ALCUNE PROPRIETÀ DEI NUMERI PRIMI E LORO APPLICAZIONI ALLA CRITTOGRAFIA

Alessandro Zaccagnini

30.11.2001 — 7.12.2001

Docente	Alessandro Zaccagnini
Indirizzo	Dipartimento di Matematica, via Massimo d'Azeglio, 85/a, 43100 Parma
Telefono	0521 032302 (centralino 032300)
Fax	0521 032350
e-mail	`alessandro.zaccagnini@unipr.it`
pagina web	`http://www.math.unipr.it/~zaccagni/home.html`

Questo testo è disponibile all'indirizzo
`http://www.math.unipr.it/~zaccagni/psfiles/Crittografia.pdf`

Introduzione

Queste note accompagnano le mie lezioni sulle applicazioni di alcune proprietà dei numeri primi alla crittografia. Il problema che si incontra nel presentare queste idee agli studenti del Corso di Laurea in Ingegneria dipende dal fatto che la matematica studiata ad Ingegneria (con l'eccezione di qualche rudimento di calcolo combinatorio) è tutta matematica delle grandezze continue (analisi, geometria, meccanica), mentre l'informazione è per sua natura discreta. Ho dunque cercato di gettare un ponte fra il mondo continuo e quello discreto, cominciando da un problema di natura algebrica la cui soluzione dovrebbe essere già nota. Può darsi che i concetti introdotti all'inizio possano sembrare, proprio per questo motivo, piuttosto remoti dalla crittografia e qualche ripetizione è stata inevitabile, ma prego i lettori di avere pazienza e di credermi sulla fiducia. Ho ritenuto opportuno aggiungere a queste note anche delle informazioni sulla distribuzione dei numeri primi non immediatamente pertinenti alla crittografia ed alcuni esempi numerici di cui, evidentemente, non è possibile parlare durante le lezioni per mancanza di tempo.

Aritmetica e Numeri primi

L'equazione $z^n = 1$

Un possibile argomento di raccordo fra la matematica del continuo e quella del discreto è dato dallo studio delle soluzioni dell'equazione $z^n = 1$ dove n è un intero positivo fissato e $z \in \mathbb{C}$. Fissiamo dunque un numero naturale $n \geq 1$ e poniamo $\delta_n = e^{2\pi i/n}$. In questo modo, le soluzioni dell'equazione $z^n = 1$ sono ordinate nel modo "naturale" se poniamo $z_j = e^{2\pi ij/n} = \delta_n^j$, $j = 0, \dots, n-1$. (Per la precisione, dovremmo scrivere $z_{j,n}$, ma di solito per noi n è fissato e quindi lo ometteremo senza ambiguità). È ben noto che i punti z_j sono disposti ai vertici del poligono regolare con n lati (almeno per $n \geq 3$) inscritto nella circonferenza di centro l'origine e raggio 1, con un vertice nel punto $z_0 = 1$. Scriveremo $\mathcal{U}_n = \{z_0, \dots, z_{n-1}\}$.

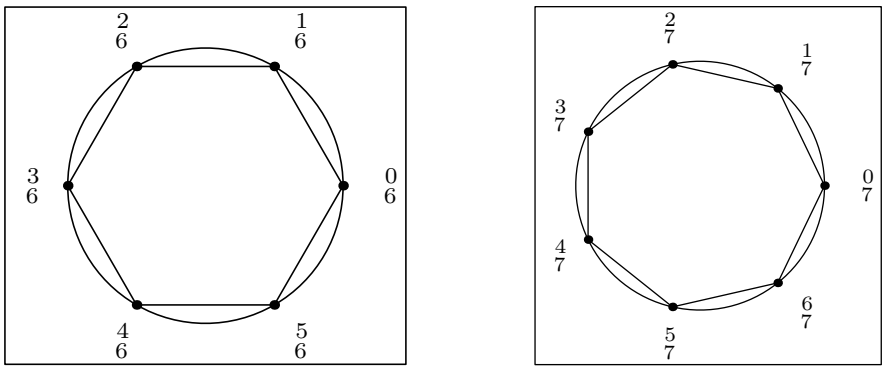


Figura 1. *Gli insiemi $\mathcal{U}_6$ ed $\mathcal{U}_7$.*

Dal punto di vista dell'Analisi Matematica, non c'è grande differenza fra $\mathcal{U}_6$ ed $\mathcal{U}_7$: ci proponiamo di mostrare che, invece, dal punto di vista dell'Aritmetica questi due insiemi sono piuttosto diversi fra loro. Cominciamo con qualche osservazione sulla struttura "moltiplicativa" dell'insieme $\mathcal{U}_n$. Ricordiamo le proprietà commutativa ed associativa del prodotto valide in tutto $\mathbb{C}$ (e dunque per il prodotto di elementi di $\mathcal{U}_n$) e che la moltiplicazione per z_j corrisponde ad una rotazione di $\mathbb{C}$ attorno all'origine in senso antiorario dell'angolo $2\pi\frac{j}{n}$.

- per $j,\, k \in \{0,\, \dots\ n-1\}$ si ha

$$z_j \cdot z_k = \delta_n^j \cdot \delta_n^k = \begin{cases} \delta_n^{j+k} & \text{se } j+k < n, \\ \delta_n^{j+k-n} & \text{se } j+k \geq n. \end{cases}$$

- per $j \in \{0,\, \dots\, ,\, n-1\}$ si ha $(\delta_n^j)^{-1} \in \mathcal{U}_n$ e

$$(z_j)^{-1} = (\delta_n^j)^{-1} = \overline{\delta_n}^j = \begin{cases} \delta_n^{n-j} & \text{se } j \neq 0, \\ \delta_n^0 = 1 & \text{se } j = 0. \end{cases}$$

Queste proprietà, insieme al fatto che $1 = z_0 \in \mathcal{U}_n$, si riassumono dicendo che $\mathcal{U}_n$ con l'operazione di prodotto è un *gruppo abeliano* o *commutativo*. In altre parole, $\mathcal{U}_n$ ha elemento neutro rispetto all'operazione di prodotto, questa operazione è commutativa ed associativa, ed ogni elemento ha inverso.

- sia $m \in \mathbb{N}$: possiamo trovare quoziente q e resto r della divisione di m per n; si ha quindi $m = qn + r$ dove $0 \leq r < n$. Dunque

$$\delta_n^m = \delta_n^{qn+r} = \left(\delta_n^n\right)^q \cdot \delta_n^r = 1^q \cdot \delta_n^r = \delta_n^r.$$

  In altre parole, le potenze di δ_n dipendono solo dal resto della divisione dell'esponente per n. Questo è implicito nel primo punto qui sopra e dipende dal fatto che la funzione $x \mapsto e^{2\pi i x}$ ha periodo 1.

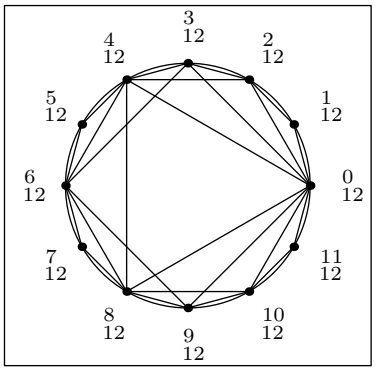
In particolare, questo significa che $z_j^{mn} = 1$ per ogni $m \in \mathbb{Z}$ e quindi che le potenze di z_j sono *periodiche* con periodo che non supera n.

- se $z \in \mathcal{U}_n \cap \mathcal{U}_m$ (cioè se $z^n = z^m = 1$) allora $z^{\lambda n + \mu m} = 1$ per ogni scelta di $\lambda,\, \mu \in \mathbb{Z}$, e quindi $z^{(n,m)} = 1$, dove (n, m) indica il *massimo comun divisore* di n ed m (Algoritmo di Euclide, *infra*). Per esempio, se $z^{120} = z^{51} = 1$ allora $z^{3\cdot 120 - 7\cdot 51} = \left(z^{120}\right)^3 \cdot \left(z^{51}\right)^{-7} = 1$, da cui $z^3 = 1$, ed infatti $(120, 51) = 3$. La stessa cosa si può vedere cosí:

$$\begin{aligned} 120 = 2 \cdot 51 + 18 \quad &\Longrightarrow \quad 1 = z^{120} = \left(z^{51}\right)^2 \cdot z^{18} = z^{18} \\ 51 = 2 \cdot 18 + 15 \quad &\Longrightarrow \quad 1 = z^{51} = \left(z^{18}\right)^2 \cdot z^{15} = z^{15} \\ 18 = 1 \cdot 15 + 3 \quad &\Longrightarrow \quad 1 = z^{18} = \left(z^{15}\right)^1 \cdot z^3 = z^3 \end{aligned}$$

- se $z \in \mathcal{U}_n$ (cioè se $z^n = 1$) allora esiste un unico intero d che divide n (scriveremo questa relazione nella forma $d \mid n$) tale che $z^d = 1$, $z^\delta \neq 1$ per ogni $0 < \delta < d$. Infatti, sia d il minimo intero positivo tale che $z^d = 1$. Dunque $d \leq n$ ed inoltre, per il punto precedente, si ha $z^{(n,d)} = 1$. Ma $(n, d) \leq d$ e quindi, per la minimalità di d, deve essere $(n, d) = d$, cioè $d \mid n$. L'intero d si dice *ordine* di z in $\mathcal{U}_n$. Per esempio, in $\mathcal{U}_{12}$

	ordine
z_0	1
z_6	2
$z_4,\, z_8$	3
$z_3,\, z_9$	4
$z_2,\, z_{10}$	6
$z_1,\, z_5,\, z_7,\, z_{11}$	12



Abbiamo visto sopra che le potenze di $z \in \mathcal{U}_n$ sono periodiche con un periodo che divide n: l'ordine di z è precisamente il *minimo* periodo delle potenze di z, cioè il minimo periodo della successione periodica $1, z, z^2, z^3, \dots$

Dal punto di vista insiemistico, dire che $z \in \mathcal{U}_n$ ha ordine d significa che $z \in \mathcal{U}_d$ e che $z \notin \mathcal{U}_\delta$ per ogni $\delta < d$, mentre dal punto di vista geometrico significa che z è uno dei vertici del poligono regolare (tra quelli presi in considerazione qui) con d lati e di nessun poligono regolare con un numero di lati minore.

Queste ultime idee ci permettono di notare le prime differenze tra i diversi valori di n: poiché l'ordine di $z \in \mathcal{U}_n$ è un divisore di n, se n è un numero primo l'ordine è necessariamente 1 oppure n, e quindi o $z = 1$ oppure il suo ordine è n. Un elemento z di $\mathcal{U}_n$ che ha ordine n si dice *generatore* (qualche volta anche *radice primitiva*) poiché ha l'importante proprietà che le sue potenze successive forniscono *tutti* gli elementi di $\mathcal{U}_n$: consideriamo gli elementi di $\mathcal{U}_n$

$$1 = z^0, z^1, z^2, z^3, \dots, z^{n-1}. \tag{1}$$

Affermo che essi sono tutti distinti: infatti, se $z^j = z^k$ per qualche coppia $0 \le j < k < n$ allora $z^{k-j} = 1$, ma $1 \le k - j < n$ e questo è impossibile. La (1) implica dunque che gli n elementi indicati sono tutti e soli gli elementi di $\mathcal{U}_n$. Si noti che questa cosa non è vera se z non è un generatore: per esempio, se prendiamo z_3 in $\mathcal{U}_{12}$ troviamo che le sue potenze successive sono $z_3^0 = 1$, $z_3^1 = z_3$, $z_3^2 = -1$, $z_3^3 = -z_3$, $z_3^4 = 1$ (in effetti $z_3 = e^{2\pi i \cdot 3/12} = i$, l'unità immaginaria).

Possiamo anche osservare che mentre le potenze successive di z_1 danno tutti gli elementi di $\mathcal{U}_{12}$ nell'ordine naturale, e le potenze di $z_{11} = z_1^{-1}$ nell'ordine inverso, le potenze successive di z_7 sono le seguenti:

$$\begin{array}{llllll} z_7^0 = z_0 & z_7^1 = z_7 & z_7^2 = z_2 & z_7^3 = z_9 & z_7^4 = z_4 & z_7^5 = z_{11} \\ z_7^6 = z_6 & z_7^7 = z_1 & z_7^8 = z_8 & z_7^9 = z_3 & z_7^{10} = z_{10} & z_7^{11} = z_5 \end{array}$$

cioè le potenze successive di z_7 ci danno un modo piuttosto semplice di "rimescolare" gli elementi di $\mathcal{U}_{12}$. Questo fatto è importante in vista delle applicazioni alla crittografia. Conviene notare che qualunque sia $n \ge 2$, $\delta_n = e^{2\pi i/n}$ è un generatore di $\mathcal{U}_n$: quindi $\mathcal{U}_n$ è un gruppo abeliano *ciclico*; come vedremo piú avanti il numero di generatori di $\mathcal{U}_n$ cresce con n (in modo irregolare e piuttosto complicato).

In definitiva, in questo paragrafo abbiamo visto che le proprietà aritmetiche di $\mathcal{U}_n$ dipendono dalla divisibilità, e questo serve a motivare la discussione che segue. Abbiamo dimostrato che l'insieme $\mathcal{U}_n = \{z_0, \dots, z_{n-1}\}$, dotato dell'operazione $\cdot$ è un *gruppo abeliano ciclico* generato da z_1 (e non solo: se $n \ge 3$ c'è almeno un altro generatore, z_{n-1}). Abbiamo anche visto che c'è una corrispondenza naturale (che i matematici chiamano *isomorfismo*) con l'insieme $\mathbb{Z}_n \overset{\text{def}}{=} \{0, \dots, n-1\}$ delle classi di resto modulo n, dotato dell'operazione di addizione con resto. La corrispondenza è, in un certo senso, l'analogo discreto del logaritmo.

