Appunti di Finite Fields and Symmetric Cryptography

michele@tumbolandia.net

10 giugno 2015

Indice

Capitolo 1

Cazzilli sui campi

Si consideri F campo.

Teorema 1.1 (Formula di Interpolazione di Lagrange)**.** *Siano $\{a_i\}_{i=0}^n$ e $\{b_i\}_{i=0}^n$ due collezioni di $n+1$ elementi, in cui gli a_i sono tutti distinti e $n \geq 0$. Allora esiste* unico *polinomio $f \in F[x]$ di grado $\deg f \leq n$ tale che $f(a_i) = b_i$:*

$$f(x) = \sum_{i=0}^{n} b_i \prod_{\substack{k=0 \\ k \neq i}}^{n} (a_i - a_k)^{-1}(x - a_k).$$

Dimostrazione. Per ogni addendo i-esimo si ha:

$$x = a_i \iff \prod_{\substack{k=0 \\ k \neq i}}^{n} (a_i - a_k)^{-1}(a_i - a_k) = 1$$

$$x \neq a_i \iff \prod_{\substack{k=0 \\ k \neq i}}^{n} (a_i - a_k)^{-1}(a_i - a_k) = 0$$

Poiché $\exists k \neq i$ s.t. $x = a_k$. □

Definizione 1.2 (Polinomio Simmetrico)**.** Un polinomio $f \in F[x_1, \ldots, x_n]$ è detto *simmetrico* sse

$$f(x_{j_1}, \ldots, x_{j_n}) = f(x_{i_1}, \ldots, x_{i_n})$$

$\forall i_k, j_k$ famiglie di permutazioni di n elementi.

Definizione 1.3. Il k-esimo *polinomio simmetrico elementare* su $F[x_1, \ldots, x_n]$ è σ_k in:

$$\begin{aligned}
\sigma_1 &= x_1 + \cdots + x_n \\
\sigma_2 &= x_1x_2 + x_1x_3 + \cdots + x_1x_n + \cdots + x_{n-1}x_n \\
&\vdots \\
\sigma_k &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k} \\
&\vdots \\
\sigma_n &= x_1x_2 \cdots x_n
\end{aligned}$$

Teorema 1.4 (Teorema fondamentale sui polinomi simmetrici)**.** *Tutti i polinomi simmetrici sono della forma $h(\sigma_1 \ldots \sigma_n)$ per un $h \in K[x_1 \ldots x_n]$ unicamente determinato.*

Teorema 1.5 (Identità di Newton)**.** *Dato un polinomio $f = \sum_i^m a_i x^i$, allora gli a_i sono σ_i delle radici del polinomio, dove i σ_i è l'i-esimo polinomio simmetrico elementare.*

Più precisamente, si definisca la successione

$$s = \begin{cases} s_0 = n \\ s_k = s_k(x_1, \ldots, x_n) = x_1^k + \cdots + x_n^k \end{cases}$$

Allora

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 + \cdots + (-1)^{m-1}s_{k-m-1}\sigma_{m-1} + (-1)^m \frac{m}{n} s_{k-m}\sigma_m = 0$$

Dimostrazione. `TODO:??` □

Definizione 1.6 (Derivata)**.** Sia $f = a_0 + \cdots + a_n x^n \in F[x]$, allora la sua derivata è $f' = a_1 + 2a_2 + \cdots + na_n x^{n-1} \in F[x]$.

Teorema 1.7. *$b \in F$ è una radice multipla di f sse è radice di f ed f'.*

Dimostrazione.

- "$\implies$" b radice multipla $\implies f = (x-b)^2 g \implies f' = (x-b)[(x-b)g' + 2g]$.
- "$\impliedby$" b radice di $f, f' \implies (x-b) \mid f,\ (x-b) \mid f' \implies f = (x-b)g,\ f' = (x-b)h \implies f' = (x-b)g' + g = (x-b)h \implies (x-b)^2 \mid f$. □

Definizione 1.8 (Discriminante)**.** Sia $f = a_0(x-\alpha_1)\cdots(x-\alpha_n) \in F[x]$ di grado $\deg f \geq 2$ con α_i nel campo di spezzamento. Il *discriminante* di f è

$$D(f) = a_0^{2n-2} \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2.$$

Esempio 1.9. Il determinante di $f(x) = ax^2+bx+c = a(x-\alpha_1)(x-\alpha_2)$ è $D(f) = a^2(\alpha_1-\alpha_2)^2 = (\alpha_1+\alpha_2)^2 - 4\alpha_1\alpha_2 = b^2 - 4ac$.

Il determinante di $f(x) = ax^3+bx^2+cx+d = a(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ è $D(f) = a^4(\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2 = b^2c^2 - 4b^3d - 4ac^3 - 27a^2d^2 + 18abcd$.

Lemma 1.10. *$D(f) \in F$.*

Dimostrazione. Sia $f = a_0x^n + \cdots + a_n$.

$$\begin{aligned} D(f) &= a^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2 \quad \text{polinomio simmetrico} \\ &= h(\sigma_1(\alpha_1, \ldots, \alpha_k), \ldots, \sigma_n(\alpha_1, \ldots, \alpha_k)) \quad \text{per qualche } h \in K[x_1, \ldots, x_n] \text{ e } \sigma_k =^{??} (-1)^k a_0 a_k^{-1} \in K \\ &= h(-a_1a_0^{-1}, \ldots, (-1)^n a_n a_0^{-1}) \in K \end{aligned}$$

□

Esiste un modo per calcolarsi il determinante manipolando solo oggetti di K che sfrutta un attrezzo chiamato *risultante*.

Gruppi ciclici

Teorema 1.11 (Proprietà dei Gruppi Ciclici)**.** *Sia $G = \langle a \rangle$ gruppo ciclico di ordine n. Allora:*

(i) ogni sottogruppo di un gruppo ciclico è ancora ciclico;

(ii) ogni sottogruppo $\langle a^k \rangle$ di un gruppo ciclico do odine n ha ordine $n/(k,n)$;

(iii) per ogni divisore positivo dell'ordine di un gruppo ciclico, esiste uno e un solo sottogruppo con tale ordine;

(iv) per ogni divisore positivo f di n esistono esattamente $\varphi(n)$ elementi di ordine f in G;

(v) ogni gruppo ciclico ammette $\varphi(n)$ generatori.

Dimostrazione.

(i) Sia H sottogruppo, quindi nonvuoto, allora $\exists n > 0$ minimo s.t. $a^n \in H$ (poiché tanto $a^{-n} \in H \implies a^n \in H$) $\implies \forall a^s \in H,\ \ s = qn + r\ \wedge 0 \leq r < n \implies r = 0$;

(ii) Sia $H = \langle a^k \rangle \subset \langle a \rangle$ allora l'ordine di H è il minimo intero z tale per cui $n \mid kz \implies z = n/(k,m)$, infatti:

$$(a^k)^{\frac{n}{(k,m)}} = a^{\frac{kn}{(k,n)}} = a^{[k,n]} = 1$$

(iii) $d \mid n \implies (n,d) = d\ \wedge \langle a^d \rangle$ ha ordine n/d, e per ogni altro sottogruppo ciclico $\langle a^k \rangle$, se ha ugual ordine $n/d \implies d = (n,k)$ per il punto (ii) $\implies d \mid k \implies a^k \in a^d$, ma i due sottogruppi hanno anche stesso numero di elementi, quindi sono lo stesso per assioma di estensione.

(iv) Scrivo $n = df$, allora per (ii) $\langle a^k \rangle$ ha ordine $f \iff (k, n) = d$. Quindi tutti i sottogruppi di $\langle a \rangle$ aventi ordine f sono gli $\langle a^k \rangle$ s.t. $(k, n) = d \iff (h, f) = 1 \quad (k := dh)$.

(v) Immediato dal punto precedente. □

Capitolo 2

Struttura dei Campi Finiti

In questo capitolo con la notazione $K[\alpha]$ intendiamo la più piccola estensione di K contenente α. Ricordiamo che tale campo è isomorfo a $K[x]/(g)$ dove $g \in K[x]$ è il polinomio minimo di α algebrico su K (algebrico perché visto come spazio vettoriale abbiamo che $\{\alpha^i\}_i$ è una base e i suoi coeff. sono elementi di K, isomorfo perché il morfismo applicazione $\tau_\alpha : K[x] \to K[\alpha]$ ha $\ker \tau_\alpha = (g)$ e quindi per il primo teorema di isomorfismo $K[x]/(g) \simeq K[\alpha]$). In particolare, se $K[\alpha]$ è un campo, allora è uguale al suo spazio quoziente $K(\alpha)$.

Proposizione 2.1. *α è algebrico su $K \iff [K(\alpha) : K] < \infty$.*

Dimostrazione. Per " $\implies$ ": α algebrico su $K \implies K(\alpha) = K(x)/(f)$ dove f è il polinomio minimo di α su $K \implies \infty > \deg(f) = [K(\alpha) : K]$.

Per " $\impliedby$ ", si nota che che $K(\alpha)$ è lo spazio vettoriale generato da $\{\alpha\}_i^n$ dove $n = [K(\alpha) : K]$. □

Lemma 2.2. *F campo finito $\implies |F| = p^m$ dove p è un numero primo e m è il grado di F sul suo sottocampo primo.*

Dimostrazione. F campo finito $\implies F$ ha come caratteristica un numero primo. Infatti, F è un anello nonvuoto con caratteristica positiva ($\exists 1 \leq k < m$ s.t. $k1 = m1 = 0 \implies (k-m)1 = 0$) senza divisori dello 0, per cui se per assurdo la caratteristica p non fosse prima, esisterebbero $p > n, m > 1$ che lo dividono, quindi $nm1 = n1 \cdot m1 = 0$, che va in contraddizione con la definizione di caratteristica.

