Il gruppo moltiplicativo $\mathbf{Z}_p^*$ delle classi resto modulo un primo p è un gruppo ciclico.

Definizione (Logaritmo discreto). Sia p un numero primo e sia $\bar{a}$ una radice primitiva in $\mathbf{Z}_p^*$. Sia $\bar{y} \in \mathbf{Z}_p^*$. Il *logaritmo discreto di* $\bar{y}$ *in base* $\bar{a}$ è un intero m per cui vale

$$\bar{a}^m \equiv \bar{y} \mod p.$$

Si indica con $m = \log_{\bar{a}} \bar{y}$ (o semplicemente con $m = \log \bar{y}$ se la base $\bar{a}$ è chiara dal contesto). Il logaritmo discreto $m = \log_{\bar{a}} \bar{y}$ è unico modulo $(p-1)$.

Osservazione.
Sia p un numero primo e sia $\bar{a}$ una radice primitiva in $\mathbf{Z}_p^*$.

(a) Il logaritmo discreto in base $\bar{a}$ è unico modulo $p-1$, ossia $\bar{y} \equiv \bar{a}^m \equiv \bar{a}^n \bmod p$ se e solo se $m \equiv n \bmod p-1$.

Dim: Se $m = n + k(p-1)$, con $k \in \mathbf{Z}$, vale $\bar{a}^m \equiv \bar{a}^n \cdot \bar{a}^{k(p-1)}$. Poiché per il Piccolo Teorema di Fermat $\bar{a}^{p-1} \equiv \bar{1} \bmod p$, allora anche $\bar{a}^{k(p-1)} \equiv \bar{1} \bmod p$ e $\bar{a}^m \equiv \bar{a}^n \bmod p$.
Supponiamo viceversa che $\bar{a}^m \equiv \bar{a}^n \bmod p$. Allora $\bar{a}^{m-n} \equiv \bar{1} \bmod p$. Poiché $\bar{a}$ è una radice primitiva, l'ordine di $\bar{a}$ in $\mathbf{Z}_p^*$ è uguale a $p-1$. Ne segue che $m - n = k(p-1)$, per $k \in \mathbf{Z}$.

(b) Il logaritmo di un prodotto è uguale alla somma dei logaritmi dei fattori (modulo $p-1$):

$$\log_{\bar{a}}(\bar{x} \cdot \bar{y}) = \log_{\bar{a}} \bar{x} + \log_{\bar{a}} \bar{y}, \qquad \forall \bar{x},\ \bar{y} \in \mathbf{Z}_p^*.$$

Dim: Siano $m = \log_{\bar{a}} \bar{x}$ ed $n = \log_{\bar{a}} \bar{y}$. Ciò significa che $\bar{a}^m \equiv \bar{x}$ ed $\bar{a}^n \equiv \bar{y}$ modulo p. Ne segue che

$$\bar{a}^m \cdot \bar{a}^n \equiv \bar{a}^{m+n} \equiv \bar{x} \cdot \bar{y} \equiv \overline{xy} \mod p,$$

ed in particolare $m + n = \log_{\bar{a}} \overline{xy} \mod (p-1)$.

(c) Sia p un numero primo. Siano $\bar{a}$ ed $\bar{b}$ due radici primitive in $\mathbf{Z}_p^*$. Allora vale la relazione

$$\log_{\bar{a}} \bar{x} = \log_{\bar{b}} \bar{x} / \log_{\bar{b}} \bar{a}, \qquad \forall \bar{x} \in \mathbf{Z}_p^*.$$

Dim: Sia $m = \log_{\bar{a}} \bar{x}$, per cui vale $\bar{x} \equiv \bar{a}^m$. Estraendo il logaritmo in base $\bar{b}$ di ambo i termini, troviamo $\log_{\bar{b}} \bar{x} = m \log_{\bar{b}} \bar{a} = \log_{\bar{a}} \bar{x} \log_{\bar{b}} \bar{a}$, da cui segue la relazione cercata.

Osservazione.
Se $\bar{a}$ e $\bar{b}$ sono generatori di $\mathbf{Z}_p^*$, allora $m := \log_{\bar{b}} \bar{a}$ è invertibile modulo $p-1$. Infatti, se $\bar{b}$ ha ordine $p-1$ in $\mathbf{Z}_p^*$, allora $\bar{a} = \bar{b}^m$ ha ordine $p-1$ se e solo se $\gcd(m, p-1) = 1$.