Il gruppo ciclico $\mathbb{Z}_n$

Vogliamo dotare l'insieme $\mathbb{Z}_n$ delle operazioni di addizione e di moltiplicazione facendole derivare dalle analoghe in $\mathbb{Z}$, come suggerito dalle proprietà di $\mathcal{U}_n$ viste sopra. Ricordiamo che dato un intero qualsiasi $m \in \mathbb{Z}$ è sempre possibile trovare altri due interi $q_m \in \mathbb{Z}$, $r_m \in \mathbb{Z}_n$ (quoziente e resto) tali che

$$m = q_m \cdot n + r_m;$$

(se si definisce $[x] = \max\{n \in \mathbb{Z} : n \leq x\}$ in modo che, per esempio, $[\pi] = 3$, $[-\pi] = -4$, allora $q_m = \left[\frac{m}{n}\right]$, $r_m = m - q_m \cdot n$, anche se $m < 0$). Questo procedimento ci dà un'applicazione fra $\mathbb{Z}$ e $\mathbb{Z}_n$ definita da

$$m \mapsto r_m.$$

Non è difficile dimostrare che questa applicazione (detta anche *riduzione modulo* n) è compatibile con le operazioni di $\mathbb{Z}$, nel senso che per ogni a, b, $c \in \mathbb{Z}$ si ha

$$\begin{aligned} a + b = c \quad &\Longleftrightarrow \quad r_{a+b} = r_c \\ a \cdot b = c \quad &\Longleftrightarrow \quad r_{a\cdot b} = r_c \end{aligned}$$

In generale, le applicazioni compatibili con le operazioni di due insiemi diversi, si chiamano *omomorfismi*. Le operazioni aritmetiche in $\mathbb{Z}_n$ corrispondono alle analoghe in $\mathbb{Z}$, con la differenza che è necessario prendere il resto della divisione per n.

Si può notare che l'applicazione $m \mapsto r_m$ non è iniettiva (per esempio, l'immagine di tutti i multipli di n è 0): l'insieme $\mathbb{Z}$ viene ripartito negli n sottoinsiemi degli elementi che hanno lo stesso valore di r, dette *classi di congruenza modulo* n. In altre parole, due interi a e b sono nella stessa classe di congruenza modulo n se $r_a = r_b$, e la classe di congruenza di a è $r_a^{-1} = \{a,\, a \pm n,\, a \pm 2n,\, \dots\}$.

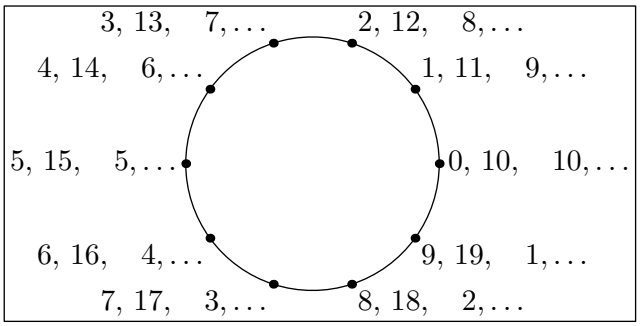


Figura 2. *Le classi di congruenza mod 10 appaiono avvolgendo la retta reale sulla circonferenza.*

Le osservazioni fatte nel paragrafo precedente consistono nell'affermazione che i due insiemi $\mathcal{U}_n$ e $\mathbb{Z}_n$ hanno la stessa struttura algebrica (sono *isomorfi*: non soltanto c'è un omomorfismo fra $\mathcal{U}_n$ e $\mathbb{Z}_n$, ma questo è anche biiettivo), nel senso che, dati z_k, $z_j \in \mathcal{U}_n$ si ha

$$z_k \cdot z_j = z_s \qquad \Longleftrightarrow \qquad r_{k+j} = r_s.$$

Conviene ora cambiare linguaggio per vedere gli stessi concetti dal punto di vista che risulta piú utile per la crittografia: fra l'altro, questo nuovo linguaggio permette di dimostrare molto facilmente tutte le proprietà enunciate qui sopra.

L'Aritmetica modulo n: le Congruenze

Fissato $n \in \mathbb{N}^*$, due interi a, $b \in \mathbb{Z}$ si dicono *congrui modulo* n se n divide $a - b$. In questo caso scriviamo

$$n \mid a - b \qquad \text{oppure} \qquad a \equiv b \mod n.$$

Se $x \in \mathbb{Z}$, si dice *minimo residuo positivo* di x modulo n l'intero a tale che $a \in \{0, \dots, n-1\} = \mathbb{Z}_n$ ed $x \equiv a \mod n$, e lo si indica con $x \mod n$. Per esempio:

$$2 \equiv 12 \equiv 22 \equiv \cdots \equiv -8 \equiv -18 \equiv \dots \mod 10$$
$$2 \equiv 2 + 10n \mod 10 \quad \text{per ogni } n \in \mathbb{Z}$$
$$-8 = (-1) \cdot 10 + 2$$

La notazione $\equiv$ è dovuta a Gauss, e ricorda che la congruenza è un'uguaglianza a meno di multipli di n; in un certo senso, un'uguaglianza approssimata. Si può anche dire che $a \equiv b \mod n$ se a e b hanno lo stesso resto della divisione per n, dove il resto della divisione di a per n è l'unico intero r tale che

$$\begin{cases} n \mid a - r, \\ r \in \{0, 1, 2, \dots, n-1\}, \end{cases} \qquad \text{cioè} \qquad a = qn + r \qquad \text{dove} \qquad q = \left[\frac{a}{n}\right].$$

Proprietà delle congruenze. La relazione di congruenza è una relazione di equivalenza. Inoltre, per ogni $c \in \mathbb{Z}$ si ha

$$a \equiv b \mod n \qquad \Longrightarrow \qquad a + c \equiv b + c \mod n \qquad \text{e} \qquad ac \equiv bc \mod n.$$

In particolare, iterando si ottiene che se $a \equiv b \mod n$ allora $a^m \equiv b^m \mod n$ per ogni $m \in \mathbb{N}$. Queste proprietà permettono di dare una dimostrazione pressoché immediata dei cosiddetti "criteri di divisibilità" per 9 e per 11. Sia

$$n = \sum_{j=0}^{k} c_j 10^j \qquad \text{dove } c_j \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

la rappresentazione decimale di n. Poiché $10 \equiv 1 \mod 9$ e $10 \equiv -1 \mod 11$, si ha

$$n \equiv \sum_{j=0}^{k} c_j \mod 9; \qquad n \equiv \sum_{j=0}^{k} (-1)^j c_j \mod 11,$$

e quindi n è divisibile per 9 se e solo se lo è la somma delle sue cifre decimali, mentre è divisibile per 11 se e solo se lo è la somma a segno alterno delle sue cifre decimali.

Teorema Cinese del Resto. *Se n_1, $n_2 \in \mathbb{Z}^*$ ed $(n_1, n_2) = 1$, il sistema*

$$\begin{cases} x \equiv a_1 \mod n_1, \\ x \equiv a_2 \mod n_2, \end{cases}$$

ha un'unica soluzione $\mod n_1 n_2$.

Per esempio, il sistema

$$\begin{cases} x \equiv 7 \mod 10 \\ x \equiv 3 \mod 21 \end{cases} \qquad \text{ha la soluzione } x \equiv 87 \mod 210.$$

Questo significa che due congruenze sono sempre *compatibili* se $(n_1, n_2) = 1$, mentre possono essere incompatibili se $(n_1, n_2) > 1$, come mostrano gli esempi che seguono:

$$\begin{cases} x \equiv 2 \mod 10 \\ x \equiv 0 \mod 4 \end{cases} \Rightarrow x \equiv 12 \mod 20; \qquad \begin{cases} x \equiv 2 \mod 10 \\ x \equiv 1 \mod 4 \end{cases} \text{ è impossibile.}$$

Resta aperta la questione delle equazioni del tipo $ax \equiv b \mod n$ e quella della legge di cancellazione del prodotto: in altre parole, sotto quali condizioni si ha che

$$ac \equiv bc \mod n \qquad \Longrightarrow \qquad a \equiv b \mod n \ ?$$

Vediamo facilmente con esempi che le risposte non sono ovvie: l'equazione $2x \equiv 1 \mod 10$ evidentemente non ha soluzione, mentre $2x \equiv 2 \mod 10$ ha le *due* soluzioni $x_1 \equiv 1 \mod 10$ ed $x_2 \equiv 6 \mod 10$ (piú semplicemente, $x \equiv 1 \mod 5$). Invece $3x \equiv 1 \mod 10$ ha l'unica soluzione $x \equiv 7 \mod 10$. Entrambe le domande sono legate alla possibilità di effettuare una divisione, cioè al calcolo dell'*inverso* di un elemento di $\mathbb{Z}_n$, sempre che questo esista. Per poter risolvere questi problemi, ma anche per dimostrare il Teorema Cinese del Resto, facciamo un passo indietro per introdurre i concetti fondamentali dell'Aritmetica: come si vedrà, tutto si basa su un semplice, ma fondamentale, Teorema di Euclide.

L'Algoritmo di Euclide

In questo paragrafo abbandoniamo temporaneamente lo stile discorsivo adottato finora per dare delle definizioni e dimostrazioni formali.

Teorema (Euclide). *Dati n, $m \in \mathbb{Z}$ si ponga $\mathcal{A}(n,m) \stackrel{\text{def}}{=} \{an + bm : a, b \in \mathbb{Z}\}$ e $d \stackrel{\text{def}}{=} (n,m)$. Si ha che $\mathcal{A} = d\mathbb{Z}$, l'insieme dei multipli interi di d, e dunque esistono λ, $\mu \in \mathbb{Z}$ tali che $d = \lambda n + \mu m$.*

Dim. È evidente che d divide ogni elemento di $\mathcal{A}$. Sia $\delta = \lambda n + \mu m$ il minimo elemento positivo di $\mathcal{A}$ (che esiste purché almeno uno fra n e m sia non nullo). Poiché $d \mid \delta$, resta da dimostrare che $\delta \mid d$. Consideriamo il resto r della divisione di n per δ (cioè l'intero r tale che $0 \leq r < \delta$ ed inoltre esiste $q \in \mathbb{Z}$ tale che $n = q\delta + r$). È chiaro che $r \in \mathcal{A}$ (poiché $r = (1 - \lambda q)n - \mu q m$) e dunque $r = 0$ (poiché altrimenti esisterebbe un elemento positivo di $\mathcal{A}$ strettamente minore di δ), cioè $\delta \mid n$. Analogamente $\delta \mid m$, e quindi $\delta \mid d$, da cui $\delta = d$. □

Per esempio, $(17, 13) = 1$ ed esistono (non unici) interi λ e μ tali che $17\lambda + 13\mu = 1$: infatti $-3 \cdot 17 + 4 \cdot 13 = 1$. Piú avanti vedremo l'Algoritmo di Euclide vero e proprio che ci permette di determinare sia $d = (n,m)$ che due interi λ e μ tali che $d = \lambda n + \mu m$. Per il momento osserviamo che se $d = 1$ questa relazione implica

$$\begin{aligned} \lambda n \equiv 1 \mod m \qquad &\text{cioè} \qquad \lambda \equiv n^{-1} \mod m \\ \mu m \equiv 1 \mod n \qquad &\text{cioè} \qquad \mu \equiv m^{-1} \mod n \end{aligned}$$

Dunque, $13^{-1} \equiv 4 \mod 17$. Una conseguenza di questo fatto è il

Corollario. *Dato $a \in \mathbb{Z}_n$, se $(a,n) = 1$ allora l'applicazione $f_a : \mathbb{Z}_n \to \mathbb{Z}_n$ definita da $f_a(x) \stackrel{\text{def}}{=} ax \mod n$ è una biiezione, con inversa $f_{a^{-1}}$.*

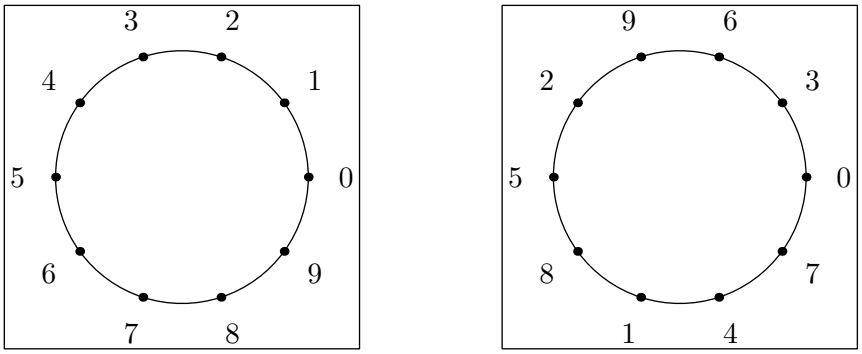


Figura 3. *Il gruppo* $\mathbb{Z}_{10}$ *è generato da* $g_1 = 1$, $g_2 = 3$, $g_3 = 7 = -g_2$, $g_4 = 9 = -g_1$, *ed è* isomorfo *al gruppo delle rotazioni del piano multiple di una rotazione di* $\alpha = \frac{2\pi}{10}$.

Per esempio, l'applicazione $n \mapsto 7n \mod 10$ è una biiezione di $\mathbb{Z}_{10}$:

n	0	1	2	3	4	5	6	7	8	9
$7n$	0	7	4	1	8	5	2	9	6	3
		$*$		$*$				$*$		$*$

Gli asterischi nell'ultima riga indicano gli interi n tali che $(n, 10) = 1$. Si noti che, invece, l'applicazione $n \mapsto 4n$ non è una biiezione di $\mathbb{Z}_{10}$: infatti $4 \cdot 0 \equiv 4 \cdot 5 \mod 10$, anche se $0 \not\equiv 5 \mod 10$. L'insieme degli elementi di $\mathbb{Z}_n$ che possiedono inverso moltiplicativo si indica con $\mathbb{Z}_n^*$ e la cardinalità di $\mathbb{Z}_n^*$ con $\varphi(n)$, detta funzione di Eulero. Per esempio, si ha $\mathbb{Z}_{10}^* = \{1, 3, 7, 9\}$, e quindi $\varphi(10) = 4$.

Un'altra conseguenza importante riguarda le equazioni lineari. Se $(a, n) = 1$ l'equazione $ax \equiv b \mod n$ è risolubile, con soluzione $x \equiv a^{-1}b \mod n$. Se invece $(a, n) = d > 1$ l'equazione $ax \equiv b \mod n$ è risolubile se e solo se $d \mid b$ ed in questo caso è equivalente a $\dfrac{a}{d}x \equiv \dfrac{b}{d} \mod \dfrac{n}{d}$.

