



PRIMI AL NOVANTA PER CENTO...

*Carlo Toffalori, Camerino
Scuola Estiva, Castellammare di Stabia
Luglio 2018*



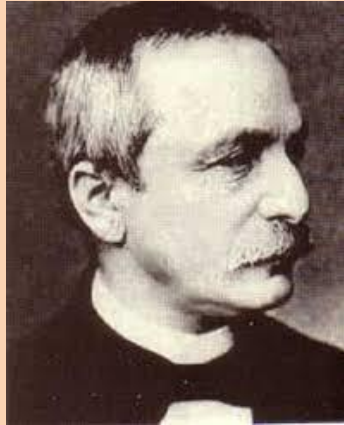
L'aritmetica

- l'insieme $\mathbf{N}$ dei numeri naturali 0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, ...
- le operazioni di addizione, moltiplicazione, *sottrazione* e *divisione*

Dante, *Convivio*

“L'arismetrica... del suo lume tutte si illuminano le scienze... l'occhio de lo 'ntelletto nol può mirare...”

Due citazioni (libere) più moderne



- L. Kronecker: *“i numeri interi sono i soli creati da Dio”*



- H. Weyl (o A. Weil?): *“i teoremi di incompletezza di Gödel dimostrano l’esistenza di Dio e anche quella del demonio”*

Le difficoltà della divisione: un numero naturale $N > 1$ è

- *primo* se gli unici divisori di N sono 1 e N ,
- *composto* altrimenti

I numeri primi: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, ...

2	3	5	7	11	13	17	19	23	29	31	37
41	43	47	53	59	61	67	71	73	79	83	89
97	101	103	107	109	113	127	131	137	139	149	151
157	163	167	173	179	181	191	193	197	199	211	223
227	229	233	239	241	251	257	263	269	271	277	281
283	293	307	311	313	317	331	337	347	349	353	359
367	373	379	383	389	397	401	409	419	421	431	433
439	443	449	457	461	463	467	479	487	491	499	503
509	521	523	541	547	557	563	569	571	577	587	593
599	601	607	613	617	619	631	641	643	647	653	659
661	673	677	683	691	701	709	719	727	733	739	743
751	757	761	769	773	787	797	809	811	821	823	827
829	839	853	857	859	863	877	881	883	887	907	911
919	929	937	941	947	953	967	971	977	983	991	997

numeri primi

Teorema fondamentale dell'aritmetica. Ogni numero naturale $N > 1$ si scompone in uno ed un solo modo (a meno dell'ordine dei fattori) nel prodotto di numeri primi.

$$15 = 3 \cdot 5, \quad 18 = 2 \cdot 3^2, \quad 24 = 2^3 \cdot 3, \quad 32 = 2^5, \dots$$

La moltitudine dei numeri primi...



Quale è il primo più grande?



Se non altro, 2 è l'unico primo pari (*"the oddest prime"*, Zassenhaus) e ci sono infiniti numeri primi

La successione dei primi è irregolare e imprevedibile

I primi gemelli: con differenza 2, dunque p e $p + 2$

- esempi: 3 e 5, 5 e 7, 11 e 13, 17 e 19, ...
- controesempi: 7 e 11, 13 e 17, 19 e 23, 23 e 29, ...

Non è chiaro quante siano le coppie di primi gemelli (infinite?).

Eppure, ***per ogni intero positivo s si possono trovare s numeri consecutivi nessuno dei quali è primo.*** Basta considerare

- $(s + 1)! + 2 > 2$, divisibile per 2, dunque composto,
- $(s + 1)! + 3 > 3$, divisibile per 3, dunque composto,
-
- $(s + 1)! + s + 1 > s + 1$, divisibile per $s + 1$, dunque composto.

Ricordare: il teorema dei numeri primi, e molto altro...

La congettura di Goldbach, 1742

haben, nicht bestanden, ob mehrere aber schon nachgeprüft sind,
 * man sieht series lauter numeros viro modo in duo quadrata
 divisibiles ganz auf solche Weise will ich auf eine conjecture
 hazardiren: daß jede Zahl welche aus zweyten numeros primis
 zusammengesetzt ist in aggregatum sovieler numerorum
 primorum sey als man will: in unitatem mit seyn übersetzend
 heißt das Congeriem omnium unitatum. z. B. 4 = 1+1+1+1
 6 = 1+1+1+1+1+1

Binneß folgen nur paar observationes so demonstrirte unter
 den Leuten:

Si v. sit functio ipsius x. cuiusmodi ut facta v = c. numero cuiusque,
 determinari possit x per c. et reliquas constantes in functione
 expressas, poterit etiam determinari valor ipsius x. in de-
 quatione $v^{n+1} = (2v+1)(v+1)$. $\left| \frac{v^{n+1}}{2v+1} - \frac{(v+1)(v+1)^{n-1}}{2v+1} \right|$ d. h. $v^{n+1} - (v+1)^n$

Si incipiatur curva cuius abscissa sit x. applicata vero sit
 summa seriei $\frac{x^n}{n \cdot 2^{2n}}$ posita n. pro exponente terminorum, hoc est
 applicata = $\frac{x}{1 \cdot 2} + \frac{x^2}{2 \cdot 2^2} + \frac{x^3}{3 \cdot 2^3} + \frac{x^4}{4 \cdot 2^4} + \text{etc.}$ dico, si fuerit
 abscissa = 1. applicata fore = $\frac{1}{2} = \frac{1}{2}$: sed hoc apparet = 4
 2. - - - - - 2. 12.
 3. - - - - - 2. 12.
 4. vel major. - - - infinitam.

Ich erwähne mit aller ungeschmälzter Beifügung
 dieses Beweisesbogens
 Moskau 7. Jun. st. 1742. J. Goldbach.

Vertical text on the left margin:
 Diese Conjectur ist schon öfters von andern Mathematikern gemacht worden, aber noch nicht bewiesen worden. Ich habe sie hier nur als eine Vermuthung aufgeführt, die man weiter untersuchen kann.

- C. Goldbach, lettera a Eulero: *ogni numero naturale $N \geq 6$ è la somma di 3 primi*
- L. Eulero, risposta a Goldbach: equivalente provare che C. Goldbach, ***ogni numero pari $M \geq 4$ è la somma di 2 primi***

La congettura di Goldbach:

- basta un controesempio per smentirla,
 - non bastano miliardi di miliardi di miliardi esempi per confermarla
- $4 = 2 + 2, 6 = 3 + 3, 8 = 3 + 5, 10 = 3 + 7, 12 = 5 + 7, 14 = 3 + 11, \dots$
- $2p = p + p$ per ogni primo p





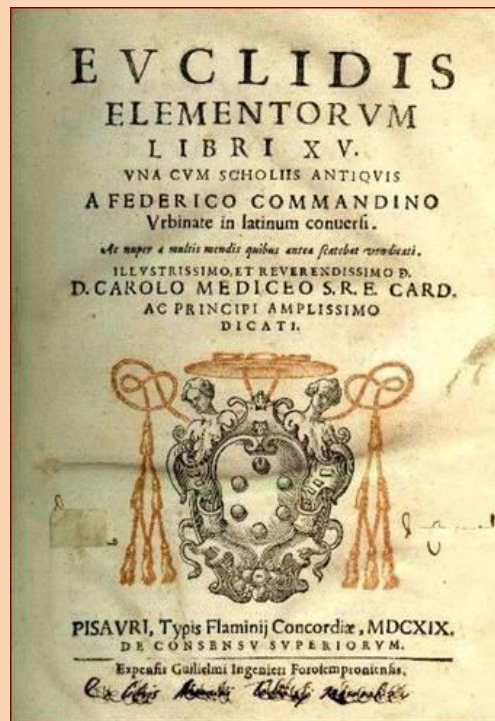
Due problemi classicissimi: primalità e fattorizzazione

- *Un input comune: un naturale $N \geq 2$*
- *L'output atteso*

Primalità: N è primo o composto?

Fattorizzazione: la decomposizione di N in fattori primi (il primo passo: per N composto dispari, un divisore proprio d di N diverso da 1 e N).

Lecito assumere N dispari, i numeri pari si trattano facilmente.



Un algoritmo elementare per entrambi i casi (noto già agli antichi Greci)

- Si prende $1 < d < N$ e si divide N per d .
- Se la divisione è esatta per qualche d , si deduce che N è composto (e anzi si trova un suo divisore d diverso da 1 e N)
- Se la divisione non è esatta per nessun d , si deduce che N è primo.



Il commento di Gauss, *Disquisitiones Arithmeticae*, 1801 (articolo 329)

“Il problema di separare i primi dai composti e di decomporre i secondi nei loro fattori primi è conosciuto essere uno dei più importanti e utili in Matematica. La dignità stessa della scienza sembra richiedere di esplorare ogni possibile mezzo per la soluzione di un problema così elegante e famoso ”

Perchè questa esigenza due millenni dopo i Greci?

