UNIVERSITÀ DEGLI STUDI DI ROMA TRE
FACOLTÀ DI SCIENZE M.F.N.

Sintesi della Tesi di Laurea in Matematica

di

Sveva Coltellacci

Fattorizzazione degli interi con il metodo delle curve ellittiche

Relatore

Prof.re Francesco Pappalardi

ANNO ACCADEMICO 2002 - 2003
Luglio 2003

Classificazione AMS : 11Y5, 94A60,
Parole Chiave : TEORIA COMPUTAZIONALE DEI NUMERI, FATTORIZZAZIONE, CRITTOGRAFIA.

L'obiettivo di questa tesi è di studiare l'algoritmo di Lenstra per fattorizzare un numero composto intero attraverso l'uso delle curve ellittiche. Il lavoro è stato suddiviso in tre parti.

Nella prima parte abbiamo introdotto il concetto di curva ellittica e di gruppo dei punti razionali delle curve ellittiche. Si tratta di nozioni base che possono essere trovate in vari testi classici. Noi abbiamo deciso di seguire il testo di C. Pomerance [7]. Sia $\mathbb{F}$ un campo finito. Definiamo una curva ellittica:

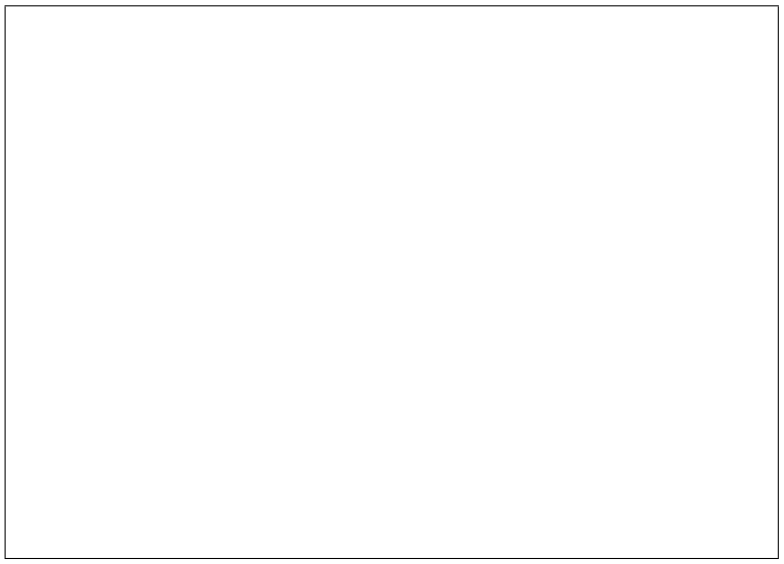
Definizione 1.2.1. *Una curva cubica non singolare del tipo* $ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0$ *con coefficienti in un campo* $\mathbb{F}$ *e con almeno un punto* P *con coordinate in* $\mathbb{F}$*, si chiama curva ellittica su* $\mathbb{F}$*.*
Se la caratteristica di $\mathbb{F}$ *è diversa da 2 e 3, è sempre possibile, attraverso un opportuno cambiamento di coordinate, ricondursi al caso* $y^2 = x^3 + ax + b$ *e* $y^2 = x^3 + Cx^2 + Ax + B$*. Tali equazioni definiscono curve ellittiche su* $\mathbb{F}$ *con la condizione che* $\Delta \neq 0$*, con* Δ *definito come segue:*
$\Delta = 4a^3 + 27b^2$*, se la curva ha equazione* $y^2 = x^3 + ax + b$ *oppure* $\Delta = 4A^3 + 27B^2 + 18ABC - A^2C^2 + 4BC^3$ *se la curva ha equazione* $y^2 = x^3 + Cx^2 + Ax + B$*.*

Le curve ellittiche hanno una notevole importanza se utilizzate con algoritmi per fattorizzare interi composti, ma tutto diventerà più chiaro dopo aver definito l'operazione di gruppo grazie alla quale $E(\mathbb{F}_p)$ diventa un gruppo abeliano.