F ha come caratteristica un numero primo, quindi ha un sottocampo K di p elementi (i.e. $\mathbf{N}1$), quindi è uno spazio vettoriale su K (finito), di dimensione $m = [F : K]$, quindi ha p^m elementi. □

Teorema 2.3 (Eulero-Fermat)**.** *F campo finito di q elementi $\implies \forall a \in F \quad a^q = a$.*

Dimostrazione. $0^q = 0$. In F^* l'ordine di ogni elemento divide l'ordine del gruppo $q - 1$ per il Teorema di Lagrange. □

Dal Teorema di Eulero-Fermat segue immediatamente che il polinomio $x^q - x = \prod_{a \in F}(x - a)$ poiché ha al più q radici, e q radici stanno nel campo.

Teorema 2.4 (Esistenza e Unicità di Campi Finiti)**.** *Per ogni primo p e ogni intero $m \geq 1$ possiamo costruire un campo con $q = p^m$ elementi unico a meno di isomorfismi.*

Dimostrazione. (Esistenza) Consideriamo il campo di spezzamento F del polinomio $f = x^q - x$. Notiamo che il polinomio non ha radici doppie poiché $f' = -1$ ha grado 0 (Teorema 1.7). Consideriamo $S = \{\, \alpha \mid \alpha^q = \alpha \,\}$; esso è un campo poiché contiene $\mathbf{0}$, $\mathbf{1}$, $(a-b)^q = a^q - b^q \in S$ (quindi è un gruppo additivo), $(ab^{-1})^q = a^q(b^q)^{-}1 = ab^{-1} \in S$ (con $b \neq 0$, quindi è un gruppo moltiplicativo). Segue che S è un campo, contiene tutte e le sole radici del polinomio f, quindi è uguale a F, che pertanto ha q elementi.

(Unicità) I campi di spezzamento sono unici a meno di isomorfismo (per la Proprietà Universale dei Polinomi). □

Teorema 2.5 (Criterio del Sottocampo)**.** *Esiste un sottocampo K per F di p^m elementi sse F ha p^n elementi dove $m \mid n$.*

Dimostrazione. È chiaro che K ha p^m elementi, dove p è primo. Inoltre, F è un sovracampo di K, quindi se visto come spazio vettoriale, deve esser che $p^m || F| \implies m \mid n$.

D'altra parte, se $m|n$ allora $n = mk \implies x^n - 1 = (x^m - 1)(x^{m(k-1)} + \cdots + 1) \implies p^m - 1 | p^n - 1 \implies x^{p^m-1} - 1 | x^{p^n-1} - 1 \implies x^{p^m} - x | x^{p^n} - x \implies$ lo splitting field di $x^{p^m} - 1$, $\mathbb{F}_{q^m}$, è un sottocampo di $\mathbb{F}_{p^n}$ perché tutte le sue radici sono ivi contenute. □

Teorema 2.6. *Per ogni campo finito $\mathbb{F}_q$ il gruppo moltiplicativo $\mathbb{F}_q^*$ è ciclico, e quindi ha un generatore.*

Dimostrazione. Assumiamo $q \geq 3$.

Sia $h = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}$ una fattorizzazione di $h = q-1$ ordine del gruppo moltiplicativo $\mathbb{F}_q^*$. Segue che $\forall 1 \leq i \leq m$ indice, $(x^{h/p_i} - 1)$ ha al più h/p_i radici, quindi esistono elementi di $\mathbb{F}_q$ che non sono radici di questo polinomio. Sia a_i un elemento di questo tipo, e $b_i = a_i^{h/p_i^{r_i}}$. Nota che $b_i^{p_i^{r_i}} = 1$ poiché otteniamo l'ordine all'esponente. D'altra parte, $b^{p_i^{r_i-1}} = a^{h/p_i} \neq 1$ per costruzione.

Affermo che $b = \prod_i b_i$ ha ordine h. Se per assurdo non fosse così, si avrebbe che $\mathrm{ord} b \mid h, b \neq h \implies \mathrm{ord} b \mid h/p_i$ (è solo una furbata per scrivere che è un divisore proprio). Facciamo che $\mathrm{ord} b \mid h/p_1$. Allora

$$1 = b^{\mathrm{ord} b} = b^{h/p_1} = b_1^{h/p_1} b_2^{h/p_1} \cdots b_m^{h/p_1}.$$

Il primo termine $b_1^{h/p_1} \neq 1$ per costruzione, mentre tutti gli altri $b_i^{h/p_1} = 1$ poiché abbiamo che $p_i^{r_i} \mid h/p_1 \ \forall i \geq 2$ (e ovviamente $i \leq m$), che sarebbe assurdo per definizione di ordine. Segue che $\mathbb{F}_q^*$ è un gruppo ciclico con generatore b. □

Definizione 2.7. Un generatore del gruppo ciclico F^* è detto *elemento primitivo.*

Teorema 2.8. *Sia $\mathbb{F}_q$ un campo e $\mathbb{F}_r$ una sua estensione finita. Allora $\mathbb{F}_r = \mathbb{F}_q(\alpha)$, dove α è un elemento primitivo di $\mathbb{F}_r$.*

Dimostrazione. Sia α un elemento primitivo di $\mathbb{F}_r$. Allora $\mathbb{F}_q(\alpha) \subset \mathbb{F}_r$ poiché è la più piccola estensione contenente un elemento di $\mathbb{F}_r$; e $\mathbb{F}_q(\alpha) \supset \mathbb{F}_r$ poiché $\langle \alpha \rangle = \mathbb{F}_r \subset \mathbb{F}_q$. Per assioma di estensione quindi $\mathbb{F}_r = \mathbb{F}_q(\alpha)$. □

Corollario 2.9. *Per ogni campo finito $\mathbb{F}_p$ ed ogni intero positivo n esiste un polinomio irriducibile su $\mathbb{F}_p[x]$ di grado n.*

Dimostrazione. Consideriamo l'estensione del campo $\mathbb{F}_r$ interpretandolo come spazio vettoriale, di dimensione $n = [\mathbb{F}_r : \mathbb{F}_p]$. Allora abbiamo $\mathbb{F}_r = \mathbb{F}_q(\alpha) \simeq \mathbb{F}_q[x]/(g)$ per qualche $\alpha \in \mathbb{F}_r$ elemento primitivo, e g il polinomio minimo per α algebrico su $\mathbb{F}_q$. □

2.1 Radici di Polinomi Irriducibili

Lemma 2.10. *Sia $f \in \mathbb{F}_q[x]$ irriducibile con radice α in un'estensione del campo. Allora $h \in \mathbb{F}_q[x]$ ha radice α sse $f|h$.*

Dimostrazione. f irriducibile $\implies$ f polinomio minimo di α $\implies$ $h(\alpha) = 0$ se e solo se $f|h$. □

Proposizione 2.11. *Sia $f \in \mathbb{F}_q[x]$ un polinomio irriducibile di grado m. Allora $f | x^{q^n} - x \iff m \mid n$.*

Dimostrazione.

" $\implies$ " $f \mid x^{q^n} - x \implies \alpha$ radice di $f \in \mathbb{F}_q[x]$ è anche radice di $x^{q^n} - x \implies \mathbb{F}_{q^n} \supset \mathbb{F}_q(\alpha) = \mathbb{F}_{q^m} \implies m|n$.

" $\impliedby$ " $m|n \implies [\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m,\ \mathbb{F}_{q^n} \ni \alpha \implies \alpha^{q^n} - \alpha = 0 \implies f \mid x^{q^n} - x$. □

Teorema 2.12. *Il polinomio irriducibile $f \in \mathbb{F}_q$ ha una radice $\alpha \in \mathbb{F}_{q^m}$. Tutte le altre radici sono $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$, e son tutte distinte.*

Dimostrazione. Sappiamo già che $\alpha \in \mathbb{F}_{q^m} = \mathbb{F}_q(\alpha) = \mathbb{F}_q[x]/(f)$ (aggiungiamo un elemento al campo, usiamo il primo teorema di isomorfismo, e vediamo l'estensione come spazio vettoriale di dimensione m).

Ora, se β soluzione, anche β^q è una soluzione:

$$\begin{aligned} & f(\alpha) = a_m \beta^m + \cdots + a_0 = 0 \\ \implies & f(\beta)^q = a_m^q \beta^{qm} + \cdots + a_0^q = a_m (\beta^q)^m + \cdots + a_0 = f(\beta^q) = 0. \end{aligned}$$

Rimane mostrare che tutte le radici sono diverse:

$$\begin{aligned}
&\alpha^{q^i} = \alpha^{q^j} \\
\implies &\alpha^{q^{i+m-j}} = \alpha^{q^{j+m-j}} = \alpha \\
\implies &\alpha \text{ radice di } x^{q^{m+i-j}} - x \\
\implies &f \mid x^{q^{m+i-j}} - x \\
\implies &m \mid m+i-j, \quad 0 < m+i-j < m \\
\implies &i = j \qquad \text{[poiché altrimenti dovrebbe esser un multiplo proprio di } m\text{]} \quad \square
\end{aligned}$$

Definizione 2.13 (Coniugato)**.** Sia $\alpha \in \mathbb{F}_{q^m}$ estensione di $\mathbb{F}_q$; $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^m}$ sono detti *coniugati* di α.

Notiamo che per il Teorema 2.12 sui coniugati tali elementi sono distinti se il polinomio minimo di $\alpha \in \mathbb{F}_q$ è di grado m, altrimenti d elementi distinti, con $d \mid m$, ripetuti m/d volte ciascuno.

Inoltre, tutti i coniugati appartengono allo stesso gruppo ciclico $\langle\alpha\rangle = \{\alpha^i\}_{i\in\mathbf{Z}}$ e hanno quindi tutti lo stesso ordine. Nel caso particolare in cui α è un elemento primitivo per il gruppo, tutte le sue potenze α^{q^i} sono elementi primitivi.