Definizione. *Un intero* $n \geq 2$ *si dice* primo *se* $d \mid n$ *implica* $|d| = 1$ *oppure* $|d| = n$.

Corollario (Euclide). *Se* p *è un numero primo e* $p \mid ab$, *allora* $p \mid a$ *oppure* $p \mid b$.

Dim. Se $p \nmid a$ allora $(a, p) = 1$ e per il Teorema di Euclide esistono interi λ e μ tali che $\lambda p + \mu a = 1$. Moltiplichiamo questa uguaglianza per b ed otteniamo $\lambda pb + \mu ab = b$. Poiché p ne divide il primo membro, deve dividere anche il secondo.□

Riassumiamo quanto abbiamo visto finora: è possibile dotare l'insieme $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$ delle solite operazioni $+$ e $\cdot$ eseguendo l'operazione desiderata in $\mathbb{Z}$ e facendola seguire dal calcolo del resto modulo n. In questo modo $\mathbb{Z}_n$ risulta essere un *anello commutativo*, e cioè ha le stesse proprietà formali di $\mathbb{Z}$, tranne la *caratteristica*: se $n \in \mathbb{N}^*$

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ volte}} = 0 \quad \text{in } \mathbb{Z}_n, \qquad \underbrace{1 + 1 + \cdots + 1}_{n \text{ volte}} \neq 0 \quad \text{in } \mathbb{Z}.$$

Questo significa fra l'altro che non è possibile *ordinare* gli elementi di $\mathbb{Z}_n$ come sono ordinati gli elementi di $\mathbb{Z}$. Un'altra differenza importante fra $\mathbb{Z}$ e $\mathbb{Z}_n$ sta nel fatto che in quest'ultimo insieme non vale necessariamente la legge di annullamento del prodotto: infatti, se n non è primo esistono a, $b \in \mathbb{Z}_n \setminus \{0\}$ tali che $ab \equiv 0 \mod n$.

Basta prendere un qualsiasi divisore a di n, con $1 < a < n$, e $b = \frac{n}{a}$. Questo spiega il curioso fenomeno che abbiamo visto prima a proposito della possibilità di risolvere le congruenze del tipo $ax \equiv b \mod n$:

$$2 \cdot 5 \equiv 0 \mod 10 \quad \text{ma } 2 \not\equiv 0 \mod 10,\ 5 \not\equiv 0 \mod 10.$$
$$2a \equiv 2b \mod 10 \quad \text{implica } a \equiv b \mod 5.$$

Inoltre, $\mathbb{Z}_n$ è anche un *gruppo ciclico*, cioè esiste almeno un elemento $g \in \mathbb{Z}_n$ (detto *generatore*) tale che ogni elemento di $\mathbb{Z}_n$ è un multiplo di g: in effetti $g = 1$ genera $\mathbb{Z}_n$ qualunque sia $n \in \mathbb{N}^*$. Un problema interessante è dunque la determinazione dei generatori di $\mathbb{Z}_n$: non è difficile convincersi del fatto che g genera $\mathbb{Z}_n$ se e solo se $(g, n) = 1$, se e solo se g è *invertibile* modulo n, cioè se e solo se esiste $h \in \mathbb{Z}_n$ tale che $hg \equiv 1 \mod n$. Infatti, se $(g, n) = d > 1$ allora tutti i numeri mg sono divisibili per d e quindi $d \mid (mg + kn)$ per ogni $k \in \mathbb{Z}$: dunque $1 \in \mathbb{Z}_n$ non è della forma mg e cioè g non è un generatore. Per esempio, in $\mathbb{Z}_{10}$ i multipli di $g = 2$ sono 0, 2, 4, 6, 8, 0, 2, ... Viceversa, se $(g, n) = 1$ allora esiste $h \in \mathbb{Z}_n$ tale che $hg \equiv 1 \mod n$, e quindi, qualunque si $r \in \mathbb{Z}_n$ si ha $(hr)g \equiv r \mod n$, e cioè r è un multiplo di g. Per esempio, preso $g = 7$ in $\mathbb{Z}_{10}^*$, si trova $h = 3$: se si vuole trovare per quale $m \in \mathbb{Z}_{10}$ si ha $mg \equiv 4 \mod 10$ basta prendere $m = 4h \equiv 2 \mod 10$. Si veda la tabella qui sopra.

Ecco ora un fatto di importanza centrale:

Osservazione. *$\mathbb{Z}_n^* = \mathbb{Z}_n \setminus \{0\}$ se e solo se n è un numero primo.*

In questo caso (e solo in questo caso) tutti gli elementi non nulli di $\mathbb{Z}_n$ hanno inverso moltiplicativo e, di conseguenza, $\mathbb{Z}_n$ ha molte delle proprietà formali di $\mathbb{R}$ o $\mathbb{C}$, cioè è un *campo*: in particolare, in questo caso ogni *equazione polinomiale* $q(x) \equiv 0 \mod p$ ha un numero di radici che non supera il grado di q. In un *campo* K, se il polinomio $q(x)$ di grado k ha le radici $\alpha_1, \ldots, \alpha_k \in K$, allora $q(x) = a(x - \alpha_1) \cdots (x - \alpha_k)$, dove $a \in K \setminus \{0\}$. La spiegazione è semplice: se $q(\alpha_1) = 0$ allora q è divisibile per $x - \alpha_1$ ("Regola di Ruffini"); ripetendo questo procedimento per $\alpha_2, \ldots, \alpha_k$, si ottiene il risultato. Questa scomposizione in fattori è valida perché possiamo effettuare le operazioni di addizione e moltiplicazione (ed è quindi valida in $\mathbb{Z}_n$ anche quando n non è primo): ma solo quando n è primo dalla congruenza $(x - \alpha_1) \cdots (x - \alpha_k) \equiv 0 \mod n$ segue che x è uno degli α_j, per la legge di annullamento del prodotto. Questa cosa è fondamentale per dimostrare che esiste un generatore di $\mathbb{Z}_p^*$ ed è invece *falsa* in $\mathbb{Z}_n$ se n non è un numero primo.

Vediamo un esempio concreto (ed importante): consideriamo l'equazione $x^2 - 1 \equiv 0 \mod n$. Naturalmente $x^2 - 1 = (x - 1)(x + 1)$: se $n = p$, un numero primo, dal fatto che $(x-1)(x+1) \equiv 0 \mod p$ segue che $p \mid (x-1)$ oppure $p \mid (x+1)$ (Corollario del Teorema di Euclide), cioè $x \equiv 1 \mod p$ oppure $x \equiv -1 \mod p$. Invece, se n non è primo, in $\mathbb{Z}_n$ non vale la legge di annullamento del prodotto e quindi non possiamo trarre la stessa conclusione (l'equazione ha comunque le radici ± 1: il punto è che ce ne sono anche altre!). Si osservi che l'equazione $x^2 - 1 \equiv 0 \mod 8$ ha 4 radici distinte (± 1, ± 3), mentre $x^2 - 1 \equiv 0 \mod 24$ ne ha 8 (± 1, ± 3, ± 5, ± 7). In generale, posto $r(2) = 1$, $r(4) = 2$, $r\left(2^\alpha\right) = 4$ per $\alpha \geq 3$, $r\left(p^\alpha\right) = 2$ per p primo dispari ed $\alpha \geq 1$, allora l'equazione $x^2 \equiv 1 \mod p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ ha $r\left(p_1^{\alpha_1}\right) \cdots r\left(p_k^{\alpha_k}\right)$ soluzioni, per il Teorema Cinese del Resto.

Proprietà aritmetiche dei numeri primi

A questo punto è opportuno inserire qualche semplice proprietà dei numeri primi.

Definizione. *Dato $n \in \mathbb{N}^*$ chiamiamo* forma canonica di n *la decomposizione*

$$n = \prod_{i=1}^{k} p_i^{\alpha_i}, \qquad \textit{dove } p_i < p_j \textit{ se } i < j,\ \alpha_i \in \mathbb{N}^* \textit{ per } i = 1, \dots, k,$$

ed i p_i sono numeri primi. Se $n = 1$ il prodotto è vuoto.

Teorema di Fattorizzazione Unica. *Ogni $n \in \mathbb{N}^*$ ha un'unica forma canonica.*

Dim. Sia $n \geq 2$ il piú piccolo numero naturale con due forme canoniche diverse

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} = \prod_{j=1}^{l} q_j^{\beta_j},$$

con le convenzioni della definizione. Per il Corollario qui sopra, se $p_1 \mid n$ allora p_1 è uno dei primi q_j, ed analogamente q_1 è uno dei primi p_i e dunque $p_1 = q_1$ (poiché entrambi sono uguali al piú piccolo fattore primo di n). Quindi anche il numero $n/p_1 = n/q_1 < n$ ha due forme canoniche distinte, contro la minimalità di n. □

Teorema (Euclide). *Esistono infiniti numeri primi.*

Dim. Sia $\{p_1, \dots, p_n\}$ un qualunque insieme finito non vuoto di numeri primi. Il numero $N \overset{\text{def}}{=} p_1 \cdots p_n + 1 > 1$ non è divisibile per alcuno dei primi $p_1, \dots, p_n$. □

Teorema (Wilson). *$n \geq 2$ è primo se e solo se $(n-1)! \equiv -1 \mod n$.*

Dim. Ci limitiamo ad illustrare la dimostrazione per mezzo di un esempio: se $n = 11$ allora si ha

$$\begin{aligned}(11-1)! &\equiv 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \\ &\equiv 1 \cdot (2 \cdot 6) \cdot (3 \cdot 4) \cdot (5 \cdot 9) \cdot (7 \cdot 8) \cdot 10 \\ &\equiv 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 10 \\ &\equiv -1 \mod 11\end{aligned}$$

Possiamo associare ciascun fattore del prodotto 10! (tranne 1 e 10) con il suo reciproco modulo 11, poiché l'equazione $x \equiv x^{-1} \mod 11$ ha due sole soluzioni, $x \equiv 1 \mod 11$ ed $x \equiv -1 \equiv 10 \mod 11$. Osserviamo che la stessa cosa non è vera se n non è primo: per esempio, $(10-1)! \equiv 0 \mod 10$, poiché $10 = 2 \cdot 5$ e questi sono fattori in 9!; in effetti, 2 e 5 non sono invertibili modulo 10. □

Teorema (Fermat). *Se p è un numero primo e $p \nmid a$, allora $a^{p-1} \equiv 1 \mod p$.*

Questo risultato garantisce l'esistenza di $p-1$ soluzioni distinte dell'equazione $x^{p-1} \equiv 1 \mod p$, e si usa nella dimostrazione del Teorema di Gauss qui sotto.

Dim. Anche in questo caso ci limitiamo ad un esempio. Per dimostrare che $10^6 \equiv 1 \mod 7$, ricordiamo che l'applicazione $n \mapsto 10n \mod 7$ è una biiezione di $\mathbb{Z}_7^*$:

n	1	2	3	4	5	6
$10n$	10	20	30	40	50	60
$10n \mod 7$	3	6	2	5	1	4

Moltiplicando i numeri sulla prima riga della tabella (o sulla terza) troviamo 6!, moltiplicando quelli sulla seconda troviamo $10^6 \cdot 6!$ e quindi $10^6 \cdot 6! \equiv 6! \mod 7$ ed il Teorema di Fermat segue dal Teorema di Wilson. □

Osserviamo che la congruenza $10^6 \equiv 1 \mod 7$ in realtà è equivalente alla periodicità dello sviluppo decimale di $\frac{1}{7}$: infatti

$$\frac{1}{7} = 0.\overline{142857} = \frac{142857}{999999} \quad \Longleftrightarrow \quad 7 \mid 999999 = 10^6 - 1.$$

Notiamo per inciso che la seconda uguaglianza a sinistra dipende dal calcolo della somma della serie geometrica di ragione $x = 10^{-6}$. In effetti è possibile dimostrare il Teorema di Fermat in generale sfruttando questo fatto, ma quella qui sopra è una dimostrazione piú semplice. Osserviamo che l'ordine di 10 modulo p (per $p \neq 2$, 5) non è altro che il periodo della frazione decimale $\frac{1}{p}$: infatti, sappiamo che le cifre decimali iniziano a ripetersi non appena troviamo di nuovo il resto 1, come ci ricorda il calcolo qui sotto.

$$\begin{array}{lll}
10^0 \equiv 1 \mod 7 & 10^0 = & 0 \cdot 7 + 1 \\
10^1 \equiv 3 \mod 7 & 10^1 = & 1 \cdot 7 + 3 \\
10^2 \equiv 2 \mod 7 & 10^2 = & 14 \cdot 7 + 2 \\
10^3 \equiv 6 \mod 7 & 10^3 = & 142 \cdot 7 + 6 \\
10^4 \equiv 4 \mod 7 & 10^4 = & 1428 \cdot 7 + 4 \\
10^5 \equiv 5 \mod 7 & 10^5 = & 14285 \cdot 7 + 5 \\
10^6 \equiv 1 \mod 7 & 10^6 = & 142857 \cdot 7 + 1
\end{array}$$

$$\begin{array}{r|l}
1\ 0 & 7 \\
\cline{2-2}
3\ 0 & \\
2\ 0 & 0.142857 \\
6\ 0 & \\
4\ 0 & \\
5\ 0 & \\
1 &
\end{array}$$

Il Teorema seguente è di importanza fondamentale perché ci dice che la struttura di $\mathbb{Z}_p^*$ è particolarmente semplice quando p è un numero primo.

Teorema (Gauss). *Se p è un numero primo, allora $\mathbb{Z}_p^*$ è un gruppo moltiplicativo ciclico, cioè esiste $g = g_p \in \mathbb{Z}_p^*$ tale che ogni elemento di $\mathbb{Z}_p^*$ è una potenza di g_p.*

La dimostrazione non è semplice: prima illustreremo questo risultato in un caso particolare, e poi daremo qualche indicazione su come utilizzare le idee qui esposte per dare la dimostrazione vera e propria. Consideriamo il numero primo $p = 11$: possiamo calcolare a mano le potenze successive dei suoi elementi: soltanto in 4 casi accade che queste potenze assumano tutti i valori possibili modulo 11. Questo fatto è illustrato dalla Tavola 4 e dalle Figure 5 e 7. In altre parole, il gruppo moltiplicativo $\mathbb{Z}_{11}^*$ è generato da $g_1 = 2$, $g_2 = 6 = g_1^{-1}$, $g_3 = 7 = g_1^7$, $g_4 = 8 = g_3^{-1} = g_1^3$, ed è *isomorfo* al gruppo additivo $\mathbb{Z}_{10}$. Si osservi che la struttura dei gruppi moltiplicativi $\mathbb{Z}_n$ quando n non è primo è in generale molto piú complessa: ci limitiamo a dare due figure relative ai casi $n = 8$ ed $n = 24$.