Definizione 1.2.2. *Sia* $E(\mathbb{F})$ *l'insieme dei punti di una curva ellittica* $y^2 = x^3 + ax + b$ *sul campo* $\mathbb{F}$ *con caratteristica diversa da* 2 *e* 3*, con l'aggiunta di un punto supplementare che chiameremo punto all'infinito, che denotiamo con* $\mathcal{O}$*. Siano due punti arbitrari della curva,* $P_1 = (x_1, y_1)$ *e* $P_2 = (x_2, y_2)$*, non necessariamente distinti, e diversi da* $\mathcal{O}$*. Definiamo un'operazione commutativa* $+$ *con operazione inversa* $-$ *nel modo seguente:*

- $-\mathcal{O} = \mathcal{O}$;
- *Se* $P_1 = (x_1, y_1)$, *allora*
  $-P_1 = (x_1, -y_1)$;
- $\mathcal{O} + P_1 = P_1$;
- *se* $P_2 = -P_1$ *allora* $P_1 + P_2 = \mathcal{O}$;
- *se* $P_2 \neq -P_1$ *allora* $P_1 + P_2 = (x_3, y_3)$ *dove*
  $x_3 = m^2 - x_1 - x_2$ *e*
  $y_3 = m(x_3 - x_1) + y_1$ *dove*

$$m = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & se\ x_2 \neq x_1; \\ \\ \frac{3x_1^2 + A}{2y_1} & se\ x_2 = x_1, y_2 = y_1. \end{cases}$$



1.1: Somma di due punti

Grazie a tale definizione possiamo enunciare un teorema classico noto dai tempi di Jacobi:

Teorema 1.2.3 (Jacobi). *L'insieme dei punti razionali di una curva ellittica $E(\mathbb{F})$ con l'operazione definita sopra è un gruppo abeliano finito.*

Abbiamo poi enunciato un teorema, vedi Pomerance [7], che ci permette di avere informazioni sull'ordine e la struttura del gruppo $E(\mathbb{F})$.

Teorema 1.2.4 (Cassels). *Nel caso di un campo finito $\mathbb{F}_p$ con p elementi, il gruppo $E(\mathbb{F}_p)$ è ciclico o isomorfo al prodotto di due gruppi ciclici, i.e.*

$$E(\mathbb{F}_p) \cong \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z}.$$

Inoltre valgono le seguenti proprietà: $d_1 \,|d_2$ e $d_1 \,|p-1$.

Il simbolo $\#E(\mathbb{F})$ denota l'ordine del gruppo $E(\mathbb{F})$, cioè il numero delle soluzioni (x, y) dell'equazione che definisce E più 1(il punto all'infinito).

Definizione 1.2.6. *Sia E una curva ellittica. Per ogni $P \in E(\mathbb{F})$ e $n \in \mathbb{Z}$ indichiamo con $[n]\,P$ il punto:*

$$[n]\,P = \underbrace{P + P + \cdots + P}_{n \text{ volte}}$$

Definiamo inoltre $[0]\,P = \mathcal{O}$ e $[-n]\,P = -\,[n]\,P$. In particolare $[\#E(\mathbb{F})]P{=}\mathcal{O}$.

Questa definizione in alcuni casi ci permette di trovare l'ordine del gruppo dei punti razionali. Vediamo un esempio:

Esempio.