Teorema 2.14. *Tutti gli automorfismi nel campo $\mathbb{F}_{q^m}$ su $\mathbb{F}_q$ sono della forma $\sigma_i : \mathbb{F}_{q^m} \to \mathbb{F}_q : \alpha \mapsto \alpha^{q^i}$ $\forall 0 \le i < m$.*

Dimostrazione.

- **"$\Longleftarrow$"** $\sigma_i(\alpha\beta) = \sigma_i(\alpha)\sigma_i(\beta)$ per le proprietà delle potenze, e $\sigma_i(\alpha+\beta) = \sigma_i(\alpha)+\sigma_i(\beta)$ per via del fatto che $\mathbb{F}_{q^m}$ ha caratteristica p (Freshman's Dream). Inoltre $\sigma_i(\alpha) = 0 \iff \alpha = 0$, pertanto σ_i è iniettiva e surgettiva (poiché $\sigma_i(a) = a$ $\forall a \in \mathbb{F}_q$). Quindi σ_i è un automorfismo.
- **"$\Longrightarrow$"** Sia σ un automorfismo arbitrario e β un elemento primitivo di $\mathbb{F}_{q^m}$ con polinomio minimo $f = a^m x^m + a_{m-1}\beta^{m-1} + \cdots + a_0$. Allora:

$$\begin{aligned}
0 = f(\beta) &= \sigma \circ f(\beta) \\
&= \sigma(a_m\beta^m + a_{m-1}\beta^{m-1} + \cdots + a_0) \\
&= \sigma(a_m\beta^m) + \sigma(a_{m-1}\beta^{m-1}) + \cdots + \sigma(a_0) && [\sigma \text{ è lineare}] \\
&= \sigma(\beta^m)a_m + \sigma(\beta^{m-1})a_{m-1} + \cdots + a_0 && [\sigma \text{ conserva il prodotto scalare}] \\
&= \sigma(\beta)^m a_m + \sigma(\beta)^{m-1}a_{m-1} + \cdots + a_0 && [\sigma \text{ è un morfismo}]
\end{aligned}$$

  $\implies \sigma(\beta)$ è una radice del polinomio minimo $\implies \sigma(\beta)$ è un elemento primitivo $\implies \sigma(\beta) = \beta^{q^i}$ e tutti gli elementi sono potenze di questo.

$\square$

2.2 Tracce, Norme, e Basi

Consideriamo $K = \mathbb{F}_q \subset F = \mathbb{F}_{q^m}$ campi. Consideriamo il polinomio minimo di α in K, f di grado $d \mid m$. Il polinomio $g = f^{m/d}$ è detto *polinomio caratteristico* di f. Per l'identità di Newton tale polinomio $f = \sum_i^m a_i x^i$ ha come coefficienti σ_i, in particolare

$$\begin{aligned}
(-1)^m a_0 &= \prod_i \alpha_i = \alpha \cdots \alpha^{q^{m-1}} \quad \text{è detto norma } \mathrm{N}_{F/K}(\alpha) \\
-a_{m-1} &= \sum_i \alpha_i = \alpha + \cdots + \alpha^{q^{m-1}} \quad \text{è detto traccia } \mathrm{Tr}_{E/K}(\alpha)
\end{aligned}$$

Tracce

Definizione 2.15. La *traccia* è la somma dei coniugati.

Proposizione 2.16 (Proprietà della Traccia)**.** *La funzione* $\mathrm{Tr}_{F/K}$ *soddisfa:*

(i) $\mathrm{Tr}(\alpha+\beta) = \mathrm{Tr}(\alpha) + \mathrm{Tr}(\beta) \quad \forall \alpha, \beta \in F$;

(ii) $\mathrm{Tr}(c\alpha) = c\,\mathrm{Tr}(\alpha) \quad \forall c \in K$;

(iii) Tr *è una mappa surgettiva lineare;*

(iv) $\mathrm{Tr}(c) = mc \quad \forall c \in K \ (m = [F : K])$;

(v) $\mathrm{Tr}(\alpha^q) = \mathrm{Tr}(\alpha) \quad \forall \alpha \in F$

Dimostrazione.

(i) $\mathrm{Tr}(\alpha+\beta) = \sum_i (\alpha+\beta)^{q^i} = \sum_i (\alpha^{q^i} + \beta^{q^i}) = \sum_i \alpha^{q^i} + \sum_i \beta^{q^i} = \mathrm{Tr}(\alpha) + \mathrm{Tr}(\beta)$;

(ii) $\mathrm{Tr}(c\alpha) = \sum_i (c\alpha)^{q^i} = c\sum_i \alpha^{q^i} = c\,\mathrm{Tr}(\alpha)$;

(iii) Si ha da (i) e (ii) che Tr è una mappa lineare. Ora, $\dim(\mathrm{Im}(\mathrm{Tr}))$ è 0 o 1. Se per assurdo fosse 0 il polinomio $\mathrm{Tr} = x^{q^{m-1}} + \cdots + x$ ha q^m radici (tutti gli elementi di F), ma questo è assurdo poiché ha grado q^{m-1}. Ne dobbiamo conseguire che la mappa ha dimensione dell'immagine 1 quindi è surgettiva;

(iv) $\mathrm{Tr}(c) = \sum_i c^{q^i} = \sum_i^{m-1} c = mc$;

(v) α è coniugato di α^q, quindi hanno stessa somma dei coniugati.

□

Teorema 2.17. *Sia $f : \mathbb{F}_{q^m} \to \mathbb{F}_q$ una mappa $\mathbb{F}_q$-lineare. Allora $\exists! \beta \in \mathbb{F}_{q^m}$. $f(\alpha) = \mathrm{Tr}(\beta\alpha) = L_\beta(\alpha)$.*

Dimostrazione.

" $\Longleftarrow$ " L_β è lineare, infatti: $L_\beta(\alpha+\varepsilon) = \mathrm{Tr}(\beta\alpha+\beta\varepsilon) = L_\beta(\alpha) + L_\beta(\varepsilon)$, e $L_\beta(c\alpha) = \mathrm{Tr}(c\beta\alpha) = cL_\beta(\alpha)$;

" $\Longrightarrow$ " per dimostrare che tutte le funzioni lineare un L_β, mostro che sono uguali in numero. Noto anzitutto che tutte le funzioni $\mathbb{F}_q$-lineari su $\mathbb{F}_{q^m}$ sono tutte le matrici di $1 \times m$ con elementi in $\mathbb{F}_q$, quindi q^m. Poi,

$|\{f : \mathbb{F}_{q^m} \to \mathbb{F}_q, f \text{ lineare}\}| = q^m = |\{L_\beta : \beta \in \mathbb{F}_{q^m}\}| \iff (L_\beta = L_\gamma \implies \beta = \gamma)$

È sufficiente a questo punto notare che $0 = L_\beta(\alpha) - L_\gamma(\alpha) = \mathrm{Tr}(\beta\alpha) - \mathrm{Tr}(\gamma\alpha) = \mathrm{Tr}((\beta-\gamma)\alpha) \iff \beta - \gamma = 0$ poiché, come abbiamo già notato, la dimensione dell'immagine della traccia non è 0. □

Teorema 2.18 (Transitività della traccia)**.** *Siano $K \subset F \subset E$ campi. Allora:*

$$\mathrm{Tr}_{F/K}(\mathrm{Tr}_{E/F}(\alpha)) = \mathrm{Tr}_{E/K}(\alpha).$$

Dimostrazione. Assegniamo variabili ai gradi: $[F : K] = m, [E : F] = n \implies [E : K] = mn$.

$$\mathrm{Tr}_{F/K}(\mathrm{Tr}_{E/F}(\alpha)) = \sum_i^{m-1} \Big(\sum_j^{n-1} \alpha^{q^{m^j}} \Big)^{q^i} = \sum_{i=0}^{m-1} \Big(\sum_{j=0}^{n-1} \alpha^{q^{(mj+i)}} \Big) = \sum_i^{mn-1} \alpha^{q^i} = \mathrm{Tr}_{F/K}(\alpha)$$

□

Norme

Definizione 2.19 (Norma)**.** La norma è il prodotto dei coniugati:

$$\mathrm{N}_{F/K}(\alpha) = \prod_i \alpha^{q^i} = \alpha^{1+\cdots+q^{m-1}} = \alpha^{\frac{1-q^m}{1-q}}.$$

È facile notare che la norma non è una mappa lineare (la somma non funziona); tuttavia $\mathrm{N}(\alpha) = 0 \iff \alpha = 0$. Inoltre, ricordiamo che il codominio della norma vive in K, e questo può esser osservato dai coefficienti del polinomio minimo per α.

Proposizione 2.20 (Proprietà della Norma)**.** *La norma soddisfa le seguenti proprietà:*

(i) $\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta) \quad \forall \alpha, \beta \in F$;

(ii) la norma ristretta al gruppo moltiplicativo è un morfismo surgettivo;

(iii) $\mathrm{N}(a) = a^m \quad \forall a \in K$;

(iv) $\mathrm{N}(\alpha^q) = \mathrm{N}(\alpha)$.

Dimostrazione.

(i) banale dalla definizione di potenza;

(ii) per (i) sappiamo già che è un morfismo, rimane mostrare che è surgettivo. Sappiamo che il kernel di N sono tutte le radici di $x^{\frac{q^m-1}{q-1}} - 1$ quindi al più $\frac{q^m-1}{q-1}$. Ma allora, per il primo teorema di isomorfismo:

$$|\operatorname{Im}(\mathrm{N})| = \left|\frac{\mathbb{F}_{q^m}^*}{\ker \mathrm{N}}\right| = \frac{|\mathbb{F}_{q^m}^*|}{|\ker \mathrm{N}|} \geq q^m - 1 \cdot \frac{q-1}{q^m-1} = q - 1 = |\mathbb{F}_q^*|;$$

(iii) banale;

(iv) $\mathrm{N}(\alpha^q) = \mathrm{N}(\alpha)^q = \mathrm{N}(\alpha)$.