Gauss: *“Le tecniche conosciute in precedenza richiederebbero una fatica intollerabile anche per i più instancabili **calcolatori**. ”*



Il costo dell'algoritmo "elementare"

- $N - 2$ divisioni, una per ogni possibile $d = 2, \dots, N - 1$,
- da riferire non a N , ma alla **lunghezza** di N

Ma la lunghezza di N in base 10 (o 2) è all'incirca il suo logaritmo in quella base!

- Un miliardo 1.000.000.000 si compone di 10 cifre
- 100 è piccolo, ma 2^{100} supera abbondantemente l'età del mondo in secondi



Nel caso specifico:

- ogni singola divisione ha costo quadratico rispetto alla lunghezza di N ,
- le divisioni richieste nei casi peggiori sono quasi a 10 elevati alla lunghezza di N .

Possibili scorciatoie

- Esplorare $1 < d < \sqrt{N}$: se $N = d \cdot d'$ con $d, d' > 1$, uno tra d e d' è $\leq \sqrt{N}$.
- Se d non funziona, neppure $2d, 3d, \dots$ vanno bene.

Ma i guai restano! Sono prevedibili fino a $\sqrt{N} - 1$ divisioni, ossia una quantità ancora esponenziale rispetto alla lunghezza di N .



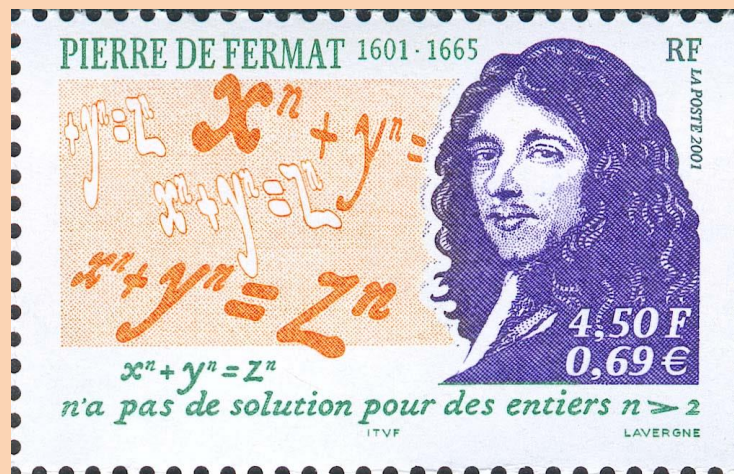
Scindere primalità e fattorizzazione

- ***Teorema di Wilson:*** per ogni intero $N > 1$, N è primo se e solo se $(N-1)! \equiv -1 \pmod{N}$.
- ***Commenti***

Si controlla la *primalità* di N senza bisogno di considerarne la *fattorizzazione*.

La verifica richiede comunque il calcolo di $(N-1)!$, dunque $N-2$ moltiplicazioni (“troppe” rispetto alla lunghezza di N).

A proposito di congruenze



Il piccolo teorema di Fermat: se N è primo, allora ogni intero a primo con N soddisfa $a^{N-1} \equiv 1 \pmod{N}$.

Una dimostrazione rapida:

- si considera il gruppo moltiplicativo del campo $\mathbb{Z}_N$ che ha ordine $N - 1$ e si compone degli interi modulo N primi con N , e dunque invertibili modulo N ;
- si applica il teorema di Lagrange.



Da notare: in riferimento alla lunghezza di N , è “rapido” controllare

- a, N primi tra loro (l’algoritmo di Euclide richiede tempi al più quadratici),
- $a^{N-1} \equiv 1 \pmod{N}$ (basta una divisione, dunque un tempo quadratico).

Pure l’elevamento a potenza $a \rightarrow a^N - 1 \pmod{N}$ ammette procedure veloci di calcolo.

Curiosità numero 1. Sia $N > 1$ un intero dispari. Fissiamo a primo con N , $1 < a < N$, ad esempio $a = 2$ (visto che N è dispari). Se vale $a^{N-1} \equiv 1 \pmod{N}$, allora N è primo?

Da ricordare: la congruenza $a^{N-1} \equiv 1 \pmod{N}$ è “rapida” da verificare.

La risposta: NO, ci sono numeri N composti che soddisfano $a^{N-1} \equiv 1 \pmod{N}$ (pseudoprimi in base a)

Un esempio (Sarrus): $341 = 11 \cdot 31$ è pseudoprimo in base 2 (anche se non in base 3).