$E : y^2 = x^3 + 3x$ sul campo $\mathbb{F}_7$.
Poichè x e y si suppone siano in $\mathbb{F}_7 = \{0, 1, 2, 3, 4, 5, 6\}$, possiamo considerare i sette possibili valori di x, sostituirli nell'equazione e verificare quando il risultato è un quadrato in $\mathbb{F}_7$. Otteniamo 8 punti, incluso il punto all'infinito $\mathcal{O}$, cioè

$$\mathbb{E}(\mathbb{F}_7) = \{\mathcal{O}, (0,0), (1, \pm 2), (2,0), (3, \pm 1), (5,0)\}\,.$$

Quindi $E(\mathbb{F}_7)$ è un gruppo abeliano di ordine otto, ciclico o prodotto di due gruppi ciclici, uno di ordine 2 e uno di ordine 4. Per determinare in quale di questi due casi ci troviamo, basta verificare che nessun un punto ha ordine 8. Infatti, facendo i calcoli, usando le formule della definizione 1.2.2, si verifica che i punti $(0,0),(2,0),(5,0)$ hanno ordine 2, mentre i punti $(1,\pm 2),(3,\pm 1)$ hanno ordine 4. In conclusione, si ha

$$E(\mathbb{F}_7) \cong \mathbb{Z}_2 \times \mathbb{Z}_4.$$

Nella pratica vedremo che il problema più rilevante è quello di determinare quale sia l'ordine del gruppo.
Enunciamo, quindi, l'importante teorema sull'ordine del gruppo dei punti razionali di una curva ellittica E:

Teorema 1.3.2 (Hasse,Weil). *Sia E una curva non singolare definita sul campo finito $\mathbb{F}_p$, allora il numero dei punti su $E(\mathbb{F}_p)$ con coordinate in $\mathbb{F}_p$ è $p+1+t$ dove t è tale che $|t| \leq 2\sqrt{p}$. Esplicitamente*

$$|(\#E) - (p+1)| \leq 2\sqrt{p}.$$

Dopo aver trovato che l'ordine del gruppo è, dunque, un intero nell'intervallo $((\sqrt{p}-1)^2,(\sqrt{p}+1)^2)$, abbiamo studiato alcuni importanti algoritmi, proposti da Shanks-Menstre [6] e Schoof [18], su come determinare tale ordine. Per meglio comprendere il calcolo della complessità di questi algoritmi, nell'Appendice B abbiamo introdotto alcune nozioni basilari su come calcolare la complessità.

La seconda parte è stata divisa in due sottosezioni, una riguardante gli algoritmi fondamentali della Teoria computazionale dei numeri, l'altra è una rassegna di alcuni tra i più importanti algoritmi per fattorizzare interi composti, al fine di confrontarli al metodo di Lenstra [11] che usa le curve ellit-

tiche.
Subito dobbiamo ricordare il teorema che è alla base della fattorizzazione.

Teorema 2.0.4 (Teorema Fondamentale dell'aritmetica).

Sia n un intero maggiore di 1. Allora n si può fattorizzare nel prodotto di un numero finito di primi $p_1, \ldots, p_s$:

$$n = p_1^{h_1} p_2^{h_2} \cdots p_s^{h_s},$$

dove p_j, con $j = 1, \ldots, s$, sono tutti distinti, gli esponenti h_j sono positivi. Inoltre tale fattorizzazione è unica a meno dell'ordine dei primi $p_1, \ldots, p_s$.

Abbiamo iniziato con l'*Algoritmo dei quadrati successivi* (§2.1.1), che si usa per calcolare $a^k (\mathrm{mod} n)$ con $k \in \mathbb{N}$. Il metodo consiste nel calcolare l'espansione k in base 2 e calcolare le varie potenze $a^{2^j} \bmod n$; alla fine avremo $a^k \bmod n$ con un numero di operazioni in $\mathbb{Z}/n\mathbb{Z}$ al più pari a $2 \lg_2 k$.
Famosissimo è l'*Algoritmo di Euclide* per calcolare il massimo comun divisore (§2.1.2). Si basa sulla divisione di due interi a e b e procede con successive divisioni tra i resti fino ad arrivare ad una divisione con resto zero; il resto precedente a quello nullo sarà il massimo comune divisore tra a e b e quindi il più grande intero che divide entrambi. Il numero di divisioni successive è pari al più a $2 \lg_2 \max\{2a, 2b\}$.
Per quanto riguarda gli algoritmi di classe esponenziale abbiamo iniziato con l'antico *Crivello di Eratostene* (§2.2.1) che ha, come scopo, quello di trovare tutti i primi minori di un certo numero n. Si basa sulla cancellazione di tutti i primi conosciuti e dei loro multipli fino ad arrivare al più grande primo minore di $\sqrt{n}$; a quel punto l'algoritmo termina. Infatti tutti gli interi rimasti sono esattamente i primi.
Il *Metodo ρ di Pollard* [15] (§2.2.2) è una motivazione per introdurre il problema del paradosso del compleanno. Infatti grazie ad esso potremmo in seguito spiegare un'ottimizzazione del metodo di Lenstra.