□

Teorema 2.21 (Transitività della Norma)**.** *Siano $K \subset F \subset E$ campi. Allora:*

$$\mathrm{N}_{F/K}(\mathrm{N}_{E/F}(\alpha)) = \mathrm{N}_{E/K}(\alpha).$$

Basi

Abbiamo detto che se $\alpha \in \mathbb{F}_{q^m} = F$, possiamo interpretare il campo come uno spazio vettoriale, e pertanto esiste un'unica combinazione lineare che lo descrive data la base $\alpha_1, \ldots, \alpha_m$:

$$\alpha = c_1\alpha_1 + \cdots + c_m\alpha_m = c_1(\alpha)\alpha_1 + \cdots + c_m(\alpha)\alpha_m$$

in cui ogni coefficiente è una funzione $\mathbb{F}_q$-lineare, quindi per il Teorema 2.17

$$c_i(\alpha) = L_{\beta_i}(\alpha) = \operatorname{Tr}(\beta_i\alpha) \quad \forall \alpha \in F.$$

Vale la pena notare che l'insieme $\beta_i, \ldots, \beta_m$, posto $\alpha = \alpha_i$, è tale che $\operatorname{Tr}(\alpha_i\beta_j)$ è 0 se $i \neq j$, altrimenti 1 (si dimostra per sostituzione diretta). In più, $\{\beta_i\}_i$ è una base di $\mathbb{F}_{q^m}$. Data una combinazione lineare su dei d_i:

$$d_1\beta_1 + \cdots + d_m\beta_m = 0 \implies \alpha_i(d_1\beta_1 + \cdots + d_m\beta_m) = 0 \quad \forall \alpha_i \implies d_i = 0 \quad \forall i.$$

Definizione 2.22 (Base Duale)**.** Sia $\mathbb{F}_{q^m}$ un campo estensione finita per $\mathbb{F}_q$, e $\alpha_1, \ldots, \alpha_m$ una base. Allora $\beta_1, \ldots, \beta_m$ è una detta *base duale* o *complementare* sse:

$$\operatorname{Tr}(\alpha_i\beta_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{altrimenti} \end{cases}$$

Osservazione 2.23. I β_i della base duale sono univocamente determinati, poiché per il Teorema 2.17 le mappe lineari che descrivono i coefficienti sono unicamente determinate da un β_i.

Definizione 2.24 (Base Auto-Duale)**.** Una base è detta *auto-duale* se $\{\alpha_i\}_i = \{\beta_i\}_i$.

Definizione 2.25 (Base Polinomiale)**.** Diremo che una base è *polinomiale* sse composta dalle potenze consecutive del campo $\mathbb{F}_{q^m}$ su $\mathbb{F}_q$: $\langle 1, \alpha, \alpha^2, \ldots, \alpha^{m-1}\rangle$.

Definizione 2.26 (Base Normale)**.** Diremo che una base è *normale* sse composta della forma $\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}$ per qualche $\alpha \in \mathbb{F}_{q^m}$ che abbia tali potenze distinte.

Lemma 2.27 (Lemma di Artin)**.** *Siano $\{\psi_i : G \to \mathbb{F}^*\}_{i=1}^m$ morfismi da un gruppo qualsiasi sul gruppo moltiplicativo, e, siano $\{a_i \in \mathbb{F}\}_{i=1}^m$ elementi del campo non tutti nulli. Allora*

$$\exists g \text{ s.t. } a_1\psi_1(g) + a_2\psi_2(g) + \cdots + a_m\psi_m(g) \neq 0.$$

Dimostrazione. Per induzione:

($m = 1$) banale, $a_1 \neq 0 \ \wedge \ \psi_1(g) \in \mathbb{F}^*$;

($m - 1 \implies m$) se $a_1 = 0$ non c'è nulla da dimostrare, possiamo sfruttare l'ipotesi induttiva. Altrimenti, se per assurdo:

$$a_1\psi_1(g) + \cdots + a_{m-1}\psi_{m-1}(g) + a_m\psi_m(g) = 0 \quad \forall g \in G$$
$$\implies a_1\psi_1(h)\psi_1(g) + \cdots + a_{m-1}\psi_{m-1}(h)\psi_{m-1}(g) + a_m\psi_m(h)\psi_m(g) = 0 \quad \forall g \in G \quad [\exists h \text{ s.t. } \psi_m(h) \neq \psi_1(h)]$$

moltiplicando per $\psi_m(h)^{-1}$:

$$b_1\psi_1(g) + \cdots + b_{m-1}\psi_{m-1}(g) + a_m\psi_m(g) = 0 \quad \forall g \in G \qquad [b_i \coloneqq a_i\psi_i(h)\psi_m(h)^{-1}]$$

sottraendo l'equazione iniziale:

$$c_1\psi_1(g) + \cdots + c_{m-1}(g) = 0 \qquad [c_i \coloneqq a_i - b_i \ \forall 0 \leq i < m]$$

che va in contraddizione con l'ipotesi induttiva. □

Analogie con Algebra Lineare. Dato T operatore lineare, allora $f(x)$ è detta annullare T sse $f(T) = a_0T^0 + \cdots + a_nT^n = 0$. Se f è il polinomio di grado minimo per cui T risulta radice, ed è monico, allora è detto *polinomio minimo*. Tale polinomio divide il *polinomio caratteristico*, che per il Teorema di Cayley-Hamilton è dato da $g(x) = det(xI_n - T)$.

Inoltre, un vettore $\alpha \in V$ è detto *ciclico* sse $\{ \alpha T^k \mid k \in \mathbf{N} \}$ genera lo spazio vettoriale $\mathbb{F}_{q^m}$.

Lemma 2.28. *Sia T un operatore lineare su uno spazio vettoriale finito. Allora T ha un vettore ciclico sse il polinomio caratteristico e quello minimo coincidono.*

Dimostrazione. `TODO: non dimostrato nel libro.` □

Teorema 2.29 (Teorema delle Basi Normali)**.** *Per ogni campo finito K ed F una sua estensione finita, esiste una base normale F su K.*

Dimostrazione. Si dimostra usando il morfismo di Frobenius.

Consideriamo la mappa di Frobenius $\sigma^i : \mathbb{F}^* \to \mathbb{F}^* : \alpha \mapsto \alpha^{q^i}$. Essa è banalmente una trasformazione lineare, quindi un automorfismo. Notiamo che $\sigma^m = \mathbf{1}$, quindi σ è radice di $x^m - 1$, e per il lemma di Artin gli automorfismi $\mathbf{1}, \ldots, \sigma^{m-1}$ sono linearmente indipendenti, quindi $x^m - 1$ è il polinomio minimo per σ. Inoltre tale polinomio è caratteristico - ricordiamo che il polinomio caratteristico è un polinomio di grado m monico, multiplo di m.

Si usa poi il lemma di sopra per mostrare che esiste un elemento α tale per cui $\alpha, \sigma(\alpha), \ldots, \sigma^{m-1}(\alpha)$ è una base di $\mathbb{F}$. □

Discriminante

Definizione 2.30 (Discriminante)**.** Il *discriminante* $\Delta_{F/K}(\alpha_1, \ldots, \alpha_m)$ degli elementi $\alpha_1, \ldots, \alpha_m$ è il determinante della matrice $m \times m$: $\Delta_{F/K} = \det\left(\left[\operatorname{Tr}(\alpha_i\alpha_j)\right]_{i,j}\right)$.

Osservazione 2.31. Il discriminante Δ sta sempre nel campo K poiché tutti gli elementi della matrice stanno nel campo.

Teorema 2.32. *Sia F un'estensione del campo K di grado m. Allora $\alpha_1, \ldots, \alpha_m \in F$ sono una base per F sse $\Delta_{F/K}(\alpha_1, \ldots \alpha_m) \neq 0$.*

Dimostrazione.

"$\implies$" Abbiamo che gli a_i sono linearmente indipendenti, quindi

$$\begin{aligned}
&\sum_i c_i\alpha_i = 0 \implies c_i = 0 \\
\implies &\sum_i c_i\alpha_i\alpha_j = 0 \implies c_i = 0 \quad (\forall\ 1 \leq j \leq m) && \text{[F dominio]} \\
\implies &\sum_i c_i \operatorname{Tr}(\alpha_i\alpha_j) = 0 \implies c_i = 0 \quad (\forall\ 1 \leq j \leq m) && \text{[Tr morfismo]}
\end{aligned}$$

che vuol dire che ogni riga della matrice $\left[\operatorname{Tr}(\alpha_i\alpha_j)\right]_{i,j}$ è linearmente indipendente, quindi la matrice ha rango massimo, quindi il determinante è non nullo.

"$\impliedby$" $\Delta(\alpha_1, \ldots, \alpha_m) \neq 0$, vogliamo mostrare che gli $\{\alpha_i\}$ sono linearmente indipendenti.

$$\begin{aligned}
&\exists c_1, \ldots, c_m \in K \text{ s.t. } c_1\alpha_1 + \cdots + c_m\alpha_m = 0 \\
\implies &c_1\alpha_1\alpha_j + \cdots + c_m\alpha_m\alpha_j = 0 \quad \forall 1 \leq j \leq m && \text{[moltiplicando per } \alpha_j\text{]} \\
\implies &c_1 \operatorname{Tr}(\alpha_1\alpha_j) + \cdots + c_m \operatorname{Tr}(\alpha_m\alpha_j) = 0 \quad \forall 1 \leq j \leq m && \text{[applicando la traccia]} \\
\implies &c_1 = c_2 = \cdots = c_m = 0 && \text{[definizione di determinante]}
\end{aligned}$$

poiché gli elementi di ogni riga sono non nulli, visto che il determinante Δ è non-nullo. □

Possiamo in realtà fare una cosa più furba: notiamo che $\operatorname{Tr}(\alpha_i\alpha_j) = \sum_{k=0}^{m-1}(\alpha_i\alpha_j)^{q^k} = \alpha_i\alpha_j + \alpha_i^q\alpha_j^q + \cdots + \alpha_i^{q^{m-1}}\alpha_j^{q^{m-1}}$, che è proprio la (i,j)-esima cella del prodotto A^TA dove $A = \left[\alpha_j^{q^{i-1}}\right]_{i,j} \implies \det(AA^T) = \det(A^2) = \det(A)^2 \neq 0 \iff \alpha_1, \ldots, \alpha_m$ è una base per F.