Per dimostrare il Teorema di Gauss abbiamo bisogno del Teorema di Fermat che garantisce che l'equazione $x^{p-1} \equiv 1 \mod p$ ha $p - 1$ radici distinte (cioè tutti gli elementi di $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$). Un altro ingrediente fondamentale è l'identità

$$n = \sum_{d \mid n} \varphi(d) \tag{2}$$

h	0	1	2	3	4	5	6	7	8	9	10
2^h	1	2	4	8	5	10	9	7	3	6	1
6^h	1	6	3	7	9	10	5	8	4	2	1
7^h	1	7	5	2	3	10	4	6	9	8	1
8^h	1	8	9	6	4	10	3	2	5	7	1
		$*$		$*$				$*$		$*$	

Tavola 4. *Le potenze successive dei generatori di* $\mathbb{Z}_{11}^*$*, indicati da* $*$ *nell'ultima riga. I generatori compaiono in corrispondenza degli esponenti* h *che sono primi con 10, e cioè dei generatori di* $\mathbb{Z}_{10}$*. Le potenze dei generatori sono periodiche con periodo* $10 = \varphi(11)$*.*

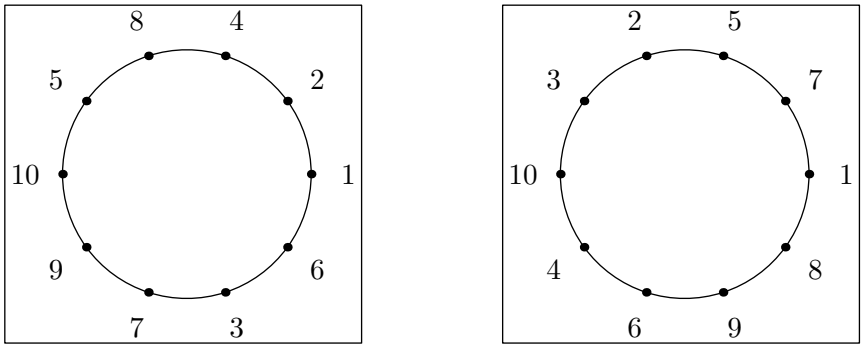


Figura 5. *Il gruppo* $\mathbb{Z}_{11}^*$ *è generato da* $g_1 = 2$*,* $g_2 = 6 = g_1^{-1}$*,* $g_3 = 7 = g_1^7$*,* $g_4 = 8 = g_3^{-1} = g_1^3$*, ed è* isomorfo *al gruppo* $\mathbb{Z}_{10}$*. I generatori hanno ordine massimo, cioè* 10*.*

che si dimostra confrontando le cardinalità degli insiemi

$$\left\{ \frac{h}{n} \colon h \in \{1, \dots, n\} \right\} = \bigcup_{d \mid n} \left\{ \frac{a}{d} \colon a \in \{1, \dots, d\} \text{ e } (a,d) = 1 \right\}.$$

A sinistra ci sono tutte le frazioni con denominatore n e numeratore $\in \{1, \dots, n\}$, a destra ci sono le stesse frazioni ridotte ai minimi termini. Per esempio, quando $n = 10$, d può avere i valori 1, 2, 5, 10 e quindi si ha

$$\left\{ \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}, \frac{5}{10}, \frac{6}{10}, \frac{7}{10}, \frac{8}{10}, \frac{9}{10}, \frac{10}{10} \right\} = \\ = \left\{ \frac{1}{1} \right\} \bigcup \left\{ \frac{1}{2} \right\} \bigcup \left\{ \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5} \right\} \bigcup \left\{ \frac{1}{10}, \frac{3}{10}, \frac{7}{10}, \frac{9}{10} \right\}$$

Osserviamo per inciso che la relazione (2) può essere utilizzata per calcolare esplicitamente il valore di $\varphi(n)$ qualunque sia n: si dimostra che $\varphi(nm) = \varphi(n)\varphi(m)$ se $(n,m) = 1$ e che $\varphi\big(p^\alpha\big) = p^{\alpha-1}(p-1)$ se p è primo ed $\alpha \geq 1$. Dunque $\varphi\big(p_1^{\alpha_1} \cdots p_k^{\alpha_k}\big) = (p_1 - 1)p_1^{\alpha_1 - 1} \cdots (p_k - 1)p_k^{\alpha_k - 1}$ se i p_j sono primi distinti.

Il punto cruciale della dimostrazione del teorema di Gauss è che per $d \mid \varphi(n)$ l'equazione $x^d \equiv 1 \mod n$ ha esattamente d soluzioni, di cui $\varphi(d)$ *primitive*, cioè soluzioni che hanno *ordine* esattamente d. Nel caso $p = 11$ questo fatto è illustrato dalla Tavola 8, in cui le soluzioni dell'equazione $x^{10} \equiv 1 \mod 11$ sono classificate secondo il loro ordine.

Osserviamo che se g genera $\mathbb{Z}_p^*$ e $x_1 = g^h$, allora x_1 ha ordine $d = (p-1)/(h, p-1)$. In altre parole, se g genera $\mathbb{Z}_p^*$, allora g^h genera $\mathbb{Z}_p^*$ se e solo se $(h, p-1) = 1$. Per

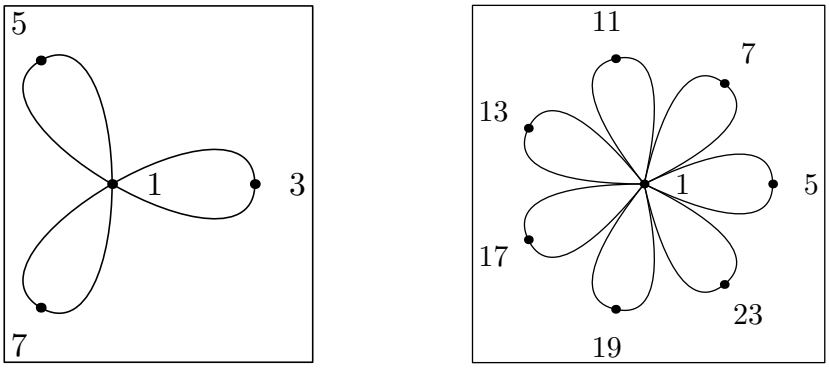


Figura 6. *Come nella Figura 5, gli archi connettono le potenze successive dello stesso elemento; ogni elemento* $x \in \mathbb{Z}_8^*$ *o* $\in \mathbb{Z}_{24}^*$ *soddisfa* $x^2 = 1$ *(e quindi ha ordine* 1 *o* 2*): dunque le sue potenze successive sono* $1, x, 1, x, \ldots$

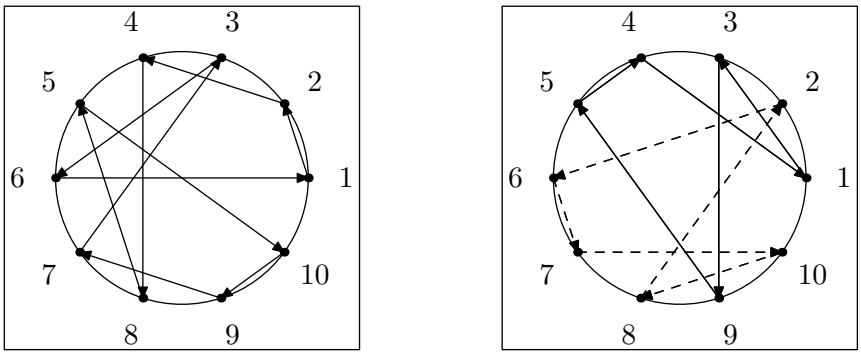


Figura 7. *A sinistra, le potenze di* 2 *esauriscono gli elementi di* $\mathbb{Z}_{11}^*$*; a destra, le potenze di* 3 *toccano solo metà degli elementi di* $\mathbb{Z}_{11}^*$*. L'altra metà si ottiene considerando la successione* $2 \cdot 3^h$*.*

esempio, sappiamo che 2 genera $\mathbb{Z}_{11}^*$ e che $2^{10} \equiv 1 \mod 11$ per il Teorema di Fermat. Vogliamo vedere che 2^r genera $\mathbb{Z}_{11}^*$ se e solo se r è invertibile modulo 10. Infatti, se r è invertibile modulo 10 allora l'applicazione $x \mapsto rx \mod 10$ è una biiezione, e quindi gli esponenti r, $2r \mod 10$, $3r \mod 10$, $\ldots$, $9r \mod 10$, $10r \mod 10$ sono, in un ordine diverso, gli interi $0, 1, 2, \ldots, 8, 9$. Per esempio, se $r = 3$:

$$2^3 \equiv 8 \qquad 2^6 \equiv 9 \qquad 2^9 \equiv 6 \qquad 2^{12} \equiv 2^2 \equiv 4 \qquad 2^{15} \equiv 2^5 \equiv 10$$
$$2^{18} \equiv 2^8 \equiv 3 \qquad 2^{21} \equiv 2^1 \equiv 2 \qquad 2^{24} \equiv 2^4 \equiv 5 \qquad 2^{27} \equiv 2^7 \equiv 7 \qquad 2^{30} \equiv 2^0 \equiv 1$$

Abbiamo detto che $\mathbb{Z}_p^*$ è *isomorfo* a $\mathbb{Z}_{p-1}$ e che la corrispondenza fra i due è l'analogo del logaritmo: per maggiore chiarezza vediamo questa cosa in dettaglio quando $p = 11$. Si faccia riferimento di nuovo alla seconda riga della Tavola 4.

$\mathbb{Z}_{10}$ è un gruppo additivo ciclico con generatori 1, 3, 7, 9

$\mathbb{Z}_{11}^*$ è un gruppo moltiplicativo ciclico con generatori 2^1, 2^3, 2^7, 2^9

Dunque, per esempio

$$\begin{aligned} &\text{in } \mathbb{Z}_{10} \text{ si ha} && 6 + 8 \equiv 4 \mod 10 \\ &\text{in } \mathbb{Z}_{11}^* \text{ si ha} && 2^6 \cdot 2^8 \equiv 2^4 \mod 11 \end{aligned}$$

Infatti, $2^6 \equiv 9 \mod 11$, $2^8 \equiv 3 \mod 11$, $2^{14} \equiv 5 \mod 11$ e $9 \cdot 3 \equiv 5 \mod 11$. In definitiva, moltiplicare elementi di $\mathbb{Z}_{11}^*$ corrisponde a sommare elementi di $\mathbb{Z}_{10}$ (il loro *logaritmo discreto* in base 2).

Equazione	Soluzioni	primitive	h
$x \equiv 1 \mod 11$	$x = 1$	1	0
$x^2 \equiv 1 \mod 11$	$x = 1, 10$	10	5
$x^5 \equiv 1 \mod 11$	$x = 1, 3, 4, 5, 9$	3, 4, 5, 9	8, 2, 4, 6
$x^{10} \equiv 1 \mod 11$	$x = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$	2, 6, 7, 8	1, 9, 7, 3

Tavola 8. *Le soluzioni di* $x^{10} \equiv 1 \mod 11$ *classificate secondo il loro ordine: in questo caso* $p = 11$, $g = 2$. *All'estrema destra sono indicati i valori di* h *corrispondenti alle soluzioni primitive: si vedano gli esponenti di* 2 *nella seconda riga della Tavola 4. Si noti che detto* d *il grado dell'equazione all'estrema sinistra, si ha* $d = 10/(10, h)$ *per le soluzioni primitive.*

Si può generalizzare il Teorema di Fermat al caso in cui l'esponente non è primo:

Teorema (Eulero). *Se* $n \geq 2$ *ed* $(a, n) = 1$ *allora si ha* $a^{\varphi(n)} \equiv 1 \mod n$.

La dimostrazione è analoga a quella del Teorema di Fermat.

Teorema (Gauss). *Il gruppo moltiplicativo* $\mathbb{Z}_n^*$ *è ciclico per* $n = 1, 2, 4$, *e per* $n = p^\alpha$, $2p^\alpha$, *dove* p *è un numero primo dispari ed* $\alpha \geq 1$.

La dimostrazione nel caso p^α sfrutta il fatto che esiste g_p che genera $\mathbb{Z}_p^*$, per costruire un intero g_p^* che genera *tutti* i gruppi $\mathbb{Z}_{p^\alpha}^*$.

Numeri pseudocasuali. Facciamo una breve digressione per vedere un'applicazione pratica di queste idee: se g genera $\mathbb{Z}_p^*$ allora per $n = 1, \dots, p-1$, i numeri $g^n \mod p$ coincidono con i numeri $1, 2, \dots, p-1$, in un altro ordine. L'applicazione $n \mapsto \big((g^n \mod p) - 1\big)/(p-1)$ dà quindi una successione periodica di periodo $p-1$ di numeri razionali nell'intervallo $[0, 1)$, sostanzialmente imprevedibile. Questo fatto viene sfruttato dai programmatori per costruire in modo piuttosto semplice delle funzioni che generano numeri pseudocasuali: per esempio, sapendo che 75 genera $\mathbb{Z}_{65537}^*$ si ottiene una successione di periodo $65536 = 2^{16}$. Si osservi inoltre che non è necessario calcolare ogni volta $g^n \mod p$, ma è sufficiente memorizzare $x = g^{n-1} \mod p$ e poi calcolare $(gx) \mod p$. Inoltre, se si vuole avere un valore iniziale diverso da 1, dato il *seme* m si può partire da $g^m \mod p$.

Problemi. Concludiamo il paragrafo indicando quattro problemi che rimangono aperti, a cui daremo risposta nel Capitolo sugli Algoritmi:

- dato $g \in \mathbb{Z}_n^*$ trovarne l'inverso $h \in \mathbb{Z}_n^*$;
- trovare la soluzione di un sistema di congruenze;
- determinare un generatore g di $\mathbb{Z}_p^*$;
- dato un generatore g di $\mathbb{Z}_p^*$ ed $a \in \mathbb{Z}_p^*$, determinare h in modo che $g^h \equiv a \mod p$ (h è il *logaritmo discreto* di a in $\mathbb{Z}_p^*$ in base g).

