L'importanza di essere primo

Alessandro Zaccagnini*

Dedicato alla memoria di Franco Conti

1 Introduzione

I numeri primi sono spesso considerati enti "astratti" di interesse esclusivamente matematico. Scopo di questo saggio è la discussione di alcune proprietà dei numeri primi che hanno trovato importantissime applicazioni pratiche con l'avvento della crittografia a chiave pubblica negli anni '70 e con la diffusione di Internet e del commercio elettronico di questi ultimi anni.

Ci occuperemo di alcune proprietà peculiari dei numeri primi che sono "profonde" senza per questo essere particolarmente difficili, e senza che siano richieste complesse nozioni preliminari: infatti, l'unica cosa di cui abbiamo bisogno è il concetto di congruenza, che è strettamente connesso alla divisibilità, ed è quindi, a tutti gli effetti, una nozione elementare. Come applicazione concreta delle idee qui esposte, abbiamo incluso la descrizione di un popolare sistema di crittografia a chiave pubblica (ElGamal).

Per non intralciare il discorso, i fatti principali relativi alle congruenze sono raccolti nell'Appendice A, insieme alle dimostrazioni formali dei risultati piú importanti. Nell'Appendice B descriviamo un algoritmo di scomposizione in fattori (diverso dalla divisione ripetuta) che sfrutta in modo essenziale l'idea di congruenza. Infine, diamo una certa quantità di riferimenti bibliografici, in maggioranza in lingua italiana, per eventuali approfondimenti, non essendo ovviamente possibile esaurire in questa sede tutti gli argomenti interessanti.

*Dipartimento di Matematica, Università di Parma, Via M. d'Azeglio, 85/a, 43100 Parma.
e-mail: alessandro.zaccagnini@unipr.it
url: http://www.math.unipr.it/∼zaccagni/home.html

2 Le congruenze e i numeri primi

Il Teorema di Wilson. Nel XVIII secolo John Wilson dette una condizione necessaria e sufficiente per la primalità di n molto semplice da enunciare.

Teorema 2.1 (Wilson) *L'intero $n \geq 2$ è primo se e solo se n divide $(n-1)!+1$.*

Possiamo verificare direttamente che questo è vero per valori piccoli di n: per esempio $6!+1 = 721$ ed in effetti 7 divide 721, mentre $7!+1 = 5041$ non è divisibile per 8. Si osservi che già per $n = 11$, il numero $(n-1)!+1 = 3628801$ è piuttosto grande. Per la dimostrazione è necessario il concetto di congruenza, che è brevemente spiegato nell'Appendice A: qui ricordiamo solamente che la scrittura $a \equiv b \bmod n$ significa che l'intero $a-b$ è divisibile per l'intero n. Illustriamo questo concetto vedendo la dimostrazione della prima implicazione del Teorema di Wilson nel caso $p = 17$, osservando che, come suggerito dal Lemma A.2, si ha $2\cdot 9 \equiv 1$, $3\cdot 6 \equiv 1$, $4\cdot 13 \equiv 1$, $5\cdot 7 \equiv 1$, $8\cdot 15 \equiv 1$, $10\cdot 12 \equiv 1$, $11\cdot 14 \equiv 1$, dove tutte le congruenze sono modulo 17. Possiamo quindi disporre il calcolo di 16! cosí:

$$\begin{aligned} 16! &= 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10\cdot 11\cdot 12\cdot 13\cdot 14\cdot 15\cdot 16 \\ &= 1\cdot (2\cdot 9)\cdot (3\cdot 6)\cdot (4\cdot 13)\cdot (5\cdot 7)\cdot (8\cdot 15)\cdot (10\cdot 12)\cdot (11\cdot 14)\cdot 16. \end{aligned} \tag{1}$$

Prendiamo questa uguaglianza modulo 17: dato che i prodotti fra parentesi valgono tutti 1 mod 17, si ha $16! \equiv 1\cdot 1^7 \cdot 16 \equiv 16 \bmod 17$. Dunque $16!+1 \equiv 16+1 \equiv 0 \bmod 17$. È evidente che stiamo sfruttando il fatto che gli interi 2, 3, ..., 15, hanno un inverso modulo 17 diverso da sé stessi, cioè è possibile associare ciascuno di questi interi al suo inverso, esaurendo completamente la lista nella (1), esclusi il primo e l'ultimo elemento, ed usando ciascun fattore una ed una sola volta.

Il problema principale del criterio di Wilson è che il numero delle operazioni necessarie per la verifica è dell'ordine di n: quindi, tentare di determinare la primalità o meno di un intero in questo modo richiederebbe un numero di operazioni ben superiore alla ricerca del piú piccolo fattore primo di n mediante divisioni successive per 2, 3, ... Infatti, se n non è un numero primo, allora ha almeno un fattore che non supera la sua radice quadrata (se a e b sono entrambi maggiori di $\sqrt{n}$, allora il loro prodotto ab è maggiore di n): procedendo per tentativi, è sufficiente considerare gli interi fra 2 e la radice quadrata di n, estremi inclusi. Se n non è divisibile per nessuno di questi interi, allora è certamente un numero primo.

Il Teorema di Fermat. Nel XVII secolo, Pierre de Fermat (noto al grande pubblico piú che altro per il suo cosiddetto "Ultimo Teorema," che è stato dimostrato solo nel 1995) scoprí un'altra interessante relazione che coinvolge numeri primi.