Notiamo inoltre che $\left[\alpha_j^{q^{i-1}}\right]_{i,j}$ è una matrice di Vandermonde, quindi il suo determinante è $\prod_{i<j}(\alpha_i - \alpha_j)$.

Corollario 2.33. *Siano $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_{q^m}$. Allora $\langle \alpha_1, \alpha_2, \ldots, \alpha_m \rangle$ è una base per $\mathbb{F}_{q^m}$ su $\mathbb{F}_q$ se e soltanto se*

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \alpha_1^q & \alpha_2^q & \cdots & \alpha_m^q \\ \vdots & \vdots & \vdots & \\ \alpha_q^{q^{m-1}} & \alpha_2^{q^{m-1}} & \cdots & \alpha_m^{q^{m-1}} \end{vmatrix} \neq 0 \tag{2.1}$$

Teorema 2.34. *$\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}$ è una base per F su K sse i polinomi $x^m - 1$ e $\alpha x^{m-1} + \cdots + \alpha^{q^{m-1}}$ sono coprimi.*

Dimostrazione. Si usa il risultante dei due polinomi. Applicando una trasformazione lineare alla matrice di cui il risultante è il determinante, otteniamo una matrice nella forma 2.1, che è nonzero sse i due polinomi sono coprimi sse $\alpha_1, \alpha_2, \cdots, \alpha_m$ è una base per F/K. `TODO: più formale` □

2.3 Polinomi Ciclotomici

Definizione 2.35. Il campo di spezzamento del polinomio $x^n - 1$ è detto *n-esimo campo ciclotomico* su K ed è denotato con $K^{(n)}$. Le radici di tale polinomio sono dette *n-esime radici dell'unità* e il loro insieme viene denotato con $E^{(n)}$.

Teorema 2.36. *Sia n un intero positivo e K un campo di caratteristica p. Allora:*

- *$p \nmid n \implies E^{(n)}$ è un gruppo ciclico di ordine n rispetto alla moltiplicazione in $K^{(n)}$;*
- *$p \mid n \implies n = mp^e$ con $(m, p) = 1$ e $K^{(n)} = K^{(m)}$, $E^{(n)} = E^{(m)}$ e le radici del polinomio ciclotomico sono tutte ivi contenute, ciascuna con molteplicità p^e.*

Dimostrazione. (Escludiamo il caso banale in cui $n = 1$ in cui non c'è nulla da dimostrare).

Pongo $f = x^n - 1$; dato che $f' = nx^{n-1}$, se $p \nmid n$ allora non vi sono radici multiple, quindi $E^{(n)}$ ha n elementi. Mostro che $E^{(n)}$ è un sottogruppo: $1 \in E^{(n)}$, $\forall \eta, \zeta \in E^{(n)} (\eta \zeta^{-1})^n = \eta^n (\zeta^n)^{-1} = 1$. Per mostrare che tale gruppo è ciclico, uno adotta lo stesso argomento del Teorema 2.6: si scompone n in fattori primi $n = \prod_i^k p_i^{e_i} \implies \exists \alpha_i$ *non radice* di $x^{n/p_i} - 1$, $\implies \beta_i = \alpha_i^{n/p_i^{e_i}} \implies \beta = \prod_i^k \beta_i$ genera $E^{(n)}$.

Il secondo punto segue immediatamente da $x^n - 1 = x^{mp^e} - 1 = (x^m - 1)^{p^e}$ perché K ha caratteristica p. □

Definizione 2.37 (Radice Primitiva dell'Unità)**.** Una radice dell'unità che è anche generatore del gruppo ciclico $E^{(n)}$ è detta *radice primitiva dell'unità*.

Osservazione 2.38. Dal Teorema 1.11 sappiamo che le radici *primitive* n-esime dell'unità sono esattamente $\varphi(n)$. Infatti, se ζ è una radice primitiva dell'unità, allora anche ζ^s lo è, dove $1 \leq s \leq n$, ed $(s, n) = 1$.

Definizione 2.39 (n-esimo Polinomio Ciclotomico)**.** Sia $K = \mathbb{F}_{q^m}$, n un intero positivo non divisibile per p, e ζ una radice primitiva dell'unità. Il polinomio

$$Q_n(x) = \prod_{\substack{s=1 \\ (s,n)=1}}^{n} (x - \zeta^s) = \prod_{\substack{\omega \in E^{(n)} \\ \mathrm{ord}(\omega) = n}} (x - \omega)$$

di grado $\varphi(n)$ è detto n-esimo polinomio ciclotomico di K e tutte le sue radici sono i generatori di $E^{(n)}$.

Teorema 2.40. *Se la caratteristica del campo è zero, allora $Q_n(x)$ è irriducibile su K, e $[K^{(n)} : K] = \varphi(n)$.*

Dimostrazione. Se la caratteristica del campo K è zero, allora non divide mai n, e pertanto $K^{(n)} = K(\zeta)$ dove ζ è una n-esima radice primitiva dell'unità. Inoltre $Q_n(x)$ è il suo polinomio minimo poiché ha uniche radici essa e i suoi coniugati aventi stesso ordine, pertanto $[K^{(n)}] : K] = \deg Q_n(x) = \varphi(n)$ □

Teorema 2.41. *Il numero di polinomi monici irriducibili in $\mathbb{F}_q[x]$ di grado m ed ordine e è esattamente $\varphi(n)/m$, dove m è il minimo intero per cui $q^m \equiv 1 \pmod{n}$. Se $e = m = 1$ allora in due; altrimenti 0.*

Dimostrazione. Se $e = m = 1$ posiamo facilente contarli, e sono x, $x - 1$. Se la caratteristica è un primo, allora una generica radice dell'unità η appartiene al campo $\mathbb{F}_{q^k}$ sse $\eta^{q^k} = \eta$, ossia se $n \mid q^k - 1 \implies q^k \equiv 1 \pmod{n}$ (per definizione di radice primitiva, che ha ordine n). Sia m il minimo numero che soddisfa tale proprietà; notiamo allora che $[K(\eta) : K] = m$ e tale numero non dipenede dalla radice considerata. Ne consegue che $Q_n(x)$ fattorizza in $\varphi(n)/m$ polinomi irriducibili, aventi tutti il medesimo grado m. □

Teorema 2.42 (Proprietà di Q_n)**.** *I polinomi ciclotomici godono delle seguenti proprietà:*

(a) $Q_{mp}(x) = Q_m(x^p)/Q_m(x)$ *se* p *è primo e* $m \in \boldsymbol{N}$ *s.t.* $p \nmid m$*;*

(b) $Q_{mp}(x) = Q_m(x^p)$ *se* p *è primo e* $m \in \boldsymbol{N}$ *s.t.* $p \mid m$*.*

Dimostrazione. `TODO: dimostrati negli appunti, è abbastanza facile e si va per induzione.` □

Teorema 2.43. *Sia* K *campo di caratteristica* p *ed* $n > 0$ *s.t.* $n \nmid p$*:*

(i) $x^n - 1 = \prod_{d|n} Q_d(x)$*;*

(ii) i coefficienti di $Q_n(x)$ *appartengono al sottocampo primo di* K*, e a* $\mathbf{Z}$ *se* $K = \mathbf{Q}$*.*

Dimostrazione.

(i)
$$x^n - 1 = \prod_{s=1}^{n} (x - \zeta^s) = \prod_{d|n}^{n} Q_d$$
viene semplicemente raggruppando i termini della radice dell'unità aventi stesso periodo, ed esso è unico perché una potenza generica ζ^s ha ordine $d/(s,d)$.

(ii) Per induzione: se $n = 1$ allora $Q_1 = x - 1 \in K[x]$ banalmente; altrimenti suppongo $Q_d(x) \in K[x] \ \ \forall 1 \leq d < n \implies$
$$Q_n = \frac{x^n - 1}{\prod_{\substack{d|n \\ d<n}} Q_d(x)} \in K[x]$$
per ipotesi induttiva (e perché banalmente il numeratore ha coefficienti nel campo). □

Teorema 2.44. *Il campo finito* $\mathbb{F}_q$ *è il campo ciclotomico* $(q-1)$*-esimo su qualsiasi suo sottocampo.*

Dimostrazione. Per Eulero-Fermat $\mathbb{F}_q^*$ sono le soluzioni di $x^{q-1} - 1 = 0$. Non è possibile che tutti gli elementi di $\mathbb{F}_q^*$ stiano in un sottocampo proprio di $\mathbb{F}_q$. □

2.4 Teorema di Weddeburn

Definizione 2.45. Un *corpo*, o *finite division ring*, è un anello che ha tutte le proprietà dei campi eccetto la commutatività sulla moltiplicazione.

Teorema 2.46 (Teorema di Weddeburn)**.** *Ogni corpo finito è un campo.*

Capitolo 3

Polinomi su Campi Finiti

Teorema 3.1. *Sia* $f \in \mathbb{F}_q[x]$ *s.t.* $f(0) \neq 0, \deg f = m \implies \exists 1 \leq e < q^m$ *s.t.* $f \mid x^e - 1$.