Pseudoprimi e numeri di Carmichael

È importante notare che il Teorema di Wilson dà una condizione necessaria e sufficiente affinché n sia primo (ma molto inefficiente, dato che sono necessarie $n - 2$ moltiplicazioni). Il Teorema di Fermat, invece, dà solo una condizione necessaria, ma la verifica corrispondente può essere effettuata in un tempo relativamente breve dato che esiste un algoritmo efficiente per il calcolo delle potenze, come vedremo piú avanti. In altre parole, se vogliamo verificare se n è primo o meno, possiamo calcolare $a^{n-1} \mod n$ per qualche $a \in [2, n-1]$ che sia primo con n (d'altra parte,

$341 = 11 \cdot 31$	$2^{10} \equiv 1\,(341)$	$10 \mid 340$	$561 = 3 \cdot 11 \cdot 17$	$5^{80} \equiv 1\,(561)$	$80 \mid 560$
$561 = 3 \cdot 11 \cdot 17$	$2^{40} \equiv 1\,(561)$	$40 \mid 560$	$35 = 5 \cdot 7$	$6^2 \equiv 1\,(35)$	$2 \mid 34$
$645 = 3 \cdot 5 \cdot 43$	$2^{28} \equiv 1\,(645)$	$28 \mid 644$	$217 = 7 \cdot 31$	$6^6 \equiv 1\,(217)$	$6 \mid 216$
$91 = 7 \cdot 13$	$3^6 \equiv 1\,(91)$	$6 \mid 90$	$25 = 5^2$	$7^4 \equiv 1\,(25)$	$4 \mid 24$
$703 = 19 \cdot 37$	$3^{18} \equiv 1\,(703)$	$18 \mid 702$	$561 = 3 \cdot 11 \cdot 17$	$7^{80} \equiv 1\,(561)$	$80 \mid 560$
$15 = 3 \cdot 5$	$4^2 \equiv 1\,(15)$	$2 \mid 14$	$9 = 3^2$	$8^2 \equiv 1\,(9)$	$2 \mid 8$
$85 = 5 \cdot 17$	$4^8 \equiv 1\,(85)$	$8 \mid 84$	$21 = 3 \cdot 7$	$8^2 \equiv 1\,(21)$	$2 \mid 20$
$561 = 3 \cdot 11 \cdot 17$	$4^{20} \equiv 1\,(561)$	$20 \mid 560$	$45 = 3^2 \cdot 5$	$8^4 \equiv 1\,(45)$	$4 \mid 44$
$217 = 7 \cdot 31$	$5^6 \equiv 1\,(217)$	$6 \mid 216$	$65 = 5 \cdot 13$	$8^4 \equiv 1\,(65)$	$4 \mid 64$

Tavola 9. *Alcuni pseudoprimi nelle basi* $2, \ldots, 8$. *La prima colonna contiene la fattorizzazione dello pseudoprimo, la seconda la congruenza soddisfatta con il minimo esponente possibile. Per motivi di spazio la congruenza* $\alpha \equiv \beta \mod n$ *è stata scritta nella forma alternativa* $\alpha \equiv \beta(n)$.

se troviamo $a \in [2, n-1]$ con $d \stackrel{\text{def}}{=} (a, n) > 1$ abbiamo anche trovato un fattore non banale di n, e cioè d). Il Teorema di Fermat garantisce che se $a^{n-1} \not\equiv 1 \mod n$ allora n è certamente composto, senza produrre esplicitamente fattori di n; ma se, viceversa, $a^{n-1} \equiv 1 \mod n$, non è detto che n sia primo.

Che il Teorema di Fermat dia solo una condizione necessaria può essere visto per mezzo di esempi numerici: in effetti $2^{340} \equiv 1 \mod 341$, ma $341 = 11 \cdot 31$. Possiamo dimostrare questo fatto senza quasi fare calcoli: poiché $2^{10} \equiv 1 \mod 11$ per il Teorema di Fermat e $2^5 = 32 \equiv 1 \mod 31$, si ha $2^{10} \equiv 1 \mod 31$ e quindi per il Teorema Cinese del Resto $2^{10} \equiv 1 \mod 341$ (le due congruenze $x \equiv 1 \mod 11$ ed $x \equiv 1 \mod 31$ sono compatibili e dunque hanno una soluzione simultanea, che evidentemente è $x \equiv 1 \mod 11 \cdot 31$). Ma $10 \mid 340$ e quindi $2^{340} = \big(2^{10}\big)^{34} \equiv 1^{34} \equiv 1 \mod 341$. Queste considerazioni giustificano la necessità della seguente

Definizione. *Diciamo che* $n \in \mathbb{Z}$ *è uno pseudoprimo in base* $a \in \mathbb{N}^*$ *se è composto ed* $a^{n-1} \equiv 1 \mod n$.

Per esempio, qualunque sia $n \geq 2$, $4n^2 - 1$ è pseudoprimo in base $2n$: $(2n)^2 \equiv 1 \mod \big(4n^2 - 1\big)$ e $2 \mid 4n^2 - 2$. Si veda anche la Tavola 9, che nella prima colonna contiene uno pseudoprimo n con la sua fattorizzazione, nella seconda una congruenza del tipo $a^d \equiv 1 \mod n$, dove d ha il minimo valore possibile, e nella terza la verifica che $d \mid n - 1$ e quindi che n è uno pseudoprimo in base a.

Si potrebbe sperare che gli pseudoprimi siano piuttosto rari (magari, fissata la base a, che ce ne sia solo un numero finito). In effetti, però, vale il

Teorema (Cipolla). *Dato* $a \in \mathbb{N}^*$ *esistono infiniti* $n \in \mathbb{N}$ *pseudoprimi in base* a.

Infatti, se $p \nmid a\big(a^2 - 1\big)$ allora $n_p \stackrel{\text{def}}{=} \dfrac{a^{2p} - 1}{a^2 - 1} = \dfrac{a^p - 1}{a - 1} \cdot \dfrac{a^p + 1}{a + 1}$ è uno pseudoprimo in base a. La dimostrazione non è molto difficile, e si basa sul Teorema di Fermat e su alcune identità algebriche.

A questo punto si potrebbe almeno sperare che gli insiemi degli pseudoprimi in base 2 ed in base 3, per esempio, siano disgiunti, ma anche questo è falso: infatti 1105 è simultaneamente pseudoprimo in base 2 ed in base 3. La tabella degli

pseudoprimi riprodotta qui mostra anche che 561 è pseudoprimo in base 2, 4, 5, 7. Non è difficile dimostrare che 561 (ed anche 1105) è uno pseudoprimo in ogni base a tale che $(a, 561) = 1$ (risp. $(a, 1105) = 1$). Infatti, per il Teorema di Fermat, se $(a, 561) = 1$, allora $a^2 \equiv 1 \mod 3$, $a^{10} \equiv 1 \mod 11$ e $a^{16} \equiv 1 \mod 17$; dato che il minimo comune multiplo di 2, 10 e 16 è 80, si ha

$$\begin{cases} a^{80} = \left(a^2\right)^{40} & \equiv 1 \mod 3 \\ a^{80} = \left(a^{10}\right)^{8} & \equiv 1 \mod 11 \\ a^{80} = \left(a^{16}\right)^{5} & \equiv 1 \mod 17 \end{cases} \qquad \Longrightarrow \qquad a^{80} \equiv 1 \mod 3 \cdot 11 \cdot 17,$$

per il Teorema Cinese del Resto. Poiché $80 \mid 560$ si ha anche $a^{560} \equiv 1 \mod 561$, e quindi 561 è uno pseudoprimo in ogni base a tale che $(a, 561) = 1$.

Definizione. *Gli interi n che sono pseudoprimi in tutte le basi a tali che $(a, n) = 1$ si dicono numeri di Carmichael.*

Qui sopra abbiamo dimostrato che 561 è un numero di Carmichael (in effetti è il piú piccolo). I successivi sono 1105, 1729, 2465, 2821, 6601, ... , ed oggi è noto che sono infiniti. Vale la pena di osservare che la nostra dimostrazione del fatto che 561 è un numero di Carmichael può essere estesa in generale per dimostrare il

Criterio di Korselt. *L'intero n è di Carmichael se e solo se è composto, libero da fattori quadrati e $p - 1 \mid n - 1$ per ogni $p \mid n$.*

Dunque se n è di Carmichael allora è dispari ed ha almeno tre fattori primi distinti.

Evidentemente l'esistenza degli pseudoprimi e dei numeri di Carmichael pone un limite alla possibilità di utilizzare il Teorema di Fermat come criterio di primalità, ma non per questo tutto è perduto. Esiste infatti un criterio basato sul vero inverso del Teorema di Fermat: questo garantisce che i numeri che lo soddisfano sono primi a tutti gli effetti, e non semplicemente degli pseudoprimi.

Teorema (Lucas). *Se $a^d \not\equiv 1 \mod n$ per ogni $d \mid n - 1$ tale che $d < n - 1$ ed inoltre $a^{n-1} \equiv 1 \mod n$, allora n è primo.*

La dimostrazione dipende dal fatto che un tale elemento $a \in \mathbb{Z}_n$ ha ordine esattamente $n - 1$ in $\mathbb{Z}_n^*$, e questo può accadere se e solo se n è primo (ricordiamo che l'ordine di a in $\mathbb{Z}_n^*$ divide $\varphi(n)$, e che $\varphi(n) \leq n - 2$ se $n \geq 4$ non è primo).

Il Teorema di Lucas permette di stabilire se l'intero n è primo, ma è necessario conoscere la fattorizzazione completa di $n - 1$. Esistono varianti di questo Teorema che permettono di ottenere lo stesso risultato (in un modo piú complicato) conoscendo solo una fattorizzazione parziale. In qualche caso ci si accontenta di sapere che n è *probabilmente* primo, verificando la condizione di Fermat per un certo numero di basi scelte a caso, e "sperando" di non aver trovato un numero di Carmichael. Si osservi che, come vedremo qui sotto, il calcolo delle potenze modulo n può essere effettuato in modo molto efficiente dal punto di vista computazionale (il numero di iterazioni necessarie è essenzialmente il logaritmo in base 2 dell'esponente), ed in modo che tutti i risultati parziali del calcolo siano $\leq n$ in valore assoluto.

Concludiamo il Capitolo indicando l'esistenza di altri criteri di primalità anch'essi basati essenzialmente sulla struttura di $\mathbb{Z}_m^*$, la cui descrizione ci costringerebbe però ad introdurre nuovi concetti e ad allungare ulteriormente la discussione: il Lettore interessato è invitato a consultare il Capitolo 2 del libro di Ribenboim.

Crittografia

Applicazioni alla Crittografia

Qui assumiamo che siano noti i problemi e le definizioni relative alla crittografia: ci limitiamo a ricordare che il problema principale è lo scambio di informazioni per mezzo di un canale non sicuro, come può essere Internet. Gli utenti di un sistema crittografico devono concordare fra loro l'*alfabeto* in cui sono scritti i messaggi che si scambieranno: per i nostri scopi è sufficiente sapere che ogni messaggio può essere trasformato in una sequenza piú o meno lunga di interi (per esempio, il codice ASCII dei singoli caratteri). Fissiamo dunque un insieme di *messaggi* $\mathfrak{M}$: solitamente $\mathfrak{M} = \mathbb{Z}_N$ dove $N \in \mathbb{N}$ è molto grande (tipicamente al giorno d'oggi $N \approx 2^{512} \approx 10^{154}$). Per noi un messaggio è un elemento di $\mathbb{Z}_N$. Nella pratica, si deve trasformare ogni testo alfanumerico in uno o piú messaggi di questo tipo. Le *funzioni crittografiche* che consideriamo sono biiezioni $f\colon \mathfrak{M} \to \mathfrak{M}$ (nel linguaggio del calcolo combinatorio, *permutazioni* dell'insieme $\mathfrak{M}$). Nelle applicazioni pratiche queste funzioni dipendono da uno o piú parametri, parte dei quali sono tenuti segreti da ciascun utente del sistema, mentre altri sono resi pubblici.

Crittosistemi a chiave pubblica: RSA. L'idea dei crittosistemi a chiave pubblica è semplice: ciascun utente sceglie una funzione crittografica che dipende da alcuni parametri, ma rende noti solo quelli che permettono di codificare i messaggi a lui diretti, mantenendo segreti quelli necessari alla decodifica. In questo modo, chiunque può spedire un messaggio all'utente in questione senza che questo, se intercettato da terzi, possa essere compreso. Vediamo ora come questo possa essere realizzato nella pratica.

Ogni utente (diciamo A) compie le seguenti operazioni una volta sola

- A sceglie due numeri primi grandi p e q;
- calcola $n = p \cdot q$;
- calcola $\varphi(n) = (p-1)(q-1) = n - p - q + 1$;
- sceglie $e \in \mathbb{N}$ tale che $\big(e, \varphi(n)\big) = 1$;
- determina $d \in \mathbb{Z}_n^*$ tale che $e \cdot d \equiv 1 \mod \varphi(n)$;
- rende nota la coppia (n, e).

La *funzione crittografica* di A è

$$f_A(x) = x^e \mod n$$

che può essere calcolata da tutti gli utenti del crittosistema. La funzione che A utilizza per decifrare è

$$f_A^{-1}(y) = y^d \mod n$$

per calcolare la quale è necessario conoscere d, e quindi $\varphi(n)$ e quindi la fattorizzazione di n. La sicurezza di questo sistema dipende in modo essenziale dalla difficoltà di scomporre n nei suoi fattori primi. La conoscenza di p e q permette di determinare d se è noto e e quindi di leggere i messaggi destinati ad A.

A deve tenere segreti p, q e d. L'insieme dei messaggi è $\mathfrak{M} = \mathbb{Z}_n$. Chi voglia inviare un messaggio $M \in \mathfrak{M}$ ad A calcola $C = f_A(M) = M^e \mod n$ e lo trasmette. Per leggere il messaggio originale, A calcola $f_A^{-1}(C) = C^d \mod n$: infatti $C^d \equiv \big(M^e\big)^d \equiv M^{ed} \equiv M \mod n$ per il Teorema di Eulero.