Teorema 2.2 (Fermat) *Se p è un numero primo ed a è un intero non divisibile per p, allora $a^{p-1} \equiv 1 \bmod p$. In altre parole, p divide $a^{p-1}-1$.*

Apparentemente, anche in questo caso il numero delle operazioni da svolgere è dell'ordine di grandezza del numero di cui vogliamo sapere se sia primo o meno, dato che dobbiamo calcolare la $(p-1)$-esima potenza di a modulo p: in effetti, però, le cose non stanno cosí. Per esempio, per calcolare 2^{100} sono sufficienti 8 moltiplicazioni; basta disporre il calcolo come segue:

$$2^{100} = \left(\left(2\cdot\left(((2\cdot 2^2)^2)^2\right)^2\right)^2\right)^2.$$

Quindi, se vogliamo calcolare $2^{100} \bmod 101$ possiamo procedere cosí:

$$\begin{aligned}
2^2 \bmod 101 &\equiv 2^2 = 4\\
2^3 \bmod 101 &\equiv 2\cdot 2^2 = 8\\
2^6 \bmod 101 &\equiv \left(2^3\right)^2 = 64\\
2^{12} \bmod 101 &\equiv \left(2^6\right)^2 = 4096 \equiv 56 \bmod 101\\
2^{25} \bmod 101 &\equiv 2\cdot\left(2^{12}\right)^2 = 2\cdot 3136 \equiv 10 \bmod 101\\
2^{50} \bmod 101 &\equiv \left(2^{25}\right)^2 \equiv 100 \bmod 101\\
2^{100} \bmod 101 &\equiv \left(2^{50}\right)^2 \equiv 10000 \equiv 1 \bmod 101.
\end{aligned} \tag{2}$$

Il procedimento è noto come "algoritmo dei quadrati ripetuti." Questo calcolo suggerisce (ma non dimostra, come spieghiamo qui sotto) che 101 possa essere un numero primo. Osserviamo anche che il calcolo risulta leggermente semplificato (e i numeri piú piccoli) se, ogni volta che troviamo un risultato parziale > 50, da questo sottraiamo 101: infatti questa operazione non fa cambiare la classe di congruenza del risultato, ma è piú semplice e piú efficiente fare calcoli con numeri piccoli piuttosto che con numeri grandi. Con questo accorgimento, i risultati parziali nella (2) sarebbero rispettivamente 4, 8, -37, -45, 10, -1, 1.

In generale, è possibile calcolare la n-esima potenza di un intero eseguendo un numero di moltiplicazioni che non supera il doppio del numero di cifre binarie di n, cioè circa 7 volte il numero delle cifre decimali di n. Se n è un intero molto grande, questo numero è molto piú piccolo di quello delle moltiplicazioni necessarie per verificare il criterio di Wilson.

C'è però un problema: la condizione di Fermat è solamente *necessaria* per la primalità di un intero, e non è certo una condizione *sufficiente*. Probabilmente l'esempio piú semplice possibile è il seguente: per $a = 4$ ed $n = 15$ si ha $4^2 = 16 \equiv 1 \bmod 15$ e quindi $4^{14} = (4^2)^7 \equiv 1 \bmod 15$, ma, evidentemente, 15 *non è* un numero primo. Le coppie di interi (a,n), dove $a > 1$ ed n non è primo, e tali che $a^{n-1} \equiv 1 \bmod n$ sono relativamente rare, ed è effettivamente possibile modificare l'enunciato qui sopra in modo da ottenere una condizione necessaria e sufficiente, ma complicata, per la primalità: purtroppo, però, per ogni a esistono infiniti valori *non primi* di n per cui $a^{n-1} \equiv 1 \bmod n$, detti *pseudoprimi in base a*. Per i dettagli rimandiamo i lettori ai riferimenti contenuti nell'Appendice C.

Il Teorema di Gauss. Il risultato astratto piú importante di cui parleremo qui è un Teorema di Gauss: ne vedremo anche un'applicazione alla crittografia. Ricordiamo che se n è un intero qualsiasi $\mathbb{Z}_n$ indica l'insieme $\{0, 1, 2, \ldots, n-1\}$, e che se p è primo $\mathbb{Z}_p^*$ indica l'insieme $\{1, 2, \ldots, p-1\}$ (si veda l'Appendice A).

Teorema 2.3 (Gauss) *Se p è un numero primo allora esiste almeno un intero $g \in \mathbb{Z}_p^*$ le cui potenze successive g, g^2, $\ldots$, g^{p-1} sono tutte distinte modulo p.*

Per esempio, per i primi piú piccoli possiamo scegliere i valori qui indicati.

p	2	3	5	7	11	13	17	19	23	29	31	37
g_p	1	2	2	3	2	2	3	2	5	2	3	2

Possiamo formulare il Teorema di Gauss in un modo alternativo: scelto comunque un numero primo p, esiste un intero $g \in \mathbb{Z}_p^*$ tale che l'applicazione $x \mapsto g^x \bmod p$ è una biiezione (o corrispondenza biunivoca) fra $\mathbb{Z}_{p-1}$ e $\mathbb{Z}_p^*$.

La dimostrazione del Teorema di Gauss, anche se da un certo punto di vista può essere considerata "elementare," è tecnicamente troppo complessa per essere inserita in questo articolo: si veda per esempio il Capitolo 3 di [5]. Ci limitiamo ad utilizzarlo per le applicazioni alla crittografia, segnalando che i Teoremi di Wilson e Fermat sono ingredienti fondamentali della dimostrazione.