Dimostrazione. L'anello quoziente $\mathbb{F}_q[x]/(f)$ ha $q^m - 1$ classi di equivalenza nonzero; l'insieme $\{x^j + (f) : 0 \leq j \leq q^m - 1\}$ ha q^m elementi, e sono tutti non nulli. Allora:

$$\begin{aligned} & \exists i, j \quad x^i + (f) = x^j + (f) \\ \implies & \exists i, j \quad x^i = x^j \pmod f \\ \implies & \exists i, j \quad x^{i-j} = 1 \pmod f && \text{[poiché } (x, f) = 1\text{]} \\ \implies & \exists 0 \leq e \leq q^m - 1 \quad x^e = 0 \pmod f. && \square \end{aligned}$$

Definizione 3.2 (Ordine di un polinomio)**.** Sia $f \in \mathbb{F}[x]$ non identicamente nullo. Allora *l'ordine di* f, detto anche *periodo* o *esponente*, e denotato $\text{ord}(f)$, è l'intero più piccolo per cui $f \mid x^e - 1$ se $f(0) \neq 0$; altrimenti $f = x^h g$ con $g(0) \neq 0$ e $\text{ord}(f) = \text{ord}(g)$.

Teorema 3.3. *L'ordine di un polionomio irriducibile è uguale all'ordine moltiplicativo di una radice presa nell'estensione.*

Dimostrazione. Sia f polinomio di grado m su $\mathbb{F}_q$ tale che $\deg f = m$, irriducibile. Allora lo splitting field di f è $\mathbb{F}_{q^m}$ ed $\exists \alpha \in \mathbb{F}_{q^m}$ radice di f. Allora $\text{ord}(\alpha) = e \iff \alpha^e = 1 \iff \alpha$ radice di $x^e - 1 \iff f \mid x^e - 1 \iff \text{ord}(f) = e$ (con e minimo). $\square$

Segue immediatamente da questo che $\text{ord}(f) \mid q^m - 1$: se $f = cx \implies \text{ord}(f) = 1$, altrimenti $\text{ord}(f) = \text{ord}(\alpha) \mid \text{ord}(\mathbb{F}^*_{q^m})$.

Teorema 3.4. *Il numero di polinomi monici, irriducibili, di grado* m *ed ordine* e *è uguale a* $\varphi(e)/m$ *per* $e \geq 2$. *Se* $e = m = 1$ *sono* 2*, altrimenti* 0.

Dimostrazione. Immediato per il Teorema 2.41: se abbiamo un polinomio irriducibile f di ordine e, allora $f \mid Q_e$ e ogni fattore di Q_e ha lo stesso grado m (dove ricordiamo m è il minimo intero per cui $q^m \equiv 1 \pmod n$). Nel caso in cui $e = m = 1$ si contano: $x - 1$ e x monici irriducibili. $\square$

Lemma 3.5. *Sia* $c > 0$. $f \in \mathbb{F}_q[x], f(0) \neq 0$. *Allora* $f \mid x^c - 1 \iff$ *ord*$(f) \mid c$.

Dimostrazione. Fisso $e = \text{ord}(f)$.

- **"** $\impliedby$ **"** $\text{ord}(f) \mid c \implies f \mid x^e - 1,\ e \mid c \implies f \mid x^e - 1,\ x^e - 1 \mid x^c - 1 \implies f \mid x^c - 1$.
- **"** $\implies$ **"** si fa usando la divisione con resto: $f \mid x^c - 1$ implica che $c \geq e$, dunque $c = qe + r$ con $0 \leq r < e$. Riscrivo $x^c - 1 = (x^{qe+r} - 1) = (x^{qe} - 1)x^r + (x^r - 1)$. Segue che $f \mid x^r - 1$, che è possibile solo per $r = 0$. $\square$

Corollario 3.6. $\gcd(x^{e_1} - 1, x^{e_2} - 1) = x^d - 1$ *dove* $d = (e_1, e_2)$.

Dimostrazione. Chiamo $f_1 = x^{e_1} - 1$ e $f_2 = x^{e_2} - 1$, ed $f = (f_1, f_2)$ il loro massimo comun divisore. Per definizione di massimo comun divisore si ha che $x^d - 1 \mid f_1,\ x^d - 1 \mid f_2 \implies x^d - 1 \mid f$. D'altra parte, per il Lemma 3.5 di sopra, $f \mid f_1,\ f \mid f_2 \implies \text{ord}(f) \mid e_1,\ \text{ord}(f) \mid e_2 \implies \text{ord}(f) \mid d \implies f \mid x^d - 1$. $\square$

Teorema 3.7. *Sia* $\mathbb{F}_q$ *un campo di caratteristica* p*, e* $g(x)$ *un polinomio irriducibile. Allora* $ordg^b = ord(g)p^t$*, dove* t *è il più piccolo intero pr cui* $p^t \geq b$.

Dimostrazione. Sia $f = g^b$, $c = \text{ord}f$ ed $e = \text{ord}g$. Vogliamo dunque mostrare che $c = ep^t$.

Anzitutto, possiamo notare che $g \mid f \implies e \mid c$ per il Lemma 3.5. Inoltre, e non è un multiplo di p, poiché $e \mid q^m - 1$ dove m è il grado di g.

Dunque, $g \mid x^e - 1 \implies g^b \mid (x^e - 1)^b \implies f \mid (x^e - 1)^b \implies f \mid (x^e - 1)^{p^t}$. Quindi $\text{ord}(f)$ è della forma ep^u, per qualche $0 \leq u \leq t$. Tuttavia, poiché tutte le radici di $(x^e - 1)$ sono semplici, visto che $p \nmid e$, si ha che le radici di $(x^{ep^u} - 1)$ hanno tutte molteplicità p^u, ma $g^b \mid (x^{ep^u} - 1) \implies p^u \mid b \implies u \geq t \implies u = t$. □

Teorema 3.8. *Sia $f = g_1 \cdots g_k$ con g_i polinomi relativamente primi tra loro, non identicamente nulli. Allora $ord(f) = [ord(g_1), \ldots, ord(g_k)]$.*

Dimostrazione. Anzitutto, si nota che se $x \mid g_i$ non cambia nulla, quindi possiamo assumere $g(0) \neq 0$.

Fisso $e = \text{ord}(f)$, $e_i = \text{ord}(g_i)$, $c = [e_1, \ldots, e_k]$. $e_i \mid c \implies g_i \mid x^c - 1 \implies f = \prod_i g_i \mid x^c - 1$ poiché i g_i sono coprimi tra loro per ipotesi. Segue che $\text{ord}(f) \mid c$.

D'altra parte, $g_i \mid f \mid x^e - 1 \implies e_i \mid e \implies c \mid e \implies c \mid \text{ord}(f)$. □

Dai due teoremi precedenti, segue che se $f \in \mathbb{F}_q$ è un polinomio, $f = ag_1^{e_1} \cdots g_k^{e_k}$ la sua fattorizzazione canonica. Allora $\text{ord}(f) = [\text{ord}(g_1), \ldots, \text{ord}(g_k)]p^t$ dove p è l'ordine del campo e t è il minimo intero positivo tale per cui $p^t \geq \max\{e_1, \ldots, e_k\}$.

Teorema 3.9. $f \in \mathbb{F}_q[x] \implies ord(f) = ord(f^*)$.

Dimostrazione. Nota: il reciproco di $f(x)$, $x^{\deg(f)}f(x^{-1})$ è denotato con f^*.

Se $f(0) \neq 0$, allora $f \mid x^e - 1 \iff f^* \mid x^e - 1$. Altrimenti, $f = x^h g$ con $g(0) \neq 0$ e $\text{ord}(f) = \text{ord}(g) = \text{ord}(g^*) = \text{ord}(f^*)$.

La storia $f \mid x^e - 1 \iff f^* \mid x^e - 1$ si dimostra facilmente usando la definizione di prodotto inteso come sequenza dei coefficienti.

$$f \mid x^e - 1 \iff \exists g \text{ s.t. } fg = x^e - 1. \text{ Allora,}$$
$$(fg)_i = \sum_{k+j=i} f_k g_j \quad \text{e} \quad (f^*g^*)_i = \sum_{k+j=i} f_k^* g_j^* = \sum_{k+j=k} f_{n-k} g_{n-j} = \sum_{k+j=n-i} f_k g_j.$$

Ma $n = \deg(fg) = e$, ed $x^e - 1$ è palindromo, se visto come vettore dei suoi coefficienti. □

Definizione 3.10 (Polinomio Primitivo)**.** Un polinomio di grado m su $\mathbb{F}_q[x]$ è detto *primitivo* se è il polinomio minimo di un elemento primitivo di $\mathbb{F}_{q^m}$.

Teorema 3.11 (Caratterizzazione dei polinomi primitivi)**.** *Un polinomio primitivo di grado m su $\mathbb{F}_q$ è tale sse il suo ordine è $q^m - 1$ per il Teorema 3.3.*

Dimostrazione. **"$\implies$"** f primitivo $\implies$ f polinomio minimo di α generatore del gruppo ciclico $\mathbb{F}_{q^m} \implies$ f monico, irriducibile $\implies$ f monico, $f(0) \neq 0$, e ha ordine uguale all'ordine di α, cioé $q^m - 1$.

"$\impliedby$" Se $\text{ord}(f) = q^m - 1$, e per assurdo f fosse riducibile, allora o $f = g^b \implies \text{ord}(f) = \text{ord}(g)p^t \implies p \mid \text{ord}(f)$ che è assurdo; oppure $f = g_1 g_2$ con $\text{ord}(g_i) \leq q^m - 1$ (Teorema 3.1) e di grado $m_i = \deg(g_1)$, e allora $\text{ord}(f) = \text{ord}(g_1 g_2) \leq \text{ord}(g_1)\text{ord}(g_2) \leq (q^{m_1} - 1)(q^{m_2} - 1) < q^{m_1+m_2} - 1 = q^m - 1 = \text{ord}(f)$ che è ancora un assurdo.

□

Teorema 3.12. *Sia $f \in \mathbb{F}_q[x]$ tale che $f(0) \neq 0$. Allora $x^r \equiv a \pmod f$ con $a \in \mathbb{F}_q^* \implies ord(f) = rh$ dove $h = ord(a)$.*

Dimostrazione. Pongo $e = \text{ord}(f)$. Da un lato abbiamo che $x^{hr} = a^h = 1 \implies e \leq hr$.