Curiosità numero 2. Se vale $a^{N-1} \equiv 1 \pmod{N}$ per **ogni** intero a primo con N , $1 < a < N$, allora N è primo?

Da notare: le congruenze da controllare possono arrivare, per N primo, a $N - 1$ (troppe rispetto alla lunghezza di N).

La risposta: NO, ci sono numero composti dispari N che soddisfano $a^{N-1} \equiv 1 \pmod{N}$ per ogni intero a primo con N , $1 < a < N$: gli *pseudoprimi di Carmichael* (1912)

Esempi: 561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341, 41041, 46657, 52633, 62745, 63973, ... (**infiniti!** Alford-Granville-Pomerance, 1994)





Altri algoritmi di primalità basati sul piccolo teorema di Fermat?

Carl Pomerance: *“la congruenza di Fermat è così semplice che è un peccato doverci rinunciare”*

Un'altra premessa utile: le radici quadrate modulo un primo N

Per N primo dispari, ogni quadrato modulo N primo con N ha esattamente 2 radici quadrate modulo N :

se $x^2 \equiv a^2 \pmod{N}$, allora $x \equiv \pm a \pmod{N}$, e inoltre $a \not\equiv -a \pmod{N}$.

Infatti un primo N che divide $x^2 - a^2 = (x - a) \cdot (x + a)$ divide $x - a$ o $x + a$. Inoltre N non può dividere $a - (-a) = 2a$ perché dispari e primo con a .

In particolare le radici quadrate di 1 modulo N sono ± 1 .

Non sempre vero per N composto! Per esempio $x^2 \equiv 1 \pmod{8}$

- per $x \equiv \pm 1 \pmod{8}$
- ma anche per $x \equiv \pm 3 \pmod{8}$.

*Un approccio sorprendente e discusso: gli **algoritmi probabilistici** (le classi BPP , RP , $coRP$, ZPP della teoria della complessità computazionale)*

*L'idea: **sacrificare la precisione per aumentare la velocità***

- Si ricorre all'aiuto di testimoni casuali
- Si ammettono risposte sbagliate o incerte purché i tempi di lavoro siano accelerati
- Si cerca di rendere minima la probabilità di errore o di insicurezza (al variare dei testimoni)



Émile Borel: *un evento che ha probabilità $< 10^{-50}$ non accadrà mai, e se anche accade non sarà mai rilevato.*



Algoritmi probabilistici

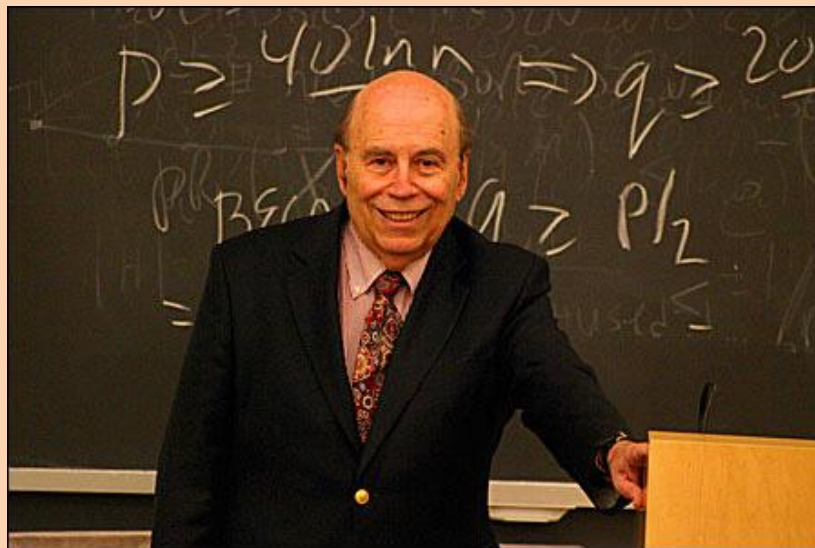
- *tipo Montecarlo*: risposte probabilmente vere in tempi certamente rapidi
- *tipo Las Vegas*: risposte certamente vere in tempi probabilmente rapidi

In altre parole

- nel primo caso è ammessa la possibilità che un testimone sbagli,
- nel secondo caso è ammessa la possibilità che un testimone dica “*non lo so*”,

ed entrambe sono da rendere minime!

Un esempio di algoritmo Montecarlo per i primi: l'algoritmo di Miller-Rabin, 1978, fallibile ma efficiente!