La Figura 1 illustra il caso $p = 31$, dove si può prendere $g_p = 3$, come è possibile verificare direttamente con un po' di pazienza calcolandone le potenze successive. È anche utile (e molto piú facile) verificare che 2 non ha la stessa proprietà: infatti, le potenze successive di 2, ridotte modulo 31, valgono 1, 2, 4, 8, 16, 1, 2, 4, $\ldots$, e quindi *non* si ripetono con periodo 30, ma con periodo 5.

In generale, se scegliamo un intero a appartenente a $\mathbb{Z}_p^*$ e ne calcoliamo le potenze successive modulo p, queste si ripetono periodicamente, perché ci sono solo un numero finito di valori distinti possibili.[1] Inoltre, il Teorema di Fermat garantisce che il periodo non può superare $p-1$ (in effetti, il periodo *divide* $p-1$): il Teorema di Gauss, d'altra parte, afferma che c'è almeno un elemento di $\mathbb{Z}_p^*$ che ha periodo esattamente uguale a $p-1$, cioè che ha il massimo periodo possibile.

Quest'ultimo fatto è importantissimo per le applicazioni alla crittografia, in quanto fornisce un metodo semplice ed efficiente per "nascondere" un messaggio segreto da trasmettere: è l'argomento del prossimo paragrafo. Invitiamo i lettori a verificare con esempi concreti che se n non è un numero primo, allora non esistono interi $g \in \mathbb{Z}_n^*$ con la proprietà del Teorema di Gauss: l'esistenza di elementi di $\mathbb{Z}_n^*$ con periodo $n-1$ è una caratteristica peculiare dei numeri primi.

[1] È per questo stesso motivo che le rappresentazioni decimali dei numeri $1/p$, dove p è un numero primo diverso da 2 e da 5, sono periodiche. La lunghezza del periodo è uguale alla lunghezza del periodo della successione delle potenze 1, 10, 10^2, 10^3, $\ldots$, ridotte modulo p.

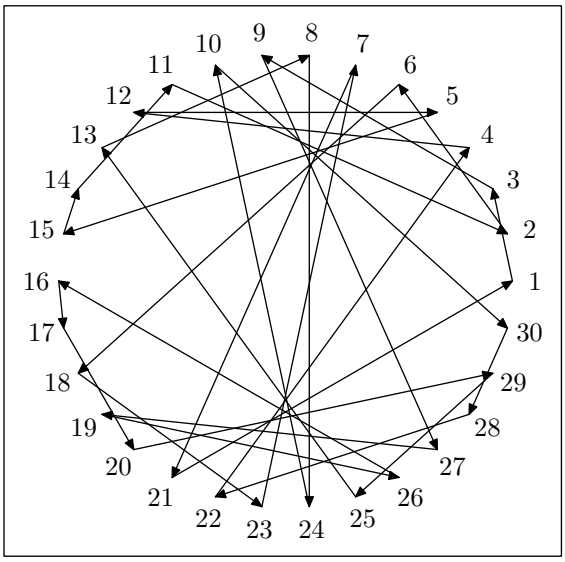

Figura 1: Le potenze successive di 3 "rimescolano" gli interi da 1 a 30 in un modo apparentemente casuale. Partendo da 1 e seguendo le frecce si trovano i numeri 1, 3, 3^2, ..., ridotti modulo 31: si noti che *tutti* gli elementi di $\mathbb{Z}_{31}^*$ sono raggiunti.

3 Applicazioni alla crittografia

Se vogliamo trasmettere un messaggio segreto via Internet, come prima cosa dobbiamo trasformarlo in una sequenza di numeri, per esempio associando ad ogni carattere il suo codice ASCII[2]. Non ci occuperemo in questa sede di come si possa materialmente eseguire questa trasformazione, ma immaginiamo che sia già stata effettuata, e quindi concentriamo la nostra attenzione esclusivamente sul problema di trasmettere dei numeri in modo sicuro. Evidentemente, chi riceve il messaggio cifrato, nella sua forma numerica, dopo averlo decifrato deve essere in grado di eseguire l'operazione inversa, e trasformare di nuovo i numeri in caratteri.

Prima di utilizzare i risultati descritti qui sopra, è necessario puntualizzare rigorosamente i termini del problema: abbiamo un insieme di *messaggi in chiaro* $\mathfrak{M}$ che due o piú soggetti vogliono poter scambiare utilizzando un mezzo di trasmissione intrinsecamente non sicuro come può essere Internet o una linea telefonica. Per poter avere la certezza di non rivelare i propri segreti ad altri, gli utenti di un sistema crittografico si accordano su un insieme di *messaggi cifrati* $\mathfrak{C}$ e su una biiezione f (nella realtà, una famiglia) fra questi due insiemi. Dunque, invece di inviare il messaggio segreto $m \in \mathfrak{M}$, l'utente A invierà $f(m) \in \mathfrak{C}$, che l'utente B può decifrare utilizzando f^{-1}, la funzione *inversa*.

[2]Non si prenda questo troppo alla lettera: in molte applicazioni reali è opportuno suddividere i messaggi in blocchi di caratteri adiacenti, o in altro modo, prima di trasformarli in numeri.