Osservazione.
Sia p un numero primo e sia $\bar{a}$ una radice primitiva in $\mathbf{Z}_p^*$. Si ha

$$\log_{\bar{a}} \bar{a} = 1, \qquad \log_{\bar{a}} \bar{1} = 0, \qquad \log_{\bar{a}} \overline{-1} = \frac{p-1}{2}.$$

Dim: I primi due logaritmi seguono immediatamente dalla definizione.
Per dimostrare che $\log_{\bar{a}} \overline{-1} = \frac{p-1}{2}$, dobbiamo verificare che $\bar{a}^{\frac{p-1}{2}} \equiv \overline{-1} \bmod p$. Dal Piccolo Teorema di Fermat abbiamo che $\bar{a}^{p-1} \equiv 1 \bmod p$. Poiché p è primo ed il polinomio $X^2 - 1 \in \mathbf{Z}_p[X]$ ha ± 1 come uniche radici in $\mathbf{Z}_p$, segue che $\bar{a}^{\frac{p-1}{2}} \equiv \bar{1} \bmod p$ oppure $\bar{a}^{\frac{p-1}{2}} \equiv \overline{-1} \bmod p$. D'altra parte l'ordine di $\bar{a}$ è uguale a $p-1$, per cui necessariamente $\bar{a}^{\frac{p-1}{2}} \equiv \overline{-1} \bmod p$.

Il calcolo dell'indice.

Sia $p > 0$ un primo. Sia $a \in \mathbf{Z}_p^*$ una radice primitiva.
Diamo una descrizione schematica dell'algoritmo detto *calcolo dell'indice* per risolvere il logaritmo discreto in $\mathbf{Z}_p^*$.
(Per semplicità di notazione, da ora in poi omettiamo la barra che indica le classi resto).

(1)
Fissare un ordine di B*-smoothness*, dove B è un intero positivo.
Ad esso è associata una *factor base* $F = \{l_i \text{ primi}, \ l_i \le B\}$, data dall'insieme dei numeri primi minori o uguali a B.
La cardinalità di F si stima con $\alpha := \#F \sim B/\ln B$.

(2)
Precalcolare i logaritmi $\log_a l_i$*, per tutti i primi* $l_i \in F$*.*
Per far ciò, prendere a caso un intero della forma $\prod l_i^{e_i} > p$, con $l_i \in F$ e $e_i \in \mathbf{N}$, e ridurlo modulo p. Sia

$$r \equiv \prod l_i^{e_i} \mod p \qquad (\#)$$

la classe così ottenuta.
Tentare la fattorizzazione di r col metodo ρ di Pollard o col metodo delle curve ellittiche, particolarmente efficaci nel determinare i fattori piccoli. (Se dopo un certo numero di iterazioni di tali algoritmi r non è stato ancora fattorizzato, verosimilmente non è B-smooth).
Se la classe r non è B-smooth, si scarta e si riparte da un altro intero a caso $\prod l_i^{e_i}$, con $p_i \in F, \ e_i \in \mathbf{N}$.
Se la classe r è B-smooth e si fattorizza come $r = \prod l_i^{f_i}$, con $l_i \in F$ ed esponenti $f_i \in \mathbf{Z}$, allora la relazione

$$\prod l_i^{e_i} \equiv \prod l_i^{f_i} \mod p$$

determina una relazione lineare fra i logaritmi dei primi della factor base

$$\sum e_i \log_a l_i \equiv \sum f_i \log_a l_i \mod (p-1) \quad \Leftrightarrow$$

$$\sum (e_i - f_i) \log_a l_i \equiv 0 \mod (p-1).$$

Tale relazione significa che il vettore $(\log_a l_1, \ldots, \log_a l_\alpha)$ è soluzione dell'equazione lineare omogenea

$$\sum (e_i - f_i) x_i \equiv 0 \mod (p-1). \qquad (*)$$

Osserviamo che l'equazione (*) è soddisfatta dal vettore dei logaritmi $(\log_b l_1, \ldots, \log_b l_\alpha)$, rispetto ad ogni radice primitiva di $\mathbf{Z}_p^*$. Tale vettore è infatti proporzionale a $(\log_a l_1, \ldots, \log_a l_\alpha)$, con fattore di proporzionalià $\log_b a$ (vedi Osservazione (c)).

