential divisors!

- There is no known alternative if one really wants to try out divisors.

- Instead, feasible algorithms are based on number theoretic properties that only the primes have. They can answer ``No, not prime" without ever giving a proper divisor!

Fermat pseudo-primality test

The idea is based on Fermat's little theorem:

If p is a prime and $1 < a < p$, then $a^{p-1} mod\ p = 1$ (that is, a^{p-1} when divided by p leaves remainder 1). E.g., $3^{5-1} = 81$.

We want to know if p is prime. Repeat the following sufficiently often:

Guess a number a with $1 < a < p$.

Test if $a^{p-1} mod\ p = 1$. If so, say YES, otherwise NO.

With enough independent repeats, we can

push the probability of error as low as we want.

https://www.omnicalculator.com/math/power-modulo

Miller's test, and Rabin's probabilistic algorithm

Fermat's test doesn't quite work: there are infinitely many "fake" primes (Carmichael numbers, e.g., $561 = 3 \cdot 11 \cdot 17$) for which each reasonable witness a yields a YES.

Miller (1976) found a modification of the main procedure that did work. Assuming the generalised Riemann Hypothesis, he showed that trying the candidates a from 2 up to $O((\log p)^2)$ is sufficient.

Rabin (1980) proved that, in Miller's modification, with a logarithmic number of trials one can push the error (false positives) as low as one wants! It could be lower than the chance of a hardware error.

AKS algorithm (2002): absolute answer, no error.

One can efficiently decide primeness (sort-of)

- The Rabin algorithm works well in practice, but is not satisfying from a theoretical point of view, because it has a arbitrarily small error of getting a false positive (i.e., saying that p is prime when it really isn't).

- The Agrawal-Kayal-Saxena (AKS) algorithm satisfies the definition from complexity theory of being efficient: it runs in ``deterministic polynomial time". But the polynomial bound on the running time isn't good in practice: about $O(n^6)$.

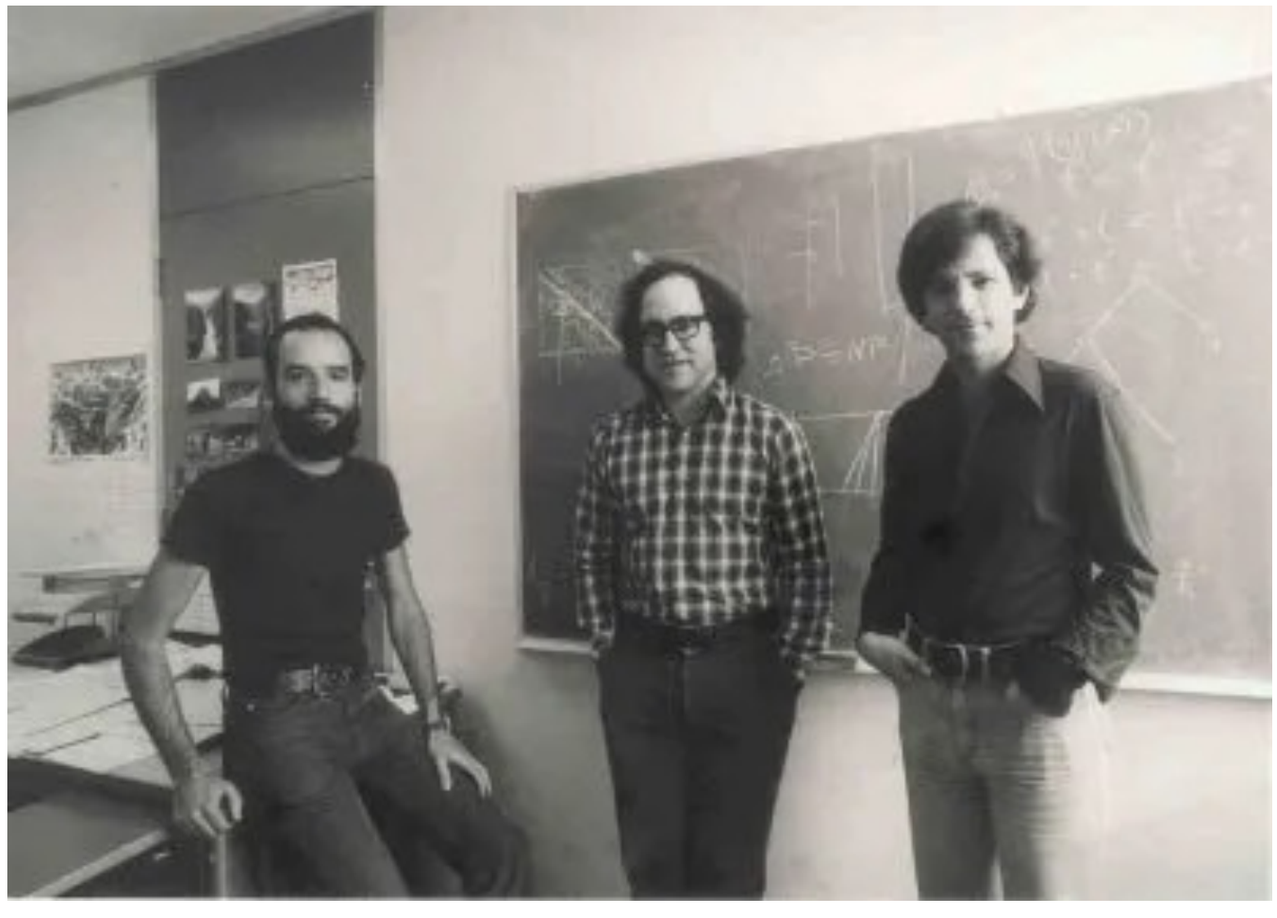
- The problem with a monster like $2^{82589933} - 1$ is that it is 800km long, so even the efficient general algorithm fails.