Dall'altro $x^r \equiv a \implies x^e \equiv x^{rs+t} = a^s x^t \equiv 1$ con $s \in \mathbf{N}$ e $0 \leq t < r \implies x^t \equiv a^{-s} \implies t = 0$ per ipotesi sul resto $\implies x^{sr} \equiv a^s \equiv 1 \pmod f \implies s \geq h \implies e \geq hr$. □

Teorema 3.13. *Sia $f \in \mathbb{F}_q[x]$ un polinomio monico di grado $m \geq 1$. Allora, f è primitivo sse $(-1)^m f(0)$ è un elemento primitivo di $\mathbb{F}_q$. Inoltre, il minimo intero r per cui $x^r \equiv a \pmod f$, per qualche $a \in (\mathbb{F}_q)^\times$ è $r = (q^m - 1)/(q - 1)$.*

Dimostrazione. " $\implies$ ". Sia f primitivo, e sia α una sua radice. Allora

$$N_{\mathbb{F}_{q^m}/\mathbb{F}_q}(\alpha) = (-1)^m f(0) = \alpha^{\frac{q^m-1}{q-1}} \tag{3.1}$$

da cui segue che l'ordine di $(-1)^m f(0)$ è $q-1$ su $\mathbb{F}_q$, pertanto è un elemento primitivo di $\mathbb{F}_q$. Inoltre, poiché f è il polinomio caratteristico di α su $\mathbb{F}_q$:

$$x^{\frac{q^m-1}{q-1}} = (-1)^m f(0) \pmod f.$$

Ciò implica che $r \leq (q^m-1)/(q-1)$, ma $q^m - 1 = \mathrm{ord}(\alpha) = \mathrm{ord}(f) \leq (q-1)r \implies r \geq (q^m-1)/(q-1)$.

" $\impliedby$ ". Date le proprietà su f di sopra, affermo che f è irriducibile. Se per assurdo non lo fosse, vi sarebbero $f_1 \cdots f_k$ polinomi monici irriducibili che fattorizzano f. Sia allora $m_i = \deg(f_i)$, quindi $\mathrm{ord} f \mid q^m - 1$ (questo per ogni $1 \leq i \leq k$), e posto

$$d = (q^{m_1}-1)(q^{m_2}-1)\cdots(q^{m_k}-1)/(q-1)^{k-1},$$

abbiamo che $\forall i\ \mathrm{ord} f_i \mid d \implies f \mid x^d - 1$. Se $k \leq 2$ allora $d < (q^{m_1+m_2+\cdots m_k}-1)/(q-1) = (q^m-1)/(q-1) = r$, che va in contraddizione proprio con la definzione di r. Dobbiamo conseguirne che $k = 1$ e duque f è irridubile. Poiché irriducibile, possiamo usare lo stesso argomento del 3.1 per dimostrare che data β radice di f, allora $\beta^r = (-1)^m f(0)$, e dunque

$$x^r \equiv (-1)^m f(0) \pmod f.$$

Poiché l'ordine di $(-1)^m f(0)$ è $q-1$, segue che $\mathrm{ord} f = q^m - 1$, dunque f è primitivo su $\mathbb{F}_q$. □

3.1 Polinomi Irriducibili (ancora)

Teorema 3.14. *Per ogni campo $\mathbb{F}_q$ e ogni intero positivo n il prodotto di tutti i polinomi irriducibili su $\mathbb{F}_q$ il cui grado divide n è $x^{q^n} - x$.*

Dimostrazione. I polinomi irriducibili che dividono $x^{q^n} - x$ sono solo e soltanto quelli il cui grado divide n per la Proposizione 2.11. Inoltre, la derivata prima di $x^{q^n} - x$ è -1, quindi non vi sono radici doppie. □

Segue immediatamente che sapendo il numero di polinomi irriducibili,

Corollario 3.15. *Se $N_q(d)$ è il numero di polinomi irriducibili su $\mathbb{F}_q[x]$ di grado d, allora*

$$q^n = \sum_{d \mid n} d N_q(d) \quad \forall n \in \boldsymbol{N}.$$

Definizione 3.16 (μ di Moebius)**.** La funzione $\mu : \mathbf{N} \to \mathbf{N}$ di Moebius è definita come:

$$\mu(n) = \begin{cases} 1 & \text{se } n = 1, \\ (-1)^k & \text{se } n = p_1 \cdots p_k \text{ primi distinti}, \\ 0 & \text{altrimenti.} \end{cases}$$

Segue immediatamente, semplicemente applicando la definizione, che se $n = 1$ tale sommatoria è uguale a 1. Altrimenti, per $n > 1$:

$$\begin{aligned} \sum_{d \mid n} \mu(d) &= \mu(1) + \sum_{i \leq k} \mu(p_i) + \sum_{0 \leq i_1 < i_2 \leq k} \mu(p_{i_1} p_{i_2}) + \cdots + \mu(p_1 p_2 \ldots p_k) \\ &= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k \\ &= (1 + (-1))^k \\ &= 0 \end{aligned}$$

dove $p_1, \ldots, p_k$ non tutti quei fattori primi che compaiono una volta sola in n.

Osservazione 3.17. Per quanto detto precedentemente:

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1 & \text{se } n = 1 \\ 0 & \text{altrimenti} \end{cases}$$

Teorema 3.18 (Formula di Inversione di Moebius)**.** *Siano $h, H : \boldsymbol{N} \to (G, +)$ gruppo abeliano. Allora:*

$$H(n) = \sum_{d|n} h(d) \quad \forall n \in \boldsymbol{N} \iff h(n) = \sum_{d|n} \mu\Big(\frac{n}{d}\Big) H(d) = \sum_{d|n} \mu(d) H\Big(\frac{n}{d}\Big) \quad \forall n \in \boldsymbol{N}.$$

Equivalentemente, nella sua versione moltiplicativa:

$$H(n) = \prod_{d|n} h(d) \quad \forall n \in \boldsymbol{N} \iff h(n) = \prod_{d|n} H(d)^{\mu(n/d)} = \prod_{d|n} H(n/d)^{\mu(d)} \quad \forall n \in \boldsymbol{N}.$$

Dimostrazione. Fissato un n qualunque,

$$\begin{aligned}
&\sum_{d|n} \mu(d) H\Big(\frac{n}{d}\Big) = \\
&\qquad = \sum_{d|n} \mu(d) \sum_{c|n/d} h(c) && \text{[applicando LHS]} \\
&\qquad = \sum_{c|n} \sum_{d|n/c} \mu(d) h(c) && \text{[tanto la somma è commutativa nel gruppo]} \\
&\qquad = \sum_{c|n} h(c) \sum_{d|n/c} \mu(d) \\
&\qquad = h(n) && \text{[tutti i termini sono nulli eccetto } c = n \text{ dove } \mu(1) = 1]
\end{aligned}$$

L'altro verso si mostra in modo simile. La versione moltiplicativa è analoga, con un cambio di notazione. □

Seguono due importanti applicazioni della formula di inversione

Teorema 3.19 (Numero di Polinomi Irriducibili)**.** *Il numero di polinomi irriducibili $N_q(d)$ su $\mathbb{F}_q$ di grado d è*

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d$$

Dimostrazione. Partendo da $q^n = \sum_{d|n} d N_q(d)$, poni $H = q^n$ e $h = dN_q(d)$ e usa la formula di inversione. □

Corollario 3.20. *Per ogni intero $n > 0$ ed ogni campo $\mathbb{F}_q$, esiste un'estensione $\mathbb{F}_{q^n}$.*

Dimostrazione. $N_q(n) \geq 1/n(q^n - q^{n-1} - q^{n-2} - \cdots - 1) = 1/n(q^n - (q^n - q)/(q-1)) > 0.$ □

Teorema 3.21. *L'n-esimo polinomio ciclotomico su un campo K di caratteristica p non divisibile con n è*

$$Q_n(x) = \prod_{d|n} (x^d - 1)^{\mu(n/d)}.$$

Dimostrazione. Parti da $x^n - 1 = \prod_{d|n} Q_d$, poni $H(n) = x^n - 1$ e $h(d) = Q_d$ definite sul gruppo moltiplicativo delle funzioni razionali su K. □

Teorema 3.22. *Il prodotto $I(q, n) \in \mathbb{F}_q[x]$ di tutti i polinomi monici irriducibili di grado n è dato da*

$$I(q, n) = \prod_{d|n} (x^{q^n} - x)^{\mu(n/d)}$$

Dimostrazione. Parti da $x^{q^n} - x$ è il prodotto di tutti i polinomi irriducibili di grado divisore di n presi una volta sola, poni $H(n) = x^{q^n} - x$ e $h(d) = I(q, d)$. □

Teorema 3.23. *Per $n > 1$ vale*

$$I(q, n) = \prod_m Q_m,$$

dove il prodotto è esteso a tutti gli m per cui n è l'ordine moltiplicativo di $q \pmod m$.

Dimostrazione. Definisco $S := \{\, \alpha \in \mathbb{F}_{q^n} \mid \text{il grado di } \alpha \text{ su } \mathbb{F}_q \text{ è } n \,\}$. Allora

$$I(q,n) = \prod_{\alpha \in S} (x - \alpha).$$

Infatti, per ogni $\alpha \in S$ esiste un polinomio monico irriducibile di grado n (il suo polinomio minimo), e d'altra parte, se α è radice di $I(q,n)$, allora vi è un polinomio monico irriducibile di cui α è radice, dunque $\alpha \in S$.

Poiché $\alpha \in \mathbb{F}_{q^n}^*$, l'ordine di α divide $q^n - 1$ per il Teorema di Lagrange, quindi $\text{ord}(\alpha) = m$ è tale che n è il minimo intero per cui $q^n \equiv 1 \pmod m$. Dunque n è l'ordine moltiplicativo di $q \pmod m$.