Ripetendo questa procedura fino ad ottenere almeno α relazioni lineari fra $\log_a l_1, \ldots, \log_a l_\alpha$, avremo costruito un sistema lineare omogeneo a coefficienti in $\mathbf{Z}_{p-1}$, di almeno α equazioni in α incognite

$$\begin{cases} (e_1^1 - f_1^1)x_1 + \ldots + (e_\alpha^1 - f_\alpha^1)x_\alpha \equiv 0 \\ (e_1^2 - f_1^2)x_1 + \ldots + (e_\alpha^2 - f_\alpha^2)x_\alpha \equiv 0 \\ \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\ (e_1^\alpha - f_1^\alpha)x_1 + \ldots + (e_\alpha^\alpha - f_\alpha^\alpha)x_\alpha \equiv 0 \end{cases} \mod (p-1)$$

di cui il vettore $(\log_a l_1, \ldots, \log_a l_\alpha)$ è soluzione *per costruzione.*

Risolvere il sistema mediante l'eliminazione di Gauss, tenendo conto del fatto che i coefficienti sono in $\mathbf{Z}_{p-1}$ che non è un campo.

(Prima di dividere un'equazione per un coefficiente, bisogna accertarsi che sia invertibile in $\mathbf{Z}_{p-1}$. Per questa ragione in generale si costruiscono un pò più di α relazioni).

Tipicamente le soluzioni del sistema formano uno spazio di dimensione 1, generato dal vettore $(\log_a l_1, \ldots, \log_a l_\alpha)$ dei logaritmi dei primi della factor base, ossia sono della forma

$$\lambda(\log_a l_1, \ldots, \log_a l_\alpha), \quad \lambda \in \mathbf{R},$$

dove il parametro λ è determinato dalla scelta della base rispetto alla quale è definito il logaritmo.
Ad esempio, se $a = l_1$, ossia la radice primitiva a coincide col primo l_1 della factor base, allora $\log_{l_1} l_1 = 1$ e le soluzioni cercate sono date dal vettore $(1, \log_{l_1} l_2, \ldots, \log_{l_1} l_\alpha)$.

(3)
Calcolare $\log_a x$, *per* $x \in \mathbf{Z}_p^*$, *a partire dai valori dei logaritmi* $\log_a l_1, \ldots, \log_a l_\alpha$ *ottenuti al passo precedente.*

Prendere a caso un prodotto della forma $x \cdot \prod l_i^{m_i} > p$, con $l_i \in F$ e $m_i \in \mathbf{N}$. Ridurlo modulo p e tentare di fattorizzare la classe trovata

$$s \equiv x \cdot \prod l_i^{m_i} \mod p.$$

Se s non è B-smooth, si riparte con un altro prodotto.
Se s è B-smooth e si fattorizza come $s \equiv \prod l_i^{n_i}$, con $l_i \in F$ ed esponenti $n_i \in \mathbf{Z}$, si ottiene

$$x \cdot \prod l_i^{m_i} \equiv \prod l_i^{n_i} \mod p,$$

da cui estraendo il logaritmo in base a e usando i valori dei logaritmi determinati al passo precedente, si ricava il logaritmo cercato

$$\log_a x \equiv \sum (n_i - m_i) \log_a l_i \mod (p-1).$$

La complessità dell'algoritmo.

Siano p un primo ed $a \in \mathbf{Z}_p^*$ una radice primitiva.
Sia B un ordine di B-smoothness;
sia $F = \{l_1, \ldots, l_\alpha\}$ la factor base corrispondente, con $\alpha \sim B/\ln B$.

(1) Costo di una relazione.

- Random $\; r \equiv \prod l_i^{e_i} \mod p$: $\qquad \mathcal{O}(\ln e_i \ln^2 p) \sim \mathcal{O}(\ln^3 p)$

In realtà, per cercare una relazione è sufficiente elevare ad un esponente casuale e_i anche un solo primo $l_i \in F$ alla volta. La potenza ottenuta dovrà essere maggiore di p, per poter essere ridotta modulo p, e al tempo stesso *vicina a zero* modulo p, per avere maggiori possibilità che sia B-smoth. Dalle relazioni

$$p < l_i^{e_i} < 2p \quad \Leftrightarrow \quad \ln p < e_i \log l_i < \ln(2p),$$

segue che l'esponente si può maggiorare con $\ln p$.