I prerequisiti (quelli già ricordati): se N è primo (dispari),

- le radici quadrate di 1 modulo N sono ± 1 ,
- (il piccolo teorema di Fermat) per ogni intero a primo con N , vale $a^N - 1 \equiv 1 \pmod{N}$.



Miller



Rabin

Il procedimento

È dato l'intero **dispari** $N > 2$. Dunque $N-1$ è pari. Si interpella un testimone intero a con $1 < a < N$.

Passo 1. Si calcola il massimo comune divisore d di a e N .

- Se $d \neq 1$, l'algoritmo risponde (correttamente) “ **N COMPOSTO**” (ha trovato un divisore non banale di N , appunto d)
- Se $d = 1$, si procede con i passi successivi.



Passo 2. Si calcola a^{N-1} modulo N .

- Se $a^{N-1} \not\equiv 1 \pmod{N}$, l'algoritmo risponde (correttamente) “***N* COMPOSTO**” (è smentito il piccolo teorema di Fermat).
- Se invece $a^{N-1} \equiv 1 \pmod{N}$, si va al passo successivo.

Passo 3. Si calcola $a^{\frac{N-1}{2}}$ modulo N (la radice quadrata di a^{N-1} modulo N).

- Se $a^{\frac{N-1}{2}} \not\equiv \pm 1 \pmod{N}$, l'algoritmo risponde (correttamente) “***N COMPOSTO***”.
- Se $a^{\frac{N-1}{2}} \equiv -1 \pmod{N}$, l'algoritmo azzarda “***N PRIMO***”.
- Idem se $a^{\frac{N-1}{2}} \equiv 1 \pmod{N}$ ma $\frac{N-1}{2}$ non è pari, cioè non si può dimezzare ulteriormente.
- Altrimenti, se $a^{\frac{N-1}{2}} \equiv -1 \pmod{N}$ e $\frac{N-1}{2}$ è pari, cioè esiste $\frac{N-1}{4}$ tra gli interi, si va ai passi successivi.

Passo 4. Come il passo 3, ma riferito ad $a^{\frac{N-1}{4}}$ modulo N (la radice quadrata di $a^{\frac{N-1}{2}}$ modulo N).

Si prosegue finché possibile.



Commenti

1. ***Attendibilità***: si noti che

- la risposta “***N COMPOSTO***” è sicura.
- quella “***N PRIMO***” può essere sbagliata.

Probabilità di errore dopo la consultazione di un testimone a : al variare di a , al più $\frac{1}{4}$

In altre parole: almeno 3 testimoni su 4 dicono la verità!

*D'altra parte, dopo la consultazione di 100 testimoni a che rispondono concordemente “***N PRIMO***”, la probabilità di errore diventa*

$$\frac{1}{4^{100}} < \frac{1}{10^{50}} .$$

2. **Tempi di lavoro:** sono da calcolare (**finché possibile**)

- dapprima un massimo comune divisore,
- poi per ogni passo un elevamento a potenza e una divisione.

Complessivamente

- I tempi limitati da un polinomio di grado al più 5 rispetto alla lunghezza di N (incluse le prevedibili iterazioni),
- la ripetizione per al più 100 testimoni non li aggrava irreparabilmente (per N grande).

Esempi

1) Sia $N = 13$, dunque $N - 1 = 12$. Scegliamo $a = 2$. Allora

- 2 è primo con 13,
- $2^{12} = 4096 \equiv 1 \pmod{13}$,
- $2^6 = 64 \equiv -1 \pmod{13}$.

Dunque l'algoritmo dichiara " **N PRIMO**" e in effetti N lo è.

2) Sia $N = 7$, così $N - 1 = 6$. Riferiamoci sempre ad $a = 2$. Allora

- 2 è primo con 7,
- $2^6 = 64 \equiv 1 \pmod{7}$,
- $2^3 = 8 \equiv 1 \pmod{7}$

dopo di che non è più possibile dividere l'esponente 3 per 2. Dunque l'algoritmo dichiara "***N PRIMO***" e in effetti N lo è.

3) Sia adesso $N = 561$ (uno pseudoprimo di Carmichael – in altre parole, per ogni intero b primo con N , $b^{560} \equiv 1 \pmod{561}$). Si sceglie $a = 359$ e con un po' di pazienza si controlla

- 359 è primo con 561,
- $359^{560} \equiv 1 \pmod{561}$,
- $359^{280} \equiv 1 \pmod{561}$,
- $359^{140} \equiv 1 \pmod{561}$,
- $359^{70} \equiv 1 \pmod{561}$,
- $359^{35} \equiv 1 \pmod{561}$.