- Lucas-Lehmer works though, for the number is Mersenne.

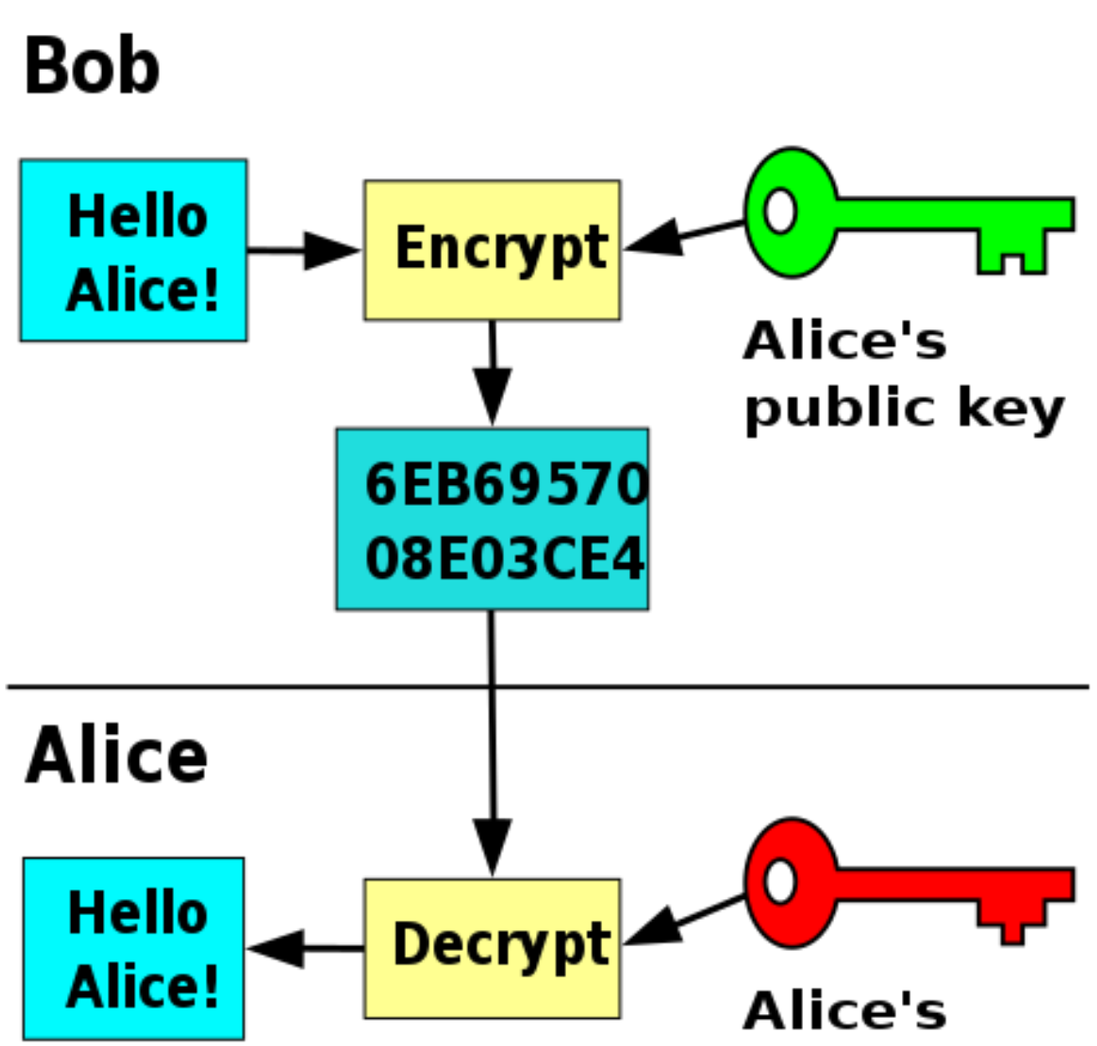
Is there an efficient algorithm for factoring?