Testo				M	$C = M^e \mod n$
M	Y	␣	M	346482	888745
I	S	T	R	232787	1201313
E	S	S	’	124768	1174612
␣	E	Y	E	787324	636449
S	␣	A	R	512117	227442
E	␣	N	O	134504	1999438
T	H	I	N	519553	483208
G	␣	L	I	188438	983073
K	E	␣	T	274489	1326351
H	E	␣	S	193488	151797
U	N	.	␣	552539	1507154

Tavola 10. *Codifica del messaggio “*MY␣MISTRESS’␣EYES␣ARE␣NOTHING␣LIKE␣THE␣SUN.␣*” per mezzo dell’alfabeto “*ABCDEFGHIJKLMNOPQRSTUVWXYZ,.’␣*” Il testo viene convertito in un equivalente numerico M: la stringa “*ABCD*” viene interpretata come il numero in base* 30 *dato da* $\mathtt{A}\cdot 30^3 + \mathtt{B}\cdot 30^2 + \mathtt{C}\cdot 30 + \mathtt{D}$*, e poi ad* A *viene assegnato il valore* 0*, a* B *il valore* 1*, e cosí via, dove* ␣ *sta per lo spazio ed ha equivalente numerico* 29*. Inoltre sono stati scelti i seguenti valori dei parametri:* $p = 1069$*,* $q = 1973$*,* $n = pq = 2109137$*,* $\varphi(n) = 2106096$*,* $e = 10001$*,* $d \equiv e^{-1} \mod \varphi(n) = 40433$*.*

Esempio pratico. La Tavola 10 illustra un esempio pratico di applicazione delle idee descritte sopra, con la codifica di un breve messaggio; il testo viene prima convertito in un equivalente numerico.

Esercizio. Per esercizio, si chiede di decifrare il messaggio qui sotto sapendo che è stato cifrato con la tecnica e con l’alfabeto descritti sopra, e che la chiave pubblica utilizzata è $(n, e) = (2109137, 10001)$. Il messaggio da decifrare è

$$744567,\ 1726777,\ 1556755,\ 957672,\ 689457,\ 858349,\ 866725.$$

In pratica, bisogna scomporre n nei suoi fattori primi p e q, determinare $\varphi(n) = n - p - q + 1$, determinare $d \equiv e^{-1} \mod \varphi(n)$ e poi calcolare $C^d \mod n$, dove C è ciascuno dei numeri qui sopra. Infine, si devono ricavare gli equivalenti alfabetici dei numeri cosí trovati.

Firma digitale: certificazione dell’identità

Un altro problema di fondamentale importanza nella comunicazione fra soggetti distanti è la certificazione dell’identità. In altre parole, ogni utente di un crittosistema ha bisogno non solo di sapere che i messaggi a lui destinati non possono essere decifrati da altri, ma anche che chi scrive sia realmente chi dice di essere. Supponiamo dunque che l’utente A, con chiave pubblica (n_A, e_A) e funzione crittografica f_A voglia convincere della propria identità l’utente B, con chiave pubblica (n_B, e_B) e funzione crittografica f_B. Per raggiungere questo scopo, l’utente A sceglie una “firma digitale” s_A che rende pubblica: in pratica A sceglie $s_A \in \mathbb{Z}_{n_A}$. Per convincere B della propria identità, in calce al proprio messaggio invia una forma crittografata della firma, e precisamente

$$m_A = f_B\big(f_A^{-1}(s_A)\big) \quad \text{se } n_A < n_B; \qquad m_A = f_A^{-1}\big(f_B(s_A)\big) \quad \text{se } n_A > n_B,$$

k		q_k	r_k	a_k	b_k	cosicché
-1			43	1	0	
0			35	0	1	
1	$43 = 1 \cdot 35 + 8$	1	8	1	-1	$8 = 1 \cdot 43 + (-1) \cdot 35$
2	$35 = 4 \cdot 8 + 3$	4	3	-4	5	$3 = (-4) \cdot 43 + 5 \cdot 35$
3	$8 = 2 \cdot 3 + 2$	2	2	9	-11	$2 = 9 \cdot 43 + (-11) \cdot 35$
4	$3 = 1 \cdot 2 + 1$	1	1	-13	16	$1 = (-13) \cdot 43 + 16 \cdot 35$
5	$2 = 2 \cdot 1 + 0$	2	0			

Tavola 11. *L'algoritmo di Euclide inizia dalla riga con $k = 1$: le prime due righe servono per completare lo schema. A sinistra eseguiamo l'algoritmo di Euclide su $(n, m) = (43, 35)$ ed usiamo i coefficienti q_k ed r_k per le operazioni a destra, mediante le formule* (3).

dove f_A^{-1} ed f_B sono definite come sopra a partire da (n_A, e_A) e (n_B, e_B) rispettivamente. Per assicurarsi dell'identità di A, B calcola

$$f_A\big(f_B^{-1}(m_A)\big) \quad \text{se } n_A < n_B; \qquad f_B^{-1}\big(f_A(m_A)\big) \quad \text{se } n_A > n_B.$$

Tutto questo funziona perché solo A può calcolare f_A^{-1}, e solo B può calcolare f_B^{-1}.

Algoritmi

L'algoritmo di Euclide

Come abbiamo visto sopra, il Teorema di Euclide implica che è possibile esprimere il massimo comun divisore d di due interi n ed m come loro combinazione lineare a coefficienti interi $d = \lambda n + \mu m$: ricordiamo che questo permette il calcolo dell'inverso moltiplicativo nel gruppo $\mathbb{Z}_n^*$.

Ora descriviamo l'Algoritmo di Euclide vero e proprio: usiamo il simbolo $\leftarrow$ per indicare l'assegnazione. Si veda la Tavola 11 per un esempio numerico.

(1) Poniamo $r_{-1} \leftarrow n$, $r_0 \leftarrow m$, $k \leftarrow 0$;
(2) se $r_k = 0$ allora $r_{k-1} = (n, m)$; l'algoritmo termina;
(3) si divide r_{k-1} per r_k trovando due interi q_{k+1} ed r_{k+1} (quoziente e resto) con la proprietà

$$r_{k-1} = q_{k+1} r_k + r_{k+1} \qquad \text{e} \qquad 0 \le r_{k+1} < r_k.$$

Si pone $k \leftarrow k + 1$. Si torna al passo 2.

L'algoritmo termina poiché la successione $(r_k) \subseteq \mathbb{N}$ è monotona decrescente. Per determinare λ e μ costruiamo le due successioni a_k e b_k:

$$a_{-1} = 1, \quad b_{-1} = 0, \quad a_0 = 0, \quad b_0 = 1.$$

Poi si calcolano a_k e b_k mediante

$$a_k = a_{k-2} - q_k a_{k-1}, \qquad b_k = b_{k-2} - q_k b_{k-1}. \tag{3}$$

Queste due successioni hanno la proprietà che $r_k = a_k n + b_k m$ per ogni $k > 0$ ed in particolare, se $r_{K+1} = 0$, per $k = K$ e quindi

$$r_K = (n, m) = a_K n + b_K m.$$

Il numero di moltiplicazioni o divisioni necessarie per l'esecuzione è $\mathcal{O}(\log m)$.

Soluzione dei sistemi di congruenze

Dato il sistema di congruenze $x \equiv a_i \mod n_i$, $i = 1, 2$, con $(n_1, n_2) = 1$, possiamo determinare i due interi λ_1, λ_2 tali che $n_1\lambda_1 + n_2\lambda_2 = 1$ per mezzo dell'Algoritmo di Euclide. Una soluzione del sistema è dunque $x_0 = a_2n_1\lambda_1 + a_1n_2\lambda_2 \mod n_1n_2$. Infatti, dato che $n_2\lambda_2 \equiv 1 \mod n_1$ si ha $x_0 \equiv a_1 \mod n_1$, ed analogamente $x_0 \equiv a_2 \mod n_2$. Se il sistema contiene piú congruenze compatibili, si possono combinare le prime due come sopra, ottenendo un nuovo sistema con una congruenza di meno, e si itera fino a rimanere con una sola congruenza.

Algoritmi di fattorizzazione

> Problema, numeros primos a compositis dignoscendi, hosque in factores suos primos resolvendi, ad gravissima ac utilissima totius arithmeticæ pertinere, et geometrarum tum veterum tum recentiorum industriam ac sagacitatem occupavisse, tam notum est, ut de hac re copiose loqui superfluum foret ... Prætereaque scientiæ dignitas requirere videtur, ut omnia subsidia ad solutionem problematis tam elegantis ac celebris sedulo excolantur.
>
> K. F. Gauss, *Disquisitiones Arithmeticae*, 1801, Art. 329.

Qui daremo una breve descrizione di alcuni algoritmi di fattorizzazione: si osservi che al giorno d'oggi si sottopone un intero N ad uno di questi algoritmi solo dopo che è stato dimostrato che non è un numero primo mediante uno dei criteri descritti qui sopra, o criteri analoghi. Quindi, nelle considerazioni che seguono, supporremo tacitamente che N sia composto.

Divisione per tentativi. Si può dimostrare che un numero intero $N \geq 2$ è primo verificando direttamente la definizione, cioè verificando che nessuna delle divisioni di N per gli interi $2 \leq m \leq N-1$ è esatta. Poiché se $N = mr$ uno fra m ed r è necessariamente $\leq \sqrt{N}$, è sufficiente effettuare $\mathcal{O}\big(N^{1/2}\big)$ divisioni. Inoltre, avendo una lista dei numeri primi $\leq \sqrt{N}$ è sufficiente provare a dividere N per ciascuno di questi numeri primi, ma in ogni caso il numero delle divisioni necessarie non è significativamente piú piccolo di $\sqrt{N}$. Naturalmente, se riusciamo a trovare m che divide N, non solo abbiamo dimostrato che N non è primo, ma abbiamo dato esplicitamente due suoi fattori non banali. Questo algoritmo ha una complessità computazionale $\mathcal{O}\big(N^{1/2}\big)$.

Fattorizzazione "alla Fermat" (Algoritmo di Lehman). Il metodo della divisione per tentativi ha certamente il vantaggio dell'estrema semplicità, ma anche l'enorme svantaggio 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 $N = 3992003 = 1997 \cdot 1999$. In questo caso è piú efficiente un altro metodo, basato su una semplice osservazione: se riusciamo a trovare x ed $y \in \mathbb{N}$ tali che $N + y^2 = x^2$, allora $N = x^2 - y^2 = (x-y)\cdot(x+y)$ e quindi N è scomposto in due fattori. Naturalmente $x - y$ ed $x + y$ non sono necessariamente primi, ed è anche possibile che $x - y$ sia proprio uguale ad 1, rendendo questa scomposizione poco interessante. In ogni modo, questa osservazione suggerisce di calcolare $N + y^2$ per alcuni valori (relativamente piccoli) di y, e di verificare se $N + y^2$ risulti essere un quadrato perfetto (osserviamo che l'algoritmo di Newton per il calcolo della radice quadrata è molto piú efficiente e piú semplice da implementare di quello insegnato di solito nelle scuole medie, visto soprattutto che qui ci interessa soltanto di sapere

se $\sqrt{n+y^2} \in \mathbb{N}$). Applicato all'esempio precedente, questo metodo funziona immediatamente: per $y = 1$ troviamo che $N+1 = 1998^2$ e quindi N ha la fattorizzazione data. Naturalmente non è possibile sapere *a priori* che le cose funzioneranno meglio con questo metodo piuttosto che con l'altro, ma è possibile "mescolarli" per ottenere un metodo di fattorizzazione piú efficiente di ciascuno dei due. In pratica si procede come segue: posto $R \stackrel{\text{def}}{=} N^{1/3}$, applichiamo la divisione per tentativi, con $m = 2$ e tutti gli interi dispari $\leq R$. Questo richiede $\mathcal{O}(R)$ divisioni. Se nessuna delle divisioni è esatta, allora N è primo oppure N è il prodotto pq di due numeri primi che soddisfano $R < p \leq q < N/R = R^2$. Si può dimostrare che se N non è primo è possibile trovare x, y e $k \in \mathbb{N}$ tali che

$$\begin{cases} x^2 - y^2 = 4kN & \text{dove } 1 \leq k \leq R \\ 0 \leq x - \sqrt{4kN} \leq \sqrt{\frac{N}{k}}(4R)^{-1} \\ p = \min\big((x+y, N), (x-y, N)\big). \end{cases}$$

Per determinare x, y e k, procediamo di nuovo per tentativi, verificando se, fissato k, esiste un valore intero di x compreso fra $x_0 \stackrel{\text{def}}{=} \lceil\sqrt{4kN}\rceil$ ed $x_1 \stackrel{\text{def}}{=} \lfloor\sqrt{4kN}+\sqrt{N/k}/4R\rfloor$ per il quale $x^2 - 4kN$ sia un quadrato perfetto. Si dimostra che anche questa parte del calcolo richiede al massimo $\mathcal{O}(R)$ operazioni, e quindi il costo totale dell'algoritmo è $\mathcal{O}(R) = \mathcal{O}\big(N^{1/3}\big)$. Senza entrare nei dettagli, se $N = pq$ con $R < p \leq q < R^2$ ed esistono r, $s \in \mathbb{N}^*$ tali che $\frac{p}{q} \approx \frac{r}{s}$ allora il numero $pqrs = (ps)(rq)$ ha due fattori quasi uguali ed è relativamente facile determinarli con il metodo visto sopra. Evidentemente, questo dà un buon algoritmo di fattorizzazione se si può dimostrare che esistono r ed s piú piccoli di p.

In pratica si procede come segue:

1. Si pone $R \leftarrow N^{1/3}$ e si divide N per $m = 2$ e per tutti gli interi dispari 3, 5, ... , fino ad R. Se qualche divisione è esatta l'algoritmo termina.
2. Si pone $k \leftarrow 1$.
3. Si pone $x_0 \leftarrow \lceil\sqrt{4kN}\rceil$, $x_1 \leftarrow \lfloor\sqrt{4kN} + \sqrt{N/k}/4R\rfloor$ e si verifica se per qualche $x \in [x_0, x_1]$ si ha che $x^2 - 4kN$ è un quadrato perfetto. Se questo accade l'algoritmo termina con il calcolo di $(x+y, N)$.
4. Si pone $k \leftarrow k+1$; se $k \leq R$ si ripete il passo 3.