Il metodo su cui Lenstra [11] si ispirò è il *Metodo* $(p-1)$ *di Pollard* [14] (§2.2.3) in cui si cerca di fattorizzare n calcolando il massimo comun divisore tra n e $a^k - 1(\mathrm{mod} n)$. Il metodo funziona se $p-1$ è il prodotto di primi ragionevolmente piccoli, altrimenti il tempo di esecuzione è nella pratica troppo elevato.
Fino ad ora abbiamo trattato degli algoritmi con una complessità esponenziale, mentre è utile accennare ad alcuni importanti metodi di fattorizzazione con una complessità sub-esponenziale. Il metodo di Lenstra è anch'esso in tale classe.
Di conseguenza ci siamo occupati del *Metodo del crivello quadratico* (§2.3.1), ideato da Pomerance [16], che si basa sull'idea di determinare due interi x, y tali che $x \neq y(\mathrm{mod} n)$ ma $x^2 \equiv y^2(\mathrm{mod} n)$. In tal modo $\gcd(x-y, n)$ è un fattore proprio di n.
Sulla stessa idea si basa il *Metodo del crivello del campo numerico* (§2.3.2) perchè anch'esso si prefigge di ottenere una relazione della forma $x^2 \equiv y^2(\ \mathrm{mod}\ n)$. In esso, però, dobbiamo utilizzare un anello di interi algebrici A e un omomorfismo ϕ, costruito ad hoc, in modo tale che se $f(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$ definisce l'anello di interi algebrici, allora $\phi : A \to \mathbb{Z}/nZ$ ed è fissata $m \in \mathbb{Z}/n\mathbb{Z}$ tale che $f(m) \equiv 0(\mathrm{mod} n)$.

La terza parte è il nucleo centrale della nostra tesi; in essa abbiamo trattato il *Metodo di Lenstra fattorizzazione degli interi.* Questo metodo usa l'idea del metodo $(p-1)$ di Pollard, con la differenza che in esso lavoriamo con l'ordine del gruppo dei punti razionali delle curve ellittiche su $\mathbb{F}_p$ e non sul gruppo $(\mathbb{Z}/p\mathbb{Z})^*$ che ha appunto $p-1$ elementi. Il risultato sarà un metodo molto più veloce. Prima di studiare l'algoritmo nel particolare dobbiamo enunciare un importante risultato di Teoria analitica dei numeri, che useremo più avanti e che determina la probabilità di riuscita dell'algoritmo. Nella presente forma l'enunciato può essere trovato sulle dispense on-line di Poonen [17].

Teorema 3.1.1 (Canfield, Erdős, Pomerance [5]). *La probabilità che un intero preso a caso in* $[1, x]$ *sia* $L(x)^{\alpha}$*-liscio è pari a*

$$L(x)^{-\frac{1}{2a}+o(1)} \text{ per } x \to \infty,$$

dove diremo che $n \in \mathbb{N}$ *è* y*-liscio se tutti i divisori primi di* n *sono minori di* y *e dove* $L(x) = e^{\sqrt{\ln x \ln \ln x}}$.