Per esempio, avendo a disposizione un numero primo grande p ed un elemento $g \in \mathbb{Z}_p^*$ con la proprietà del Teorema di Gauss, potremmo ingenuamente pensare di "nascondere" un messaggio segreto $m \in \mathbb{Z}_{p-1}$ sfruttando la biiezione descritta sopra, trasmettendo $g^m \in \mathbb{Z}_p^*$. In questo caso abbiamo scelto $\mathfrak{M} = \mathbb{Z}_{p-1}$ e $\mathfrak{C} = \mathbb{Z}_p^*$. Dal punto di vista astratto non ci sono problemi, ma questo metodo, in concreto, non funziona. Infatti, al momento attuale non sono noti metodi *efficienti* per ricavare m noti p, g e g^m: questo è il "problema del logaritmo discreto" ed è alla base della sicurezza del sistema crittografico di ElGamal, che descriveremo piú avanti. È importante sottolineare che il concetto di sicurezza è relativo allo stato dell'arte degli algoritmi esistenti: gli esperti ritengono unanimemente che il problema del logaritmo discreto sia difficile (e quindi che il crittosistema sistema di ElGamal sia sicuro) ma non è possibile escludere che qualcuno scopra in futuro un algoritmo rivoluzionario molto piú efficiente di quelli noti oggi. Considerazioni simili si applicano a tutti i crittosistemi attualmente in uso.

Il logaritmo discreto. Supponiamo di avere a disposizione un numero primo grande p ed un elemento $g \in \mathbb{Z}_p^*$ con la proprietà del Teorema di Gauss, e cioè tale che le sue potenze successive, ridotte modulo p, abbiano periodo $p-1$. Poiché vi sono esattamente $p-1$ elementi in $\mathbb{Z}_p^*$, questo significa che le potenze successive di g assumono *tutti* i valori possibili, in un qualche ordine: si veda di nuovo la Figura 1. Dato $x \in \mathbb{Z}_p^*$ ci possiamo chiedere quale sia $y \in \mathbb{Z}_{p-1}$ tale che $g^y \equiv x \bmod p$. Questo problema è *formalmente* identico al calcolo del logaritmo (che per definizione è l'esponente da dare alla base per ottenere un numero dato). Il contesto del nostro problema è un insieme *discreto*, $\mathbb{Z}_p^*$, e non *continuo* come l'insieme dei numeri reali positivi, e questo spiega il nome di "logaritmo discreto."

L'analogia formale non deve trarre in inganno: dal punto di vista pratico i due problemi sono molto diversi. L'interesse crittografico risiede nel fatto che mentre il calcolo della funzione diretta (l'esponenziale) può essere realizzato in modo efficiente anche per esponenti molto grandi mediante l'algoritmo dei quadrati ripetuti, la funzione inversa è (o meglio, si ritiene che sia) molto piú onerosa dal punto di vista computazionale. Questo è un esempio di *funzione unidirezionale*, uno dei piú importanti concetti della crittografia moderna.

Il crittosistema di ElGamal. Il crittosistema di ElGamal si basa su due fatti che abbiamo discusso sopra: sappiamo che è relativamente facile calcolare potenze, anche con esponenti grandi, ma è difficile fare l'operazione inversa. Gli utenti di questo crittosistema devono compiere un'operazione preliminare, e cioè scegliere di comune accordo un numero primo grande p ed un elemento $g \in \mathbb{Z}_p^*$ che abbia la proprietà del Teorema di Gauss. L'insieme dei messaggi in chiaro e dei messaggi cifrati è lo stesso: $\mathfrak{M} = \mathfrak{C} = \mathbb{Z}_p^*$. Ogni utente di questo crittosi-

stema sceglie la propria *chiave privata*, da mantenere assolutamente segreta: si tratta di un elemento qualsiasi x di $\mathbb{Z}_{p-1}$. L'utente in questione rende pubblica la quantità $g^x \bmod p \in \mathbb{Z}_p^*$, che si chiama *chiave pubblica*, ed è facilmente calcolabile mediante l'algoritmo dei quadrati ripetuti: è evidente che se il problema del logaritmo discreto fosse facile, allora dal valore della chiave pubblica si potrebbe ricavare efficientemente quello della chiave privata, e la sicurezza di questo metodo crittografico verrebbe a cadere.

Vogliamo ora mostrare come due utenti del sistema crittografico di ElGamal possano comunicare in sicurezza: supponiamo dunque che l'utente A abbia scelto la propria chiave privata $\alpha \in \mathbb{Z}_{p-1}$ e che abbia reso nota (per esempio sulla sua pagina web) la chiave pubblica $a \equiv g^\alpha \bmod p$. Evidentemente dobbiamo anche supporre che l'utente B abbia fatto una scelta analoga di chiave privata $\beta \in \mathbb{Z}_{p-1}$ e reso disponibile a tutti la quantità $b \equiv g^\beta \bmod p$. Queste operazioni vanno eseguite una volta sola.