Per esempio, se $N = 2881$ allora $R = 14$. N non ha fattori primi ≤ 14, ed utilizzando il metodo descritto sopra si trova che per $k = 1$ ed $x = 110$ si ha $110^2 - 4 \cdot 2881 = 24^2$, da cui deduciamo che $p = 43$ e $q = 67$. In questo caso $\frac{p}{q} \approx \frac{2}{3}$. Un esempio piú complicato è $N = 11303$ che per $k = 6$ dà $x = 521$ ed $y = 13$, da cui segue $N = 89 \cdot 127$.

Il crivello quadratico. Per brevità, ci limiteremo a parlare di un solo algoritmo subesponenziale, la cui complessità è $\mathcal{O}(N^\varepsilon)$ per ogni $\varepsilon > 0$. Questo algoritmo appartiene ad una famiglia di algoritmi simili basati con qualche variante sullo stesso schema di fondo. Lo schema di cui parliamo, dovuto a Kraitchik, si può riassumere come segue:

1. determinazione di congruenze $A_i \equiv B_i \mod N$ con $A_i \neq B_i$;
2. determinazione della scomposizione in fattori primi (parziale o completa) dei numeri A_i, B_i per un sottoinsieme delle congruenze ottenute sopra;
3. determinazione di un sottoinsieme $\mathcal{S}$ delle congruenze ottenute nel punto 2

A	$Q(A)$	Fattorizzazione	$\vec{v}(A)$	$\vec{v}(A) \mod 2$
1	200	$2^3 \cdot 5^2$	$(3,2,0,0)$	$(1,0,0,0)$
3	608	$2^5 \cdot 19$	$(5,0,0,1)$	$(1,0,0,1)$
5	1024	2^{10}	$(10,0,0,0)$	$(0,0,0,0)$
6	1235	$5 \cdot 13 \cdot 19$	$(0,1,1,1)$	$(0,1,1,1)$
19	4160	$2^6 \cdot 5 \cdot 13$	$(6,1,1,0)$	$(0,1,1,0)$
41	9880	$2^3 \cdot 5 \cdot 13 \cdot 19$	$(3,1,1,1)$	$(1,1,1,1)$
51	12800	$2^9 \cdot 5^2$	$(9,2,0,0)$	$(1,0,0,0)$

Tavola 12. *Implementazione del crivello quadratico per la fattorizzazione di* $10001 = 73 \cdot 137$. *Qui scegliamo come base di fattori l'insieme* $\mathcal{B} \stackrel{\text{def}}{=} \{2, 5, 13, 19\}$. *Nella Tavola sono riportati i valori di* A *per cui* $Q(A)$ *si fattorizza completamente in* $\mathcal{B}$, *il valore di* $Q(A)$, *i vettori* $\vec{v}(A)$ *corrispondenti, e gli stessi vettori modulo* 2. *Si osservi che i vettori negli insiemi* $\{\vec{v}(1), \vec{v}(51)\}$, $\{\vec{v}(3), \vec{v}(6), \vec{v}(19), \vec{v}(51)\}$, $\{\vec{v}(5)\}$, $\{\vec{v}(6), \vec{v}(41), \vec{v}(51)\}$, $\{\vec{v}(1), \vec{v}(3), \vec{v}(6), \vec{v}(19)\}$, $\{\vec{v}(1), \vec{v}(6), \vec{v}(41)\}$, $\{\vec{v}(3), \vec{v}(19), \vec{v}(41)\}$, *sono linearmente dipendenti* $\mod 2$, *ma solo i primi 4 portano alla scoperta di un fattore non banale di* 10001.

tale che

$$\prod_{i \in \mathcal{S}} A_i \equiv X^2 \mod N; \qquad \prod_{i \in \mathcal{S}} B_i \equiv Y^2 \mod N;$$

4. calcolo di $d = (X - Y, N)$ per ottenere un fattore di N.

Di solito, ci si assicura preliminarmente che N non abbia fattori primi molto piccoli. Gli algoritmi di questa famiglia differiscono in qualche dettaglio nella realizzazione pratica delle varie fasi indicate qui. L'obiettivo del crivello quadratico è la determinazione di una congruenza non banale $X^2 \equiv Y^2 \mod N$, dove N è il numero da scomporre in fattori. Si calcola poi $d = (X - Y, N)$ che è un fattore di N: se $1 < d < N$, allora abbiamo scomposto N nel prodotto di due fattori non banali.

La Tavola 12 illustra alcune fasi dell'algoritmo. In generale, poniamo

$$Q(A) = \left(A + \left[N^{1/2}\right]\right)^2 - N.$$

Osserviamo che per ogni A si ha $Q(A) \equiv \left(A + \left[N^{1/2}\right]\right)^2 \mod N$ e quindi un membro della congruenza cercata è sicuramente un quadrato perfetto. Si costruisce una "base di fattori" $\mathcal{B} = \{2\} \cup \{p$ dispari, p è "piccolo" e l'equazione $Q(A) \equiv 0 \mod p$ ha soluzione$\}$, e si pone $k = |\mathcal{B}|$. Per A piccolo, $Q(A) \approx 2A\sqrt{N}$ è relativamente piccolo e quindi è probabile che si riesca a scomporre in fattori primi *tutti appartenenti a* $\mathcal{B}$ numerosi valori $Q(A)$.

Se A_j è un intero per cui $Q(A_j)$ si fattorizza completamente su $\mathcal{B}$, diciamo $Q(A_j) = \prod_{p \in \mathcal{B}} p^{\alpha_{p,j}}$, costruiamo il vettore $\vec{v}(A_j) \in \mathbb{N}^k$ che ha come componenti gli esponenti $\alpha_{p,j}$, e poi riduciamo queste componenti modulo 2, ottenendo i vettori $\vec{v}_2(A_j)$. Una semplice applicazione dell'algebra lineare su $\mathbb{Z}_2$ ci permette di concludere che $k+1$ di questi vettori ridotti sono certamente linearmente dipendenti su $\mathbb{Z}_2$. Una relazione di dipendenza lineare su $\mathbb{Z}_2$ significa semplicemente che $\vec{v}_2(A'_1) + \cdots + \vec{v}_2(A'_m) \equiv \vec{0} \mod 2$ (i coefficienti della relazione di dipendenza lineare possono essere solo 0 o 1); una volta determinato un insieme $\mathcal{I}$ di indici tale che

$\{\vec{v}_2(A_j)\colon j \in \mathcal{I}\}$ sia linearmente dipendente su $\mathbb{Z}_2$, abbiamo trovato la combinazione di congruenze cercata. Infatti, per quanto osservato sopra, si ha

$$\prod_{j\in\mathcal{I}} \bigl(A_j + \bigl[N^{1/2}\bigr]\bigr)^2 \equiv \prod_{j\in\mathcal{I}} Q(A_j) \equiv \prod_{p\in\mathcal{B}} p^{\sum_{j\in\mathcal{I}} \alpha_{p,j}} \mod N$$

e, per costruzione, ciascuno degli esponenti a destra è pari. A questo punto si può passare alla quarta fase del programma, il calcolo del massimo comun divisore d. Si osservi che se $d = 1$ oppure $d = N$, è sufficiente cercare un'ulteriore fattorizzazione di qualche nuovo $Q(A)$, e ripetere il passo 3. Vi sono numerosi accorgimenti per migliorare l'efficienza dell'algoritmo, la cui complessità è stimata in una potenza di

$$L(N) = \exp\bigl((\log N \log\log N)^{1/2}\bigr).$$

Ricerca di un generatore di $\mathbb{Z}_p^*$

Per ogni p primo esiste $g \in \mathbb{Z}_p^*$ che genera $\mathbb{Z}_p^*$, cioè che ha ordine $p-1$ (in effetti ce ne sono $\varphi(p-1)$). L'algoritmo per determinare un generatore è dovuto a Gauss.

1. Si sceglie $a_1 \in \mathbb{Z}_p^*$ e si calcolano a_1, $a_1^2 \mod p$, $a_1^3 \mod p$, ... Sia r_1 l'ordine di $a_1 \mod p$: se $r_1 = p-1$ allora a_1 è un generatore ed abbiamo finito;
2. Sia $b_1 \in \mathbb{Z}_p^* \setminus \{a_1, a_1^2 \mod p, \dots, a_1^r \mod p\}$, di ordine s_1. Se $s_1 = p-1$ allora b_1 è un generatore ed abbiamo finito; altrimenti poniamo $v_1 = [r_1, s_1]$. Possiamo scrivere $v_1 = n_1 m_1$ con $(n_1, m_1) = 1$, $n_1 \mid r_1$, $m_1 \mid s_1$.
3. Sia $a_2 = a_1^{v_1/n_1} b_1^{v_1/m_1}$; si può verificare che l'ordine r_2 di a_2 è $> \max(r_1, s_1)$, e quindi abbiamo trovato un intero che ha ordine piú grande di a_1.

Si ripetono questi passi fino a trovare un generatore.

Prendiamo $p = 41$, $a_1 = 2$. Le potenze di a_1, ridotte $\mod p$, sono nell'ordine 2, 4, 8, 16, 32, 23, 5, 10, 20, 40, 39, 37, 33, 25, 9, 18, 36, 31, 21, 1, e quindi $r_1 = 20$.

Possiamo prendere $b_1 = 3$, e calcolarne le potenze successive: 3, 9, 27, 40, 38, 32, 14, 1, e quindi $s_1 = 8$. Dunque $v_1 = [20, 8] = 40$, $n_1 = 5$, $m_1 = 8$, $a_2 = 2^4 \cdot 3 \mod p = 7$, e l'ordine di 7 in $\mathbb{Z}_p^*$ è 40.

Per il numero primo $p = 65537 = 2^{16} + 1$ si può prendere $g = 75$. Poiché $p-1$ è una potenza di 2 e l'ordine di 75 deve dividere $p-1$, deve essere a sua volta una potenza di 2 ed è quindi sufficiente verificare che $75^n \not\equiv 1 \mod p$ quando n è una potenza di 2 minore di $p-1$.

Logaritmo discreto

Anche in questo caso illustriamo il funzionamento dell'algoritmo per il calcolo del logaritmo discreto per mezzo di un esempio. Prima però è opportuno mettere in guardia i lettori che conoscono l'Analisi Matematica: per calcolare con una certa approssimazione il logaritmo di un numero reale positivo si sfruttano proprietà quali continuità, derivabilità, convessità e monotonia delle funzioni esponenziale e logaritmo. Qui invece il concetto di monotonia (che si basa sulle disuguaglianze) non ha alcun senso, né, evidentemente, ne possono avere continuità e derivabilità, ed inoltre il logaritmo discreto in $\mathbb{Z}_p^*$ è un elemento di $\mathbb{Z}_{p-1}$ e sarà determinato esattamente, senza approssimazioni. Si tratta quindi di un problema di natura essenzialmente diversa da quello con lo stesso nome che conosciamo dall'Analisi.

Poiché 3 è un generatore di $\mathbb{Z}_{31}^*$, vogliamo trovare il *logaritmo discreto* di 7 in base 3, cioè l'elemento x di $\mathbb{Z}_{30}$ tale che $3^x \equiv 7 \mod 31$. Il calcolo comprende due parti.

q	r	S	A	B
		0	27	41
13	1	$0+41=41$	13	82
6	1	$41+82=123$	6	164
3	0	123	3	328
1	1	$123+328=451$	1	656
0	1	$451+656=1107$	0	1312

Tavola 13. *Si osservi che alla fine di ogni ciclo si ha sempre $m \cdot n = S + A \cdot B$.*

Precomputazione. Si calcolano i numeri $r_{j,p} \equiv 3^{30j/p} \mod 31$ per tutti i fattori primi p di 30, e per $j = 0, 1, \dots, p-1$. Questo ci dà la tabella

$$\begin{array}{lllll} r_{0,2}=1 & r_{1,2}=-1 & & & \\ r_{0,3}=1 & r_{1,3}=-6 & r_{2,3}=5 & & \\ r_{0,5}=1 & r_{1,5}=16 & r_{2,5}=8 & r_{3,5}=4 & r_{4,5}=2 \end{array}$$

Osserviamo che $r_{j,p}^p \equiv 1 \mod 31$: poiché 3 genera $\mathbb{Z}_{31}^*$, i numeri $r_{j,p}$ sono tutte e sole le radici p-esime di 1.

Il logaritmo discreto. Se $3^x \equiv 7 \mod 31$ ed $x = a + 2a'$ con $a \in \{0, 1\}$, allora

$$3^{15x} = 3^{15a+30a'} \equiv 3^{15a} \equiv 7^{15} \equiv 1 \mod 31.$$

Ora notiamo che $(7^{15})^2 \equiv 7^{30} \equiv 1 \mod 31$, cioè 7^{15} è una delle due radici quadrate di 1 calcolate sopra, ed in effetti l'ultima congruenza rivela che $7^{15} = r_{0,2}$. Poiché $3^0 \equiv 1 \mod 31$, mentre $3^{15} \equiv -1 \mod 31$, concludiamo che $a = 0$, cioè che $x \equiv 0 \mod 2$. Analogamente, se $x = b + 3b'$ con $b \in \{0, 1, 2\}$, allora

$$3^{10x} = 3^{10b+30b'} \equiv 3^{10b} \equiv 7^{10} \equiv -6 \mod 31,$$

da cui $b = 1$ cioè $x \equiv 1 \mod 3$. Infine, se $x = c + 5c'$ con $c \in \{0, 1, 2, 3, 4\}$ allora

$$3^{6x} = 3^{6c+30c'} \equiv 3^{6c} \equiv 7^6 \equiv 4 \mod 31,$$

da cui $c = 3$ cioè $x \equiv 3 \mod 5$. Troviamo cosí il sistema di congruenze

$$\begin{cases} x \equiv 0 \mod 2 \\ x \equiv 1 \mod 3 \\ x \equiv 3 \mod 5 \end{cases} \qquad \text{da cui} \qquad x \equiv 28 \mod 30$$

per il Teorema Cinese del Resto. Un algoritmo simile (ma piú complicato) funziona quando l'ordine del gruppo è divisibile per potenze di un primo piú grandi di 1. Concludiamo osservando che per eseguire questi calcoli in $\mathbb{Z}_p^*$ è necessario conoscere la completa scomposizione in fattori primi di $p-1$.