Dato n composto, $\gcd(n, 6) = 1$ e n non una potenza, vogliamo trovare un fattore non banale di n.
Scelto B_1 relativamente grande, dobbiamo scegliere la curva ellittica su cui applicare il metodo; scelti $x, y, a \in [0, n-1]$ consideriamo $b = y^2 - x^3 - ax \bmod n$. La prima condizione da verificare è che $\gcd(4a^3+27b^2, n) = 1$ cioè che sia diverso da n stesso e non sia un divisore proprio di n (nella qual cosa avremmo trovato già un fattore di n). Infatti, se il massimo comun divisore fosse uguale ad n dovremmo scegliere un'altra curva. Quindi, scelti a, x, y possiamo lavorare con $E_{(a,b)}(\mathbb{Z}/n\mathbb{Z})$ dove $E_{(a,b)}(\mathbb{Z}/n\mathbb{Z})$ è la pseudocurva di equazione $y^2 = x^3+ax+b$ definita sull'anello $\mathbb{Z}/n\mathbb{Z}$ e $P = (x, y)$. Supponiamo che $p_1, p_2, \ldots, p_k$ siano tutti primi minori o uguali a B_1 per $1 \leq i \leq k$. Dobbiamo calcolare il più grande intero a_i tale che $p_i^{a_i} \leq B_1$, con $1 \leq i \leq k$. Trovati tali a_i calcoliamo

$$\begin{cases} P_1 = [p_1^{a_1}]P, \\ \\ P_j = [p_j^{a_i}]P \quad \text{per } j = 2, \ldots, s, \end{cases}$$

per $1 \leq j \leq k$ tramite la moltiplicazione nelle curve ellittiche, che equivale a sommare P con se stesso p_j volte. Nel fare ciò ci troveremo a calcolare anche degli inversi mod p. Se questo non fosse possibile, cioè se per applicare le formule ci trovassimo a voler calcolare l'inverso di un elemento $d \in \mathbb{Z}/n\mathbb{Z}$ con $g = \gcd(n, d) \neq 1$, allora sappiamo di aver ottenuto una potenza di P che dà luogo al "'punto all'infinito modulo un divisore primo di n'". Allora

se g è diverso da n abbiamo trovato un fattore non banale di n, altrimenti se $g = n$ incrementiamo B_1 o scegliamo una nuova curva, passando ad una nuova iterazione successiva.
Chiaramente B_1 non può essere scelto casualmente ma dobbiamo avere almeno una sua approssimazione. Attraverso numerosi calcoli abbiamo trovato che B_1 deve essere circa $\exp((\sqrt{2} + o(1))\sqrt{\ln p \ln \ln p})$.
Per quanto riguarda la complessità dell'algoritmo abbiamo dimostrato che è $\mathcal{O}(w(\lg B_1)M(N))$, con w il numero delle iterazioni dell'algoritmo per ogni singola curva e $M(n)$ il limite superiore per il tempo di esecuzione di una moltiplicazione in $\mathbb{Z}/n\mathbb{Z}$.
Enunciamo un corollario, tratto da Lenstra [11], che approssima la probabilità di successo dell'algoritmo:

Proposizione 3.2.6. *Esiste una costante c positiva con la seguente proprietà. Siano $n, w, B_1 \in \mathbb{Z}$ tali che n abbia almeno due primi divisori distinti maggiori di 3 e tali che il più piccolo divisore primo p di n per cui $p > 3$ soddisfa $p \leq B_1$. Poniamo*

$$u = \{s \in \mathbb{Z} \bmod |s - (p+1)| < \sqrt{p},\ \forall q | s,\ q \leq w\}$$