If you are worried about the security of your data and credit card transactions, you might want to answer: Hopefully not.

Why? The hardness of factoring is assumed for making RSA encryption work. RSA is short for Rivest-Shamir-Adelman, the three MIT scientists who came up with the method in 1977.



Public key cryptosystem (Merkle, Diffie-Hellman)



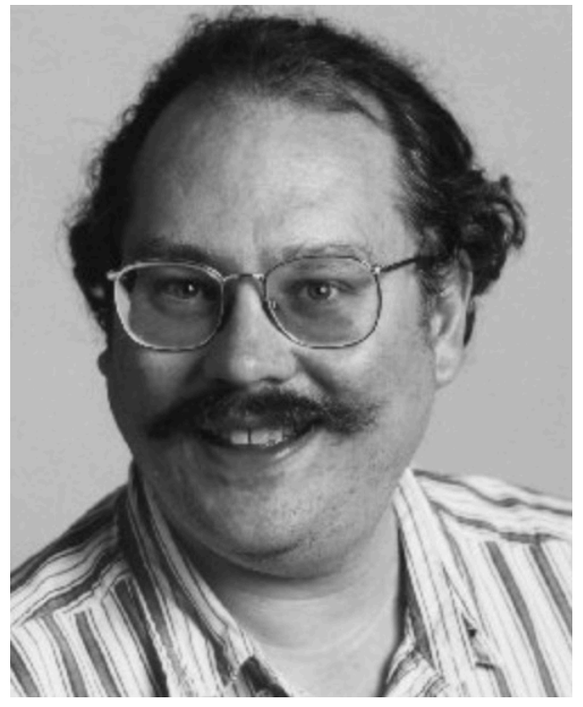
Source: wikipedia

In the RSA system, Alice starts with $N = pq$ where p, q are large primes. She uses this to obtain the two keys.

She releases N as part of the public key, and hides her private key.

Since it is hard to factor N, an adversary cannot obtain the private key from the public one and the encrypted message.

Shor's algorithm (1994)



Peter Shor, MIT
1994 plenary ICM speaker,
1998 Nevanlinna prize

Problem: Given a number N that is not prime, find a nontrivial factorization $N = ab$.