- controllare se $\; r \equiv \prod l_i^{e_i} \mod p \;$ è B-smooth: $\qquad \mathcal{O}(\sqrt{B} \ln^3 p) \quad$ (con Pollard ρ)

Per far questo è necessario fattorizzare r, o almeno capire se r ha fattori maggiori di B. Conviene usare algoritmi particolarmente efficaci nell'individuare i fattori piccoli, come Pollard ρ oppure ECM.
Consideriamo ad esempio Pollard ρ: se dopo $\sqrt{B}$ iterazioni, il numero r non è stato fattorizzato, probabilmente non è B-smooth e si scarta. La complessità dell'operazione è dell'ordine di $\mathcal{O}(\sqrt{B} \ln^3 p)$.
Se invece di Pollard ρ, si usano il metodo per tentativi o il metodo delle curve ellittiche la complessità dell'operazione è data rispettivamente da $\mathcal{O}(B \ln p)$ e da $\mathcal{O}\left(e^{\sqrt{2 \ln B \ln \ln B}}\right)$.

- il numero di tentativi necessari ad ottenere una relazione è stimabile con w^w, dove $w = \frac{\ln p}{\ln B}$.

Il numero di interi B-smooth nell'intervallo $[1, p]$ è stimato dalla funzione di Dickman $\Psi(p, B) \sim p w^{-w}$, dove $w = \frac{\ln p}{\ln B}$. Dunque per avere probabilità positiva di ottenere una classe r che sia B-smooth, il numero di tentativi si stima dell'ordine di w^w.

Da quanto detto segue che il costo totale di una relazione si può stimare con

$$\mathcal{O}\left((\ln^3 p + \sqrt{B} \ln^3 p) w^w\right) \;\sim\; \mathcal{O}\left(\sqrt{B} \ln^3 p \cdot w^w\right), \qquad w = \frac{\ln p}{\ln B}$$

(2) Costo di α relazioni (per semplicità stimiamo $\alpha \sim B$):

$$\mathcal{O}\left(B \cdot \sqrt{B} \ln^3 p \cdot w^w\right) \sim \mathcal{O}\left(B^{3/2} w^w \log^3 p\right), \qquad w = \frac{\ln p}{\ln B}$$

(3) Risolvere un sistema lineare $\alpha \times \alpha$ a coefficienti in $\mathbf{Z}_{p-1}$: $\qquad \mathcal{O}\left(B^3 \ln^2 p\right)$

Per esempio con l'eliminazione di Gauss.

(4) Calcolare $\log_a x$ a partire da $\log_a l_1, \ldots, \log_a l_\alpha$: $\qquad \mathcal{O}\left(\sqrt{B} \ln^3 p \cdot w^w\right), \qquad w = \frac{\ln p}{\ln B}$

Equivale al costo di una relazione.

TOTALE:

$$\mathcal{O}\left(B^{3/2}w^w\ln^3 p\right)+\mathcal{O}\left(B^3\ln^2 p\right)+\mathcal{O}\left(\sqrt{B}\ln^3 p\cdot w^w\right) \quad\sim\quad \mathcal{O}\left(B^{3/2}w^w\ln^3 p+B^3\ln^2 p\right),\quad w=\frac{\ln p}{\ln B}.$$

Per determinare i parametri ottimali, studiamo la funzione

$$F(w)=B^{3/2}w^w\ln^3 p=p^{\frac{3}{2w}}w^w\ln^3 p,$$

o meglio il suo logaritmo

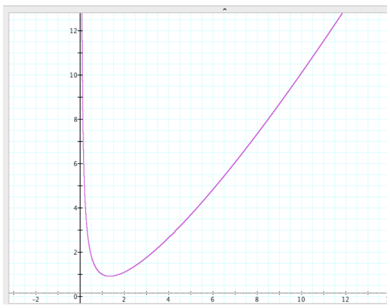
$$G(w)=\ln F(w)=\frac{3}{2}\frac{\ln p}{w}+w\ln w+\ln(\ln^3 p)\sim\frac{3}{2}\frac{\ln p}{w}+w\ln w,$$

al variare di $w=\frac{\ln p}{\ln B}$ in $\mathbf{R}^+$.

La funzione G è una funzione del tipo

$$\phi_a(w)=\frac{a}{w}+w\ln w,\qquad a>>0,$$

dove nel nostro caso $a=\frac{3}{2}\ln p$. (Vedi Nota "Smooth number estimates")



Il grafico approssimativo della funzione $G(w)=\frac{3}{2}\frac{\ln p}{w}+w\ln w$.