Se A vuole comunicare a B il messaggio $m \in \mathbb{Z}_p^*$, può procedere in questo modo: sceglie in modo del tutto arbitrario un parametro $t \in \mathbb{Z}_{p-1}$, calcola le due quantità $x \equiv g^t \bmod p$ ed $y \equiv m \cdot b^t \bmod p$ e le trasmette, in questo ordine, a B. A sua volta, B calcola $x^{-1} \bmod p$ (l'inverso di $x \bmod p$, cioè l'unico intero $z \in \mathbb{Z}_p^*$ tale che $x \cdot z \equiv 1 \bmod p$) e poi $y \cdot z^\beta \bmod p$, e può cosí leggere il messaggio m. Infatti

$$y \cdot z^\beta \equiv m \cdot b^t \cdot x^{-\beta} \equiv m \cdot g^{\beta t} \cdot g^{-\beta t} \equiv m \mod p.$$

Per usare un linguaggio volutamente pittoresco, potremmo dire che A *camuffa* il messaggio m mediante la quantità b^t, e poi invia a B il messaggio camuffato e l'informazione che consentirà a B (*solo* a B, perché solo B conosce β) di togliere la maschera e leggere il messaggio originale.

A Le congruenze

Dato un numero intero positivo n, consideriamo la proprietà di divisibilità per n: in un certo senso, vogliamo generalizzare il concetto di "pari" e "dispari", che è basato sulla divisibilità per 2. Fissiamo un intero $n \geq 2$; diciamo che due interi a e b (eventualmente negativi) sono *congrui modulo n* se la loro differenza $a - b$ è divisibile per n: in questo caso scriveremo $a \equiv b \bmod n$. Poniamo $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. Dato un qualsiasi intero m (eventualmente negativo) esiste un unico elemento r di $\mathbb{Z}_n$ tale che $m - r$ sia divisibile per n, e cioè tale che $m \equiv r \bmod n$. Per esempio, se $n = 2$, allora $\mathbb{Z}_2 = \{0, 1\}$, e possiamo scegliere $r = 0$ per gli interi m pari, ed $r = 1$ per gli interi m dispari. In questo modo, possiamo ripartire l'insieme di tutti gli interi $\mathbb{Z}$ in n classi, associando ad ogni intero m l'elemento $r \in \mathbb{Z}_n$ che ha la proprietà suddetta. L'elemento $r \in \mathbb{Z}_n$ che corrisponde ad m si

dice *residuo di m modulo n* e l'operazione che associa m ad r si dice *riduzione modulo n*; scriveremo $r = m \bmod n$ per indicare il risultato della riduzione.

L'aspetto piú importante del concetto di congruenza è la sua proprietà di *compatibilità con le operazioni*: scelti comunque a, b e $c \in \mathbb{Z}$, se $a \equiv b \bmod n$ allora

$$a+c \equiv b+c \mod n, \qquad ac \equiv bc \mod n.$$

Infatti, è del tutto evidente che se n divide $a-b$, allora n divide sia $(a+c)-(b+c) = a-b$, sia $ac-bc = (a-b)c$. Da questa proprietà segue che eseguire le operazioni in $\mathbb{Z}$ e poi ridurre modulo n oppure eseguire le stesse operazioni sui residui (e, se necessario, ridurre un'altra volta) dà esattamente lo stesso risultato: per esempio, $17 \cdot 25 = 425 \equiv 7 \bmod 11$, $17 \equiv 6 \bmod 11$, $25 \equiv 3 \bmod 11$. Inoltre $6 \cdot 3 = 18 \equiv 7 \bmod 11$.

L'idea di congruenza è importante perché fornisce un'infinità di sistemi numerici *finiti* nei quali è possibile fare "esperimenti" e verificarli esaminando tutti i casi possibili. Per esempio, è piuttosto facile dimostrare che l'equazione $x^2 = 3y^2 - 1$ non ha nessuna soluzione con x, $y \in \mathbb{Z}$: infatti, se esistesse una soluzione di questa equazione, riducendo l'uguaglianza modulo 3 troveremmo $x^2 \equiv -1 \bmod 3$, e si vede immediatamente che questa è impossibile esaminando i 3 casi $x \equiv 0$, 1, 2 mod 3. Dunque abbiamo ricondotto un problema *infinito* ad un numero finito di verifiche. Usando le congruenze è possibile dare una semplice giustificazione della validità dei "criteri di divisibilità" e della "prova del nove."

Il concetto di congruenza è relativamente recente (risale all'inizio del XIX secolo, ed è stato introdotto da Gauss), ma la divisibilità è stata studiata fino dall'alba della matematica: il primo risultato importante al proposito, che di nuovo sottolinea l'importanza dell'essere primo, è addirittura un Teorema di Euclide.

Teorema A.1 (Euclide) *Se p è un numero primo che divide il prodotto ab, allora p divide a oppure p divide b (eventualmente entrambe queste cose).*

Per la dimostrazione rimandiamo a testi piú completi. Si noti che, per esempio, 10 divide $4 \cdot 5$ ma non divide né 4 né 5.

Lemma A.2 *Se p è un numero primo ed a è un intero qualsiasi non divisibile per p allora esiste un unico intero $b \in \{1, 2, \ldots, p-1\}$ tale che $ab \equiv 1 \bmod p$.*

Dim. Moltiplichiamo i numeri 1, 2, …, $p-1$, per a e consideriamo i resti r_1, r_2, …, r_{p-1} che si ottengono dividendo ciascun prodotto per p. Vogliamo dimostrare che sono tutti diversi fra loro: supponiamo per assurdo che $r_i = r_j$ con $i > j$. Allora $ai \equiv aj \bmod p$ e quindi $a(i-j) \equiv 0 \bmod p$, che è assurdo poiché, per il Teorema di Euclide A.1, se p divide un prodotto allora divide uno dei fattori. Ma in questo caso p non può dividere a (per ipotesi) né $i-j$, perché sia i che j sono

numeri interi positivi minori di p e quindi $0 < i - j < p$. Quindi tutti i numeri r_i sono distinti, ed uno di questi è necessariamente uguale ad 1. □

L'elemento b dell'enunciato si dice *inverso di a modulo p* e qualche volta si indica con a^{-1}. Naturalmente resta aperto il problema di come determinare b una volta dato a: non vogliamo certo procedere alla cieca, come nella dimostrazione del Lemma. Un metodo molto efficiente è l'algoritmo di Euclide esteso, descritto per esempio nel §6.2 di [5].