Allora il numero N di terne $(a, x, y) \in (\mathbb{Z}/n\mathbb{Z})^3$ per cui l'algoritmo ha successo nel trovare un divisore non banale di n soddisfa

$$\frac{N}{n^3} > \frac{c}{\log p} \frac{u-2}{2\left[\sqrt{p}\right] + 1}.$$

Quanto appena descritto è l'algoritmo di Lenstra base con l'analisi della sua complessità e la descrizione della sua probabilità di successo. Nel corso degli anni ci sono stati notevoli miglioramenti ed ottimizzazioni di questo algoritmo, come per esempio:

1. Utilizzare speciali parametrizzazioni per ottenere facilmente curve in maniera casuale.

2. Scegliere curve con ordine conosciuto che sia divisibile da 12 e da 16.

3. Diminuire le operazioni ellittiche tra interi del passo 3.

4. Applicare algoritmi più veloci al passo 3.

Da queste nuove scoperte è stato possibile arricchire l'algoritmo base, che è composto da quattro passi, con un nuovo passo che sfrutta le ipotesi di P. Montgomery sulla possibilità di fare i calcoli senza calcolare le coordinate y negli inversi. Infatti nell'aritmetica di P. Montgomery un punto P viene rappresentato da una coppia $P = (x, z)$ e tramite opportuni cambiamenti di variabili vengono rese più veloci le operazioni del calcolo della somma di due punti in $E(\mathbb{F}_p)$.

Bibliografia

[1] A. O. L. Atkin and F. Morain. Finding suitable curves for the elliptic curve method of factorization. *Math. Comp.*, 60(201):399–405, 1993.

[2] I. F. Blake, G. Seroussi, and N. P. Smart. *Elliptic curves in cryptography*, volume 265 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2000. Reprint of the 1999 original.

[3] R. Brent. Some integer factorization algorithms using elliptic curves, 1986.

[4] R.P. Brent and R.E. Crandall. Three new factors of Fermat numbers. *Math. Comp.*, 69(231):1297–1304, 2000.

[5] E. R. Canfield, Paul Erdős, and Carl Pomerance. On a problem of Oppenheim concerning "factorisatio numerorum". *J. Number Theory*, 17(1):1–28, 1983.

[6] Henri Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.

[7] Richard Crandall and Carl Pomerance. *Prime numbers*. Springer-Verlag, New York, 2001. A computational perspective.

[8] Ernest S. Croot III. On non-intersecting arithmetic progressions.

[9] Pollard J.M. and A.K. Lenstra. A number field sieve.

[10] Neal Koblitz. *A course in number theory and cryptography.* Springer, second edition, 1994.

[11] H. W. Lenstra, Jr. Factoring integers with elliptic curves. *Ann. of Math. (2)*, 126(3):649–673, 1987.

[12] Peter L. Montgomery. An FFT extension of the elliptic curve method of factorization. Technical report, University of California,Los Angeles, August 1992. A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in mathematics.

[13] F. Pappalardi. Note di crittografia, 2002.

[14] J. M. Pollard. Theorems on factorization and primality testing. *Proc. Cambridge Philos. Soc.*, 76:521–528, 1974.

[15] J. M. Pollard. A Monte Carlo method for factorization. *Nordisk Tidskr. Informationsbehandling (BIT)*, 15(3):331–334, 1975.

[16] Carl Pomerance. Smooth numbers and the quadratic sieve.

[17] Bjorn Poonen. Elliptic curves. Technical report, Mathematical Sciences Research Institute, University of California,Berkeley, August 2000.

[18] René Schoof. Counting points on elliptic curves over finite fields. *J. Théor. Nombres Bordeaux*, 7(1):219–254, 1995. Les Dix-huitièmes Journées Arithmétiques (Bordeaux, 1993).

[19] Edoardo Sernesi. *Geometria vol. 1.* Bollati Boringhieri, first edition, 1989.

[20] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 199? Corrected reprint of the 1986 original.

[21] Douglas R. Stinson. *Cryptography*. CRC Press Series on Discrete Mathematics and its Applications. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2002. Theory and practice.

[22] H. Suyama. Informal preliminary report.