Siccome 35 non si può dividere ulteriormente per due, l'algoritmo, se ricorre ad $a = 359$, dichiara “***N PRIMO***” mentre N è composto.

Ma se il testimone 359 non è attendibile, $a = 2$ lo è. Infatti

- 2 è primo con 561,
- $2^{560} \equiv 1 \pmod{561}$,
- $2^{280} \equiv 1 \pmod{561}$,
- $2^{140} \not\equiv \pm 1 \pmod{561}$,

Quindi il ricorso ad $a = 2$ dichiara correttamente “***N COMPOSTO***”.

Altri algoritmi Montecarlo per i primi: l'algoritmo di Solovay-Strassen, 1977



Solovay



Strassen

Si basa sul simbolo di Legendre-Jacobi e su un teorema di Eulero: ma l'algoritmo di Miller-Rabin è preferibile sia per affidabilità che per efficienza

*Un algoritmo rapido e **infallibile**: l'algoritmo AKS (Agrawal, Kayal, Saxena), 2002*



- Agrawal: informatico teorico, con interessi alla complessità computazionale
- Kayal, Saxena: matematici, dottorandi di Agrawal.

Agosto 2002: l'algoritmo AKS

- distingue se un dato intero $N > 1$ è primo o composto
- nelle migliori implementazioni lavora in tempo poco più che polinomiale di grado 6 rispetto alla lunghezza di N
- soddisfa finalmente le richieste di Gauss (due secoli dopo la loro formulazione).

*La base lontana: ancora il **piccolo teorema di Fermat**.*



Conseguenze di AKS: l'algoritmo di Bernstein-Berrizbeitia, 2004

- probabilistico di tipo Montecarlo
- lavora in tempo poco più che polinomiale di grado 4 rispetto alla lunghezza di N
- assolutamente affidabile quando risponde “***N PRIMO***”
- talora fallibile (ma con bassa probabilità di errore) quando dichiara “***N COMPOSTO***”



Da usare in combinazione con l'algoritmo di Miller-Rabin, per ottenere un procedimento di tipo Las Vegas. Per ogni input N , si eseguono in parallelo su N i due algoritmi, ricorrendo a opportuni testimoni.

- Se l'algoritmo di Bernstein-Berrizbeitia dichiara “ **N PRIMO**”, si conclude che N è primo.
- Se l'algoritmo di Miller-Rabin dichiara “ **N COMPOSTO**”, si conclude N composto.

Il caso disgraziato (rarissimo)

- L'algoritmo di Bernstein-Berrizbeitia dichiara “ **N COMPOSTO**”,
- quello di Miller-Rabin “ **N PRIMO**”.

La conclusione obbligata: “ **$NON LO SO$** ”, necessario interpellare altri testimoni.

E il problema della fattorizzazione?

Versione “ridotta”

Input: $N > 1$ composto e dispari

Output: un divisore d di N , $1 < d < N$.

In realtà la questione chiave!

Il guaio: i tempi di lavoro degli algoritmi oggi conosciuti per la versione “ridotta” sono esponenziali (o quasi) rispetto alla lunghezza dell'input N .

Gran spreco di strumenti teorici

- curve ellittiche,
- frazioni continue, ...

nessuno capace di lavorare in tempo al più polinomiale nella lunghezza di N .

Ma attenzione alla computazione quantistica (P. Shor, 1994)!





Alcune famose decomposizioni in fattori primi

- Eulero, 1742: $F(5) = 641 \cdot 6700417$
- Clausen, 1856: $F(6) = 274177 \cdot 67280421310721$
- Landry, 1869: $M(59) = 179951 \cdot 3203431780337$
- Cole, 1903: $M(67) = 193707721 \cdot 761838257287$ (convegno dell'American Mathematical Society a S. Francisco, “*tre anni di domeniche*”)
- Poulet, 1923: $M(73) = 439 \cdot 2298041 \cdot 9361973132609$ (ma il fattore 439 era già stato scoperto da Eulero).

Solo speculazioni teoriche? Applicazioni alla sicurezza della rete

Il crittosistema a chiave pubblica più popolare: RSA (Rivest, Shamir, Adleman), 1978



Si fonda proprio su

- la facilità di generare e riconoscere i primi,
- la difficoltà di fattorizzare i composti.