Indicheremo con $\mathbb{Z}_n^*$ il sottoinsieme di $\mathbb{Z}_n$ degli elementi a per cui esiste $b \in \mathbb{Z}_n$ tale che $ab \equiv 1 \bmod n$. Per esempio, $\mathbb{Z}_{12}^* = \{1, 5, 7, 11\}$. Il Lemma precedente implica che $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ se p è un numero primo: questa è una proprietà caratteristica dei numeri primi. Infatti, se n non è primo, allora esistono interi a e α tali che $a\alpha = n$ e $1 < a \leq \alpha < n$, e non è difficile verificare che né a né α hanno la proprietà del Lemma: se esistesse b tale che $ab \equiv 1 \bmod n$, per definizione dovrebbe esistere un intero k tale che $ab - 1 = kn$, da cui $ab - kn = 1$. Il primo membro di questa uguaglianza è certamente divisibile per $a > 1$, che è assurdo. La stessa idea dimostra che se il massimo comun divisore fra m ed n è maggiore di 1, allora $m \notin \mathbb{Z}_n^*$, e, in effetti, $\mathbb{Z}_n^*$ è proprio l'insieme degli elementi di $\mathbb{Z}_n$ che non hanno fattori comuni con n, a parte 1.

Lemma A.3 *Se p è un numero primo ed x è un qualsiasi numero intero tale che $x^2 \equiv 1 \bmod p$, allora $x \equiv 1 \bmod p$ oppure $x \equiv p - 1 \bmod p$.*

Dim. Scriviamo la congruenza $x^2 \equiv 1 \bmod p$ nella forma equivalente $x^2 - 1 \equiv 0 \bmod p$, e cioè $(x-1)(x+1) \equiv 0 \bmod p$. Per il Teorema di Euclide A.1, questo significa che $x - 1 \equiv 0 \bmod p$ oppure $x + 1 \equiv 0 \bmod p$, che equivale alla tesi. □

La conseguenza piú importante per noi è che i soli elementi di $\mathbb{Z}_p^*$ che sono uguali al proprio inverso in $\mathbb{Z}_p^*$ sono 1 e $p-1$: infatti, se $x \equiv x^{-1} \bmod p$, moltiplicando ambo i membri per x troviamo $x^2 \equiv 1 \bmod p$. Anche questa è una proprietà caratteristica dei numeri primi (ed anche del numero 4 e dei numeri della forma p^α, $2p^\alpha$, $4p^\alpha$, dove p è un numero primo dispari, per la precisione): se n non è primo, l'equazione $x^2 \equiv 1 \bmod n$ può avere un numero di soluzioni maggiore di 2. Per esempio, l'equazione $x^2 \equiv 1 \bmod 8$ ha quattro soluzioni. Lasciamo ai lettori il compito di trovarle.

Dimostrazione del Teorema di Wilson. Prendiamo un numero primo $p > 2$ (se $p = 2$ la tesi è ovvia), e consideriamo il prodotto $(p-1)! = 1 \cdot 2 \cdots (p-1)$. Possiamo associare ciascun intero $2 \leq a \leq p-2$ al suo numero corrispondente b che ci è fornito dal Lemma A.2 e che è diverso da a per il Lemma A.3. Questo ci assicura che $(p-1)! \equiv 1 \cdot (p-1) \bmod p$ (poiché tutti gli altri fattori danno 1), e quindi $(p-1)! + 1 \equiv p - 1 + 1 \equiv 0 \bmod p$, cioè p divide $(p-1)! + 1$.

Non è difficile vedere che vale anche il viceversa: in altre parole, se $n \geq 2$ non è primo allora non soddisfa la condizione di Wilson. Distinguiamo due casi: n è

il quadrato di un numero primo (cioè $n = p^2$), oppure n non è il quadrato di un numero primo. In quest'ultimo caso, esistono due interi a e b tali che $1 < a < b < n$ e $ab = n$. Poiché sia a che b sono fattori di $(n-1)!$ (e sono distinti), evidentemente n divide $(n-1)!$. Nell'altro caso, se $n = 4$ si ha $3! + 1 = 7 \not\equiv 0 \bmod 4$. Se invece $n = p^2$ con $p > 2$, i numeri p e $2p$ sono fattori distinti di $(n-1)!$, ed anche in questo caso n divide $(n-1)!$. □