La funzione G ha un unico punto di minimo in $w_0\sim\sqrt{\frac{3\ln p}{\ln(\ln(p))}}$, a cui corrisponde l'ordine di B-smoothness ottimale

$$B_{best}\sim e^{\sqrt{\frac{\ln p\ln(\ln(p))}{3}}}.$$

Per questo questo valore dell'ordine di B-smoothness, la complessità del calcolo per ottenere le prime α relazioni è data da

$$(B_{best})^{3/2}w_0^{w_0}\ln^3 p\ \sim\ e^{\sqrt{3\ln p\ln(\ln(p))}}\ln^3 p.$$

Osservazione. Per i parametri ottimali risulta che $B^{3/2}$ e w^w sono all'incirca dello stesso ordine di grandezza, ossia $B^{3/2}\sim w^w$. In particolare il lavoro per ottenere le α relazioni iniziali è $\mathcal{O}\left(B^3\ln^3 p\right)$, che domina il lavoro richiesto dalla risoluzione del sistema.

Osservazione. Se per fattorizzare le classi resto modulo p, invece di Pollard ρ, si usano il metodo per tentativi o il metodo delle curve ellittiche la complessità del logaritmo discreto rimane comunque subesponeziale in $\ln p$.

Osservazioni finali.

(a) Nei sistemi crittografici basati sul logaritmo discreto, il primo p è scelto della forma $p = 1 + mQ$, dove m è un intero piccolo e Q è un grosso primo. Ad esempio $p = 1 + 2Q$. Se $p - 1$ fosse prodotto di primi piccoli q_i, tramite il Teorema Cinese del Resto, il logaritmo discreto in $\mathbf{Z}_p^*$ sarebbe riconducibile al logaritmo discreto in gruppi ciclici piccoli e quindi facilmente risolvibile.
Se $p - 1 = 2Q$, il gruppo ciclico $\mathbf{Z}_p^*$ contiene $\varphi(p-1) = (p-1) \cdot \frac{1}{2}(1 - \frac{1}{Q})$ radici primitive (quasi metà degli elementi).

(b) Per rendere l'algoritmo più efficiente nella ricerca delle relazioni conviene fare quanto segue. Sia r la classe resto modulo p ottenuta in (#): esprimere $r = \frac{z}{x}$ come rapporto fra due interi dell'ordine di grandezza di $\sqrt{p}$ (mentre r è dell'ordine di grandezza di p). Questo si può fare mediante l'algoritmo di Euclide esteso usato per verificare che $\gcd(r, p) = 1$, arrestato a metà strada. Partendo da

$$1 \cdot r + 0 \cdot p = r, \quad 0 \cdot r + 1 \cdot p = p, \ \ldots \ \ldots,$$

si arriva ad un'espressione del tipo

$$x \cdot r + y \cdot p = z,$$

dove x, y, z sono dell'ordine di grandezza di $\sqrt{p}$. Da essa si ottiene

$$r \equiv \frac{z}{x} \mod p.$$

Il vantaggio è che z e x sono più piccoli (hanno metà delle cifre di p!!), e dunque più facili da fattorizzare. Se entrambi sono B-smooth anche r lo è, e risulta

$$r = \frac{z}{x} = \frac{l_1^{e_1} \ldots l_\alpha^{e_\alpha}}{l_1^{f_1} \ldots l_\alpha^{f_\alpha}} = l_1^{e_1 - f_1} \ldots l_\alpha^{e_\alpha - f_\alpha}.$$

(c) Nel risolvere un sistema lineare omogeneo modulo $2Q$ si usa il seguente fatto: ogni equazione

$$a_1x_1 + \ldots + a_nx_n \equiv 0 \mod 2Q \quad \Leftrightarrow \quad \begin{cases} a_1x_1 + \ldots + a_nx_n \equiv 0 \mod Q \\ a_1x_1 + \ldots + a_nx_n \equiv 0 \mod 2. \end{cases}$$

Sia $(x_1^0, \ldots, x_n^0)$ una soluzione del sistema modulo Q, dove

$$x_i^0 \equiv \log_a l_i \mod Q.$$

Una soluzione $(y_1^0, \ldots, y_n^0)$ del sistema modulo 2 si ottiene cosi':

$$y_i = \begin{cases} 0 & \text{se } l_i \text{ è un quadrato modulo } p \\ 1 & \text{se } l_i \text{ non è un quadrato modulo } p \end{cases}.$$

Adesso per trovare una soluzione $(z_1^0, \ldots, z_n^0)$ modulo $2Q$ si applica il teorema cinese del resto ad ognuna delle coordinate

$$\begin{cases} z_i \equiv x_i^0 \mod Q \\ z_i \equiv y_i^0 \mod 2. \end{cases}$$