Calcolo di prodotti modulo n

L'algoritmo del prodotto $m \cdot n$ è illustrato nella Tavola 13.

1. Si assegnano i valori iniziali $S \leftarrow 0$, $A \leftarrow m$, $B \leftarrow n$;
2. Si determinano q ed r (quoziente e resto della divisione di A per 2) in modo che $A = 2 \cdot q + r$, con $r \in \{0, 1\}$. Se $r = 1$ poniamo $S \leftarrow S + B$.
3. Si pone $A \leftarrow q$, $B \leftarrow 2 \cdot B$.
4. Se $q = 0$ l'algoritmo termina ed S vale $m \cdot n$. Altrimenti si ripete il passo 2.

q	r	P	M	A
		1	23	a
11	1	$1 \cdot a = a$	11	a^2
5	1	$a \cdot a^2 = a^3$	5	a^4
2	1	$a^3 \cdot a^4 = a^7$	2	a^8
1	0	a^7	1	a^{16}
0	1	$a^7 \cdot a^{16} = a^{23}$	0	

Tavola 14. *Si osservi che alla fine di ogni ciclo si ha sempre* $a^{23} = P \cdot A^M$.

Calcolo di potenze modulo n

Volendo calcolare a^m, dove $m \in \mathbb{N}^*$, scriviamo a^m come un prodotto di potenze con base a il cui esponente sia una potenza di 2. Per esempio, per determinare a^{23} basta calcolare a^2, a^4, a^8, a^{16} (quattro elevamenti al quadrato) e poi $a \cdot a^2 \cdot a^4 \cdot a^{16}$, per un totale di sole 7 moltiplicazioni, invece delle 22 necessarie per eseguire il calcolo nel modo consueto. Il numero totale delle moltiplicazioni è $\mathcal{O}(\log m)$.

1. Poniamo $P \leftarrow 1$, $M \leftarrow m$, $A \leftarrow a$.
2. Si determinano q ed r rispettivamente quoziente e resto della divisione di M per 2. Se $r = 1$ poniamo $P \leftarrow P \cdot A$.
3. Poniamo $A \leftarrow A^2$, $M \leftarrow q$.
4. Se $M = 0$ l'algoritmo termina e $P = a^m$. Altrimenti si torna al passo 2.

Si veda la Tavola 14. Questo algoritmo è particolarmente utile quando si devono fare calcoli modulo un numero molto grande N: facendo seguire ad ogni operazione di somma o prodotto il calcolo del resto $\bmod N$, si può fare in modo che tutti i risultati parziali del calcolo siano $\leq 2N$. Inoltre, se invece di prendere il minimo resto positivo, si prende il minimo resto in valore assoluto (cioè, se quando il resto $r \in \left[\frac{1}{2}N, N\right]$ si sceglie $r' \stackrel{\text{def}}{=} r - N \in \left[-\frac{1}{2}N, 0\right]$), tutti i risultati parziali dei calcoli sono, in valore assoluto, $\leq N$.

Distribuzione dei numeri primi

Risultati quantitativi

Aggiungiamo senza dimostrazione alcuni risultati relativi alla distribuzione dei numeri primi, per prima cosa per il loro interesse intrinseco, e poi perché sono rilevanti per la determinazione della complessità di alcuni algoritmi di fattorizzazione.

Teorema dei Numeri Primi. *Posto*

$$\pi(x) \stackrel{\text{def}}{=} |\{p : p \text{ è primo e } p \leq x\}| \qquad e \qquad \mathrm{li}(x) \stackrel{\text{def}}{=} \int_2^x \frac{\mathrm{d}t}{\log t},$$

per $x \to +\infty$ *si ha*

$$\pi(x) = \mathrm{li}(x) + \mathcal{O}\Big(x \exp\{-\sqrt{\log x}\}\Big).$$

Si osservi che per $x \to +\infty$ si ha $\mathrm{li}(x) \sim x(\log x)^{-1}$, ma la relazione espressa dal Teorema dei Numeri Primi è piú precisa. Una versione piú generale di questo risultato riguarda i numeri primi nelle progressioni aritmetiche:

Teorema dei Numeri Primi nelle Progressioni Aritmetiche. *Posto*

$$\pi(x;q,a) \stackrel{\text{def}}{=} |\{p\colon p \text{ è primo, } p \le x,\ p \equiv a \mod q\}|,$$

e fissata $A > 0$, esiste una costante $C = C(A) > 0$ tale che, uniformemente per $1 \le q \le (\log x)^A$, $a \in \mathbb{Z}$ tale che $(a,q) = 1$, si ha

$$\pi(x;q,a) = \frac{1}{\varphi(q)}\,\mathrm{li}(x) + \mathcal{O}\Big(x\exp\{-C\sqrt{\log x}\}\Big).$$

In particolare, fissati $a \in \mathbb{Z}$ e $q \ge 1$, se $(a,q) = 1$ allora circa $1/\varphi(q)$ dei numeri primi nell'intervallo $[1,x]$ sono $\equiv a \mod q$. Osserviamo che in entrambi i casi è stato dimostrato un risultato piú forte, ma piú complicato da enunciare, e che si congettura che debba valere un risultato ancora piú forte, che enunciamo solo nel primo caso:

Congettura di Riemann. *Per $x \to +\infty$ si ha*

$$\pi(x) = \mathrm{li}(x) + \mathcal{O}\Big(x^{1/2}\log x\Big).$$

Siamo molto lontani dal poter dimostrare un risultato del genere, che in un certo senso è ottimale. Infatti, l'esponente $\frac{1}{2}$ non può essere certamente abbassato.

Abbiamo visto nella descrizione del crivello quadratico che è importante la distribuzione degli interi privi di fattori primi "grandi": infatti il tempo di esecuzione dipende dalla frequenza con cui si trovano interi che si scompongono completamente in fattori tutti appartenenti alla base di fattori $\mathcal{B}$. Definiamo quindi $\Psi(x,y) \stackrel{\text{def}}{=} |\{n \le x\colon p \mid n \Rightarrow p \le y\}|$: dunque $\Psi(x,y)$ è il numero degli interi $n \le x$ che non hanno fattori primi $> y$. È possibile dimostrare che il comportamento di questa funzione dipende in modo cruciale dal valore di $u = (\log x)/\log y$, quando $x \ge 1$, $y \ge 2$. Non è facile enunciare un risultato valido per ogni valore di u: il piú significativo è forse la relazione $\Psi(x,y) = xu^{-(1+o(1))\log u}$, che vale uniformemente in un'ampia regione di valori di u, e cioè $u \le y^{1-\varepsilon}$, dove $\varepsilon > 0$ è fissato, y ed u tendono ad infinito.

È stato dimostrato recentemente che esistono infiniti numeri di Carmichael, e che questi sono anche piuttosto frequenti, nel senso, che per x sufficientemente grande si ha

$$C(x) \stackrel{\text{def}}{=} |\{n \le x\colon n \text{ è di Carmichael}\}| \ge x^{2/7}.$$

Si congettura che fissato $\varepsilon > 0$, per $x > x_0(\varepsilon)$ si abbia $C(x) > x^{1-\varepsilon}$.

Abbiamo visto sopra che ogni gruppo del tipo $\mathbb{Z}_p^*$ è ciclico e quindi ha un generatore g_p, ed anche un algoritmo per determinare g_p. La Tavola 15 mostra l'ordine degli interi $n \in \{2,\dots,13\}$ modulo i primi p fra 2 e 31: i generatori vi sono indicati per mezzo di un $\star$. Concludiamo questa discussione con l'enunciato della

Congettura di Artin. *Sia $g \in \mathbb{Z}$ un intero diverso da 0, -1 e che non sia un quadrato perfetto. Allora g è un generatore di $\mathbb{Z}_p^*$ per infiniti valori di p e, piú precisamente,*

$$\lim_{x\to+\infty} \frac{|\{p \le x\colon p \text{ è primo e } g \text{ genera } \mathbb{Z}_p^*\}|}{\pi(x)} > 0.$$

p, n	2	3	4	5	6	7	8	9	10	11	12	13
2		$1^\star$		$1^\star$		$1^\star$		$1^\star$		$1^\star$		$1^\star$
3	$2^\star$		1	$2^\star$		1	$2^\star$		1	$2^\star$		1
5	$4^\star$	$4^\star$	2		1	$4^\star$	$4^\star$	2		1	$4^\star$	$4^\star$
7	3	$6^\star$	3	$6^\star$	2		1	3	$6^\star$	3	$6^\star$	2
11	$10^\star$	5	5	5	$10^\star$	$10^\star$	$10^\star$	5	2		1	$10^\star$
13	$12^\star$	3	6	4	$12^\star$	$12^\star$	4	3	6	$12^\star$	2	
17	8	$16^\star$	4	$16^\star$	$16^\star$	$16^\star$	8	8	$16^\star$	$16^\star$	$16^\star$	4
19	$18^\star$	$18^\star$	9	9	9	3	6	9	$18^\star$	3	6	$18^\star$
23	11	11	11	$22^\star$	11	$22^\star$	11	11	$22^\star$	$22^\star$	11	11
29	$28^\star$	$28^\star$	14	14	14	7	$28^\star$	14	$28^\star$	$28^\star$	4	14
31	5	$30^\star$	5	3	6	15	5	15	15	$30^\star$	$30^\star$	$30^\star$

Tavola 15. *Gli ordini degli interi $n = 2, \ldots, 13$ modulo i primi $p = 2, \ldots, 31$. I generatori sono indicati da un $\star$.*

È evidente che se $g = -1$ oppure se $g = m^2$ allora g al massimo ha ordine 2 o, rispettivamente, $\frac{1}{2}(p-1)$, e quindi non può generare $\mathbb{Z}_p^*$. Oggi è noto che, sempre ammesso che esistano, le eccezioni a questa congettura sono molto rare.

Riferimenti bibliografici

[1] L. M. Adleman, C. Pomerance, R. S. Rumely, *On distinguishing prime numbers from composite numbers*, Ann. Math. **117** (1983), 173–206.

[2] L. Childs, *A Concrete Introduction to Higher Algebra*, Springer, 1979.

[3] H. Cohen, *A Course in Computational Algebraic Number Theory*, GTM 138, 3ª ed., Springer, 1996.

[4] R. Crandall & C. Pomerance, *Prime numbers. A computational perspective*, Springer, 2001. Questo libro recentissimo contiene una dettagliata descrizione di tutti i piú moderni algoritmi che riguardano i numeri primi (inclusi quelli trattati qui), tra cui i metodi di fattorizzazione mediante curve ellittiche, ed il "Crivello con i Campi di Numeri" (Number Field Sieve) che al momento attuale sembra essere il migliore in circolazione.

[5] J. D. Dixon, *Factorization and primality tests*, Amer. Math. Monthly **91** (1984), 333–352.

[6] K. F. Gauss, *Disquisitiones Arithmeticae*, Lipsia, 1801. L'algoritmo per la determinazione di un generatore è descritto ai §§73–74.

[7] G. H. Hardy & E. M. Wright, *An Introduction to the Theory of Numbers*, 5ª ed., Oxford Science Publications, 1979. Contiene la maggior parte dei risultati teorici di cui abbiamo parlato qui: si vedano in particolare i Capp. 5–7.

[8] D. E. Knuth, *The Art of Computer Programming. Vol. 2. Seminumerical Algorithms*, 2ª ed., Addison Wesley, Reading (Mass.), 1981.

[9] N. Koblitz, *A Course in Number Theory and Cryptography*, GTM 114, 2ª ed., Springer, 1994.

[10] E. Landau, *Elementary Number Theory*, Chelsea, New York, 1960.

[11] A. Menezes, P. van Oorschot & S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996. Si vedano in particolare i Capp. 1–3 e la Bibliografia. In forma elettronica: `http://www.cacr.math.uwaterloo.ca/hac`

[12] E. A. Poe, *The Gold Bug*, in "The complete tales and poems of Edgar Allan Poe," Random House, 1975. Trad. it. *Lo scarabeo d'oro*, in "Racconti," L'Unità–Einaudi. È una vivace descrizione di come si può rompere un sistema crittografico monoalfabetico mediante un'analisi di frequenza.

[13] C. Pomerance, *Recent developments in primality testing*, Math. Intellig. **3** (1981), 97–105.

[14] ———, *Alla ricerca dei numeri primi*, Le Scienze **174** (febbr. 1983), 86–94. Una introduzione molto leggibile ai problemi di cui abbiamo parlato.

[15] ———, *The quadratic sieve factoring algorithm*, Advances in Cryptology, Proceedings of EUROCRYPT 84, LNCS 209, Springer, 1985, pp. 169–182.

[16] ———, *Factoring*, Cryptology and computational number theory, Lect. Notes AMS Short Course, Boulder, CO (USA), 1989, Proc. Symp. Appl. Math., vol. 42, 1990, 27–47. Contiene la descrizione del crivello quadratico e del metodo di fattorizzazione basato sulle curve ellittiche.

[17] ———, *A tale of two sieves*, Notices Amer. Math. Soc. **43** (1996), 1473–1485.

[18] C. Pomerance, J. L. Selfridge & S. S. Wagstaff, *The pseudoprimes to* $25 \cdot 10^9$, Math. Comp. **35** (1980), 1003–1026.

[19] P. Ribenboim, *The New Book of Prime Numbers Records*, Springer, 1996. Per i risultati teorici su pseudoprimi, numeri di Carmichael ed ulteriori estensioni di questi concetti si vedano in particolare i §§2.II.C, 2.II.F, 2.III, 2.VIII, 2.IX. Per i generatori si veda il §2.II.A, per la crittografia il §2.XII.B, mentre la Congettura di Artin è discussa nel §6.I.

[20] H. Riesel, *Prime numbers and computer methods for factorization*, 2ª ed., Birkhäuser, Boston, 1994.

[21] D. Shanks, *Solved and Unsolved Problems in Number Theory*, 4ª ed., Chelsea, New York, 1993. Contiene una discussione dettagliata della struttura dei gruppi $\mathbb{Z}_m^*$ per qualunque valore di $m \in \mathbb{Z}$ nei §§23–38: si vedano in particolare i diagrammi nel §33. Nel §35 c'è la dimostrazione del Teorema che riguarda i gruppi moltiplicativi ciclici del tipo $\mathbb{Z}_m^*$.