Dimostrazione del Teorema di Fermat. È simile a quella del Lemma A.2: moltiplichiamo i numeri 1, 2, …, $p-1$, per a e consideriamo i resti r_1, r_2, …, r_{p-1} che si ottengono dividendo ciascun prodotto per p. Sappiamo già che sono tutti diversi fra loro, e quindi, eventualmente in un altro ordine, sono i numeri 1, 2, 3, …, $p-1$. Abbiamo dunque $p-1$ congruenze del tipo $a \cdot i \equiv r_i \bmod p$: moltiplicandole tutte fra loro otteniamo

$$a^{p-1} \cdot 1 \cdot 2 \cdots (p-1) \equiv r_1 \cdot r_2 \cdots r_{p-1} = 1 \cdot 2 \cdots (p-1) \mod p,$$

per quando detto sopra. In altre parole, abbiamo trovato che $a^{p-1} \cdot (p-1)! \equiv (p-1)! \bmod p$: ricordando che $(p-1)! \equiv -1 \bmod p$ per il Teorema di Wilson, possiamo concludere la dimostrazione. □

B Scomposizione in fattori primi

Abbiamo già discusso sopra dei metodi per scomporre un numero intero n in fattori primi. Fra questi, di gran lunga il piú semplice è quello della divisione per tentativi, che però ha il problema che può richiedere quasi $\sqrt{n}$ operazioni per scomporre in fattori dei numeri n che hanno esattamente 2 fattori primi molto vicini fra loro, come per esempio $n = 7\,851\,203 = 2801 \cdot 2803$. In questo caso potrebbe sembrare piú efficiente un altro metodo, basato su una semplice osservazione: se riusciamo a trovare due numeri interi x ed y tali che $n + y^2 = x^2$, allora $n = x^2 - y^2 = (x-y) \cdot (x+y)$ e quindi abbiamo scomposto n in due fattori. Naturalmente $x - y$ ed $x + y$ non sono necessariamente primi, e, peggio, è anche possibile che uno dei due sia proprio uguale ad 1, rendendo inutile questa scomposizione.

Rimandiamo i lettori interessati alla monografia di Crandall & Pomerance [2] per una descrizione dettagliata dei piú moderni metodi di fattorizzazione: ci limitiamo ad illustrarne uno molto semplice, ma pur sempre non banale, che sfrutta le idee discusse qui sopra.

Dato un intero dispari n, il nostro obiettivo è la determinazione di due interi x ed y tali che $n = x^2 - y^2$. Dato che $x^2 = n + y^2$, sappiamo che $x \geq \sqrt{n}$, e quindi una strategia possibile è quella di ripetere i seguenti passi:

1. poniamo $x = \left[\sqrt{n}\right] + 1$;

2. determiniamo $a = x^2 - n$;
3. se esiste un intero y tale che $a = y^2$, stampa $n = (x-y)(x+y)$; l'algoritmo termina;
4. aumentiamo x di un'unità, e torniamo al passo 2.

Per motivi storici, questo metodo si chiama algoritmo di Fermat. Osserviamo che l'algoritmo termina (prima o poi) perché quando $x = \frac{1}{2}(n+1)$ si trova $y = \frac{1}{2}(n-1)$: purtroppo, questi valori danno luogo alla scomposizione banale $n = 1 \cdot n$, come i lettori possono verificare senza difficoltà. Osserviamo altresí che esistono algoritmi molto efficienti (quello di Newton, per esempio; peccato che di solito non venga insegnato nelle scuole superiori ...) per determinare la radice quadrata approssimata di un numero reale positivo. Una semplice descrizione dell'algoritmo di Newton si può trovare in [9].

La ricerca esaustiva di x ed y, descritta qui sopra, non è efficiente dal punto di vista computazionale, ma una semplice osservazione basata sulle congruenze ci permetterà di accelerare molto il calcolo. Per semplicità ci occupiamo di un caso concreto, e lasciamo ai lettori il compito di enunciare questo algoritmo in generale. Vogliamo scomporre in fattori primi l'intero $n = 55\,969\,103$ determinando due interi positivi x ed y in modo che

$$x^2 - y^2 = n \tag{3}$$

con $x - y > 1$. Osserviamo subito che dalla (3) segue immediatamente che $x^2 = y^2 + n \geq n$, e quindi che $x \geq \sqrt{n} \approx 7481.25\ldots$: dato che x è un numero intero, abbiamo che $x \geq 7482$. Consideriamo ora la stessa relazione (3) modulo 4:

$$x^2 - y^2 = n \equiv 3 \mod 4. \tag{4}$$

Qual è il vantaggio? Il vantaggio è che ora possiamo considerare solo x ed y appartenenti all'insieme $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, e determinare quali di questi soddisfano la congruenza (4). L'analisi di tutti i casi possibili rivela che $x \bmod 4 \in \{0, 2\}$, cioè x è pari. In altre parole, *ogni* soluzione intera di (3) ha un valore pari di x, ed è perfettamente inutile eseguire il passo 3 dell'algoritmo illustrato sopra quando x è dispari. Dunque, abbiamo *dimezzato* il numero di iterazioni necessarie!

Incoraggiati da questo successo, ripetiamo il ragionamento con altri valori; consideriamo la relazione (3) modulo 3:

$$x^2 - y^2 = n \equiv 2 \mod 3.$$

Questa volta cerchiamo x ed y nell'insieme $\mathbb{Z}_3 = \{0, 1, 2\}$, e troviamo che $x \equiv 0 \bmod 3$. Infine, consideriamo la relazione (3) modulo 5:

$$x^2 - y^2 = n \equiv 3 \mod 5,$$

dove x ed y appartengono all'insieme $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Un semplice calcolo rivela che x mod 5 appartiene all'insieme $\{2, 3\}$.