Shor's algorithm can do that in polynomial time on a hypothetical "quantum computer".

This means that one needs circuits of poly(log N) many quantum gates. The algorithm only finds the result with high probability.

Classical part of Shor: reduce factoring to order-finding

Let N be odd, not prime, not a prime power. E.g. N= 15.

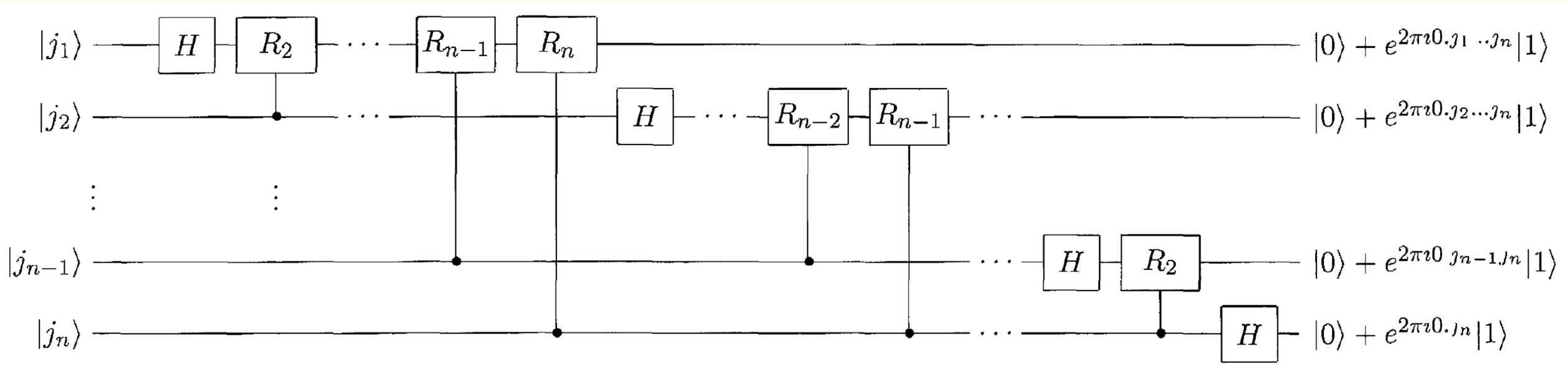
- Choose random $x < N$ such that gcd(x, n)=1. E.g. x=7
- Let p be the order of $x \bmod N$. I.e., p is least such that $x^p \equiv 1 \bmod N$. If p is odd, try other x. p=4
- Else $\left(x^{p/2}+1\right)\left(x^{p/2}-1\right) \equiv x^p - 1 \equiv 0 \bmod N$.
- So N divides this product.
- So gcd($x^{p/2}+1, N$) gcd($x^{p/2}-1, N$)= N is a factoring.

E.g. $(7^2+1)(7^2-1) \equiv 7^4 - 1 \equiv 0 \bmod 15$, and the factoring is $5 \cdot 3 = 15$

If one of the factors is 1, try another x.

One can show that the chance of x being "useless" is $\leq 2^{-m}$, where N has m prime factors.

Circuit for n-bit quantum Fourier transform



Here $R_k \equiv \begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix}$ and the given number is $j_1 \ldots j_n$ in binary.

Credit: Nielsen-Chuang, Quantum Computation and Quantum Information, 2010 edition, page 218

So, is RSA encryption still safe?

- So far, the largest number factored with Shor's algorithm is 21. (Some team has tried 35.)
- The problem is the noise/error when applying quantum gates. A paper by Y. Cai uploaded on arXiv in June 2023 claims to prove that the error is unavoidable, and so Shor cannot in practice factor numbers.
- Other algorithms work better: Schnorr's SVP. In Dec 2022, researchers (arxiv.org/2212.12372) have used this to factor a 48-bit number on a quantum device:

  261980999226229 = 15538213x16860433.

- The current RSA standard uses 2048 bits (617 digits)
- New field: post-quantum cryptography