Riassumendo quanto trovato finora, sappiamo che se x è un intero che soddisfa (3) allora $x \geq 7482$, è divisibile per 2 e per 3 (cioè è divisibile per 6) ed è congruo a 2 o 3 modulo 5. Notiamo che 7482 è multiplo di 6, ed elenchiamo i successivi multipli di 6, insieme alla loro classe di resto modulo 5:

x	7482	7488	7494	7500	7506	7512	7518	7524	7530
x mod 5	2	3	4	0	1	2	3	4	0
	$*$	$*$				$*$	$*$		

I valori di x indicati da $*$ sono ammissibili modulo 5, e quindi sono i soli valori da tentare nel passo 3 dell'algoritmo di Fermat: i primi due valori danno luogo ad y non interi, mentre il terzo ci dà

$$7512^2 - 55969103 = 461041 = 671^2$$

dalla quale ricaviamo $55969103 = (7512 - 679)(7512 + 679) = 6833 \cdot 8191$. Per completezza, si dovrebbe poi dimostrare che i due fattori trovati sono effettivamente numeri primi, ma è evidente che abbiamo in ogni caso ridotto notevolmente la complessità del problema.

Se avessimo studiato anche le congruenze modulo 7, avremmo potuto eliminare il valore $x = 7488$. Se invece di considerare congruenze modulo 4 le avessimo considerate modulo 8, avremmo scoperto che $x \equiv 0 \bmod 4$, e avremmo potuto eliminare i valori $x = 7482, 7494, 7506, 7518, 7530$. Il numero primo $p = 2$ è evidentemente speciale: per esempio, l'equazione $x^2 \equiv 1 \bmod p$ ha due soluzioni distinte modulo p (e cioè $x = 1$ ed $x = p - 1$) per $p > 2$, mentre per $p = 2$ ha una sola soluzione e cioè $x = 1$. È per questo motivo che è preferibile considerare le congruenze modulo $2^2 = 4$ oppure $2^3 = 8$, o con esponenti ancora piú grandi, mentre per tutti gli altri primi questa scelta non risulta necessaria.

Per concludere, osserviamo che tutte le operazioni coinvolte sono molto semplici, tanto che è pensabile di eseguire l'intero algoritmo senza l'aiuto di una calcolatrice. Questo algoritmo è (relativamente) efficiente poiché si basa su un procedimento di *crivello* che permette di eliminare intere classi di congruenza ad ogni iterazione, ma esistono altri algoritmi, molto piú efficaci, che si basano sempre sull'idea di crivello. Il crivello piú famoso è indubbiamente quello di Eratostene, ed è descritto nel Capitolo 5 di [1], nel Capitolo 6 di [5] e in [8]. Un algoritmo di fattorizzazione estremamente efficiente (il crivello quadratico) che si basa su una evoluzione delle idee descritte in questo paragrafo è descritto nei dettagli nel §6.1 di [2] e nel §6.6.1 di [5]: si tratta di uno dei due algoritmi che al momento attuale si contendono il primato per il *miglior* algoritmo di fattorizzazione in assoluto.

C Letture ulteriori

Per le congruenze si veda il Capitolo 2 di Conway & Guy [1], l'articolo di Pomerance [6], oppure il Capitolo 2 di Languasco & Zaccagnini [5].

Alcune importanti proprietà dei numeri primi sono descritte, a livello divulgativo, nel libro di Conway & Guy [1], in una delle conferenze di Lang della raccolta [4] e nell'articolo di Pomerance [6]. Per approfondimenti si consiglia di consultare le monografie di Hardy & Wright [3] e quella di Ribenboim [7]. In particolare, gli pseudoprimi sono trattati nel §2.VIII del libro di Ribenboim.

Criteri di primalità ed algoritmi di fattorizzazione sono descritti in grande dettaglio in Crandall & Pomerance [2] ed in qualche caso interessante anche in Languasco & Zaccagnini [5]. L'algoritmo dei "quadrati ripetuti" è spiegato in [6] e in [5], e sue generalizzazioni e miglioramenti in [2]. Le applicazioni alla crittografia sono discusse in grande dettaglio in [5].

Riferimenti bibliografici

[1] J. H. Conway, R. K. Guy. *Il libro dei numeri*. Hoepli, Milano, 1999.

[2] R. Crandall, C. Pomerance. *Prime numbers. A computational perspective*. Springer-Verlag, New York, 2001.

[3] G. H. Hardy, E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford Science Publications, Oxford, fifth edition, 1979.

[4] S. Lang. *La bellezza della matematica*. Bollati Boringhieri, 1991.

[5] A. Languasco, A. Zaccagnini. *Introduzione alla crittografia*. Ulrico Hoepli Editore, Milano, 2004.

[6] C. Pomerance. Alla ricerca dei numeri primi. *Le Scienze*, 174:86–94, febbraio 1983.

[7] P. Ribenboim. *The New Book of Prime Numbers Records*. Springer-Verlag, New York, 1996.

[8] A. Zaccagnini. Variazioni Goldbach: problemi con numeri primi. *L'Educazione Matematica, Anno XXI, Serie VI*, 2:47–57, 2000. `http://www.math.unipr.it/~zaccagni/psfiles/papers/Goldbach_I.pdf`.

[9] A. Zaccagnini. Formato A4. *L'Educazione Matematica, Anno XXIV, Serie VII*, 1:47–54, 2003. `http://www.math.unipr.it/~zaccagni/psfiles/papers/FormatoA4.pdf`.