A questo punto considero le partizioni $S \supset S_m = \{\alpha \in S : \text{ord}(\alpha) = m\}$, e pertanto:

$$I(q,n) = \prod_{\alpha \in S} (x - \alpha) = \prod_m \prod_{\alpha \in S_m} (x - \alpha) = \prod_m Q_m$$

per la definizione di polinomio ciclotomico (sono tutte radici con lo stesso ordine). □

3.2 Costruzione di Polinomi Irriducibili

Teorema 3.24. *Se m è l'ordine di $s \pmod e$, allora mt è l'ordine di $s \pmod{et}$, ammesso che $s, t \geq 2$, ogni fattore primo di t divide e, ma non $\frac{s^m - 1}{e}$ e che $t \equiv 0 \pmod 4 \implies s^m \equiv 1 \pmod 4$.*

Dimostrazione. Per induzione sui fattori di t contatti con la molteplicità.

- t primo. Pongo $d = (s^m - 1)/e$ e abbiamo che $s^m \equiv 1 \pmod e \implies s^m = de + 1 \implies$

$$s^{mt} = (de+1)^t = d^t e^t + \binom{t}{1}(de)^{t-1} + \cdots + \binom{t}{t-1}(de) + 1 \implies s^{mt} \equiv 1 \pmod{et}$$

  poiché il tutti gli addendi eccetto l'ultimo sono divisibili per et: infatti, per il primo $t \mid e \implies te \mid e^t$, e gli altri si osserva il coefficiente binomiale. Abbiamo quindi che l'ordine $k \mid mt$. D'altra parte, $s^m \equiv 1 \pmod{et} \implies s^m \equiv 1 \pmod e$ (è immediato applicando la definizione di divisione) e quindi $m \mid k$. Poiché t è primo, segue che $k = m \veebar k = mt$. Il primo caso è impossibile poiché $s^m \equiv 1 \pmod{et} \implies s^m - 1 = de \equiv 0 \pmod{et} \implies et \mid de \implies t \mid d$ che è una contraddizione con le ipotesi.

- $t = rt_0$ con r fattore primo. Per ipotesi induttiva $s^{mr} \equiv 1 \pmod{er}$, vogliamo verificare le ipotesi del teorema per usare il passo induttivo ancora su t_0. Nello specifico, vogliamo mostrare che ogni fattore primo di t_0 divide er, ma non divide $d_0 = (s^{mr} - 1)/er$.

  Il primo predicato è banale. Per il secondo, abbiamo che $(s^m - 1)^r = (s^m - 1)c$ con $c = s^{m(r-1)} + \cdots + 1 \implies d_0 = \frac{s^{mr}-1}{er} = dc/r$. Si può facilmente mostrare che c/r è un intero: $e \mid s^m - 1$ ed $r \mid e$ per ipotesi $\implies s^m \equiv 1 \pmod r$ e dunque ogni addendo in c vale 1, quindi $c \equiv r \equiv 0 \pmod r$. Quindi basta mostrare che $t_0 \nmid (c/r)$ poiché sappiamo già che $t_0 \mid t$ e $t \nmid d$ e il "divide" è una relazione di equivalenza.

  Notiamo $s^m \equiv 1 \pmod{r_0}$ $\forall r_0$ fattore primo di t_0 per ipostesi, e dunque $c \equiv r \pmod{r_0}$.

  Se $r \neq r_0$ abbiamo finito, perché $c/r \equiv 1 \pmod{r_0} \implies r_0 \nmid c/r$ e possiamo immediatamente applicare il teorema.

  Se $r = r_0 \implies s^m = 1 + br \pmod{r^2}$ per qualche $b \in \mathbf{Z} \implies s^{mj} = (1+br)^j = 1 + jbr$ $\forall j \geq 0$ perché tutti gli altri fattori sono divisibili per r^2. Quindi $c = s^{m(r-1)} + \cdots + 1 = r + br\sum_{j=0}^{r-1} j = r + br\frac{r(r-1)}{2}$. A questo punto, se r è dispari allora $c/r \equiv 1 \pmod r$ e abbiamo finito poiché $r_0 = r \nmid c/r$; se r pari abbiamo $r_0 = r = 2$ quindi $t \equiv 0 \pmod 4 \implies s^m \equiv 1 \pmod 4$ per ipotesi, e $c = s^m + 1 \equiv 2 \pmod 4 \implies c/r = c/2 = 1 \pmod{r=2} \implies r_0 \nmid c/r$. □

Teorema 3.25. *Siano $f_1(x), \ldots, f_N(x)$ tutti i distinti polinomi irriducibili di grado m e ordine e su $\mathbb{F}_q$. Sia $t \geq 2$ un intero tale che i suoi fattori primi dividono e ma non $(q^m - 1)/e$. Assumiamo anche che $t \equiv 0 \pmod 4 \implies q^m \equiv 1 \pmod 4$. Allora $f_1(x^t), \ldots, f_N(x^t)$ sono tutti i distinti polinomi irriducibili di grado mt e di ordine et su $\mathbb{F}_q$.*

Dimostrazione. Mostriamo che i $f_i(x^t)$ sono irriducibili: $f_i(x)$ ha le radici e-esime dell'unità, quindi $f_i(x) \mid Q_e(x) \implies f_i(x^t) \mid Q_e(x^t) = Q_{et}(x)$ per il Teorema 2.42 e ha grado mt che è l'ordine di $q \pmod{et}$ per il teorema precedente, quindi è irriducibile.

Mostriamo che i $t_i(x^t)$ sono gli unici irriducibili. $N = \varphi(e)/m$ sono i polinomi irriducibili di ordine e e grado $m \implies \frac{\varphi(et)}{mt} = \frac{\varphi(e)t}{mt} = N$ sono i polinomi irridicibili in $\mathbb{F}_q$ di grado mt e ordine et. □

Dato un polinomio irriducibile di ordine e, è possibile trovare tutti i polinomi irriducibili il cui ordine divide e. Possiamo restringerci al caso in cui $g(0) \neq 0$, poiché $g(x) = x$ sarà sempre tra tali polinomi. Sia α una radice di f, e per ogni $n \in \mathbf{N}$, $g_t(x)$ il polinomio minimo per α^t su $\mathbb{F}_q$. Sia $T = \{ t_1, t_2, \ldots, t_n \}$ l'insieme di tutti gli interi positivi per cui $\forall n \in \mathbf{N} \;\; \exists! t_j \;.\; t \equiv t_j \cdot q^b \pmod{e}$ per qualche $b \geq 0$.

Teorema 3.26. *Con la notazione di sopra, i polinomi $g_{t_1}, \cdots, g_{t_n}$ sono tutti i polinomi monici irriducibili il cui ordine divide e e $g(0) \neq 0$.*

Dimostrazione. Per costruzione, sappiamo già che i $g_{t_i}(x)$ sono polinomi monici, irriducibili, il cui ordine divide e e per cui $g_{t_i}(0) \neq 0$.

Siamo interessati a mostrare che essi sono tutti i polinomi di quel tipo. Sia β una radice di g_i, allora ha ordine d un divisore di e, allora è una radice e-esima dell'unità ($\beta^d = 1 \implies \beta^e = 1$), e pertanto posso esprimerla come α^t, dove α è una radice primitiva e-esima dell'unità, per qualche $t \in \mathbf{N}$. Allora $\beta = \alpha^t = [\alpha^{t_i}]^{q^b}$ è un coniugato di α^{t_j} e pertanto radice di α_{t_j}.

Siamo interessati a mostrare che essi sono distinti. Sia per assurdo $g_{t_i} = g_{t_j}$. Allora α^{t_i} e α^{t_j} sono entrambe radici dei due polinomi, pertanto $\alpha^{t_j} = \alpha^{t_i q^b}$. Ma questo è un assurdo per la costruzione di T, per cui dovrebbe essere esistere un unico $t_j \;.\; t \equiv t_j q^b \pmod{e}$, mentr in questo caso abbiamo

$$t_j \equiv t_j q^0 \pmod{e} \qquad \wedge \qquad t_j \equiv t_i q^b \pmod{e}. \qquad \square$$

È possibile orrservare che $k = deg g_t$ è l'ordine moltiplicativo di $q \pmod{t}$, e $d = \mathrm{ord} g_t = \mathrm{ord} \alpha^t = e/(e,t)$ per le proprietà dei gruppi ciclici. Peratnto esiste un modo per calcolare il grado e l'ordine dei vari g_{t_i}.

Teorema 3.27. *Sia f un polinomio monico irriducibile su $\mathbb{F}_q[x]$ di grado m. Sia $\alpha \in \mathbb{F}_{q^m}$ una sua radice, e $\forall t \in \boldsymbol{N}$ f_t il polinomio caratteristco di $\alpha^t \in \mathbb{F}_{q^m}/\mathbb{F}_q$. Allora*

$$f_t(x^t) = (-1)^{m(t+1)} \prod_{j=1}^{t} f(\omega_j x),$$

dove $\omega_1, \ldots, \omega_t$ sono le t-esime radici dell'unità su $\mathbb{F}_q$ contate con le rispettive molteplicità.

Dimostrazione. Siano $\alpha_1, \ldots, \alpha_m$ le radici di f. Per la proprietà universale dei polinomi, $\alpha_1^t, \ldots, \alpha_m^t$ sono le radici di $f(x^t)$. Notiamo che:

$$f(x^t) = \prod_{t=1}^{m} (x^t - \alpha^t) = \prod_{t=1}^{m} \prod_{j=1}^{t} (x - \alpha_i \omega_j) = \prod_{t} \prod_{j} \omega_j (\omega_j^{-1} x - \alpha_t)$$

Si noti che $x^t - 1 = \prod_j (x - \omega_j)$ implica che $\prod_j = (-1)^{t+1}$, ed è facile convincersene usando la definizione di norma. Pertanto:

$$f_t(x^t) = (-1)^{m(t+1)} \prod_{j}^{t} \prod_{i}^{m} (\omega_j^{-1} x - \alpha_j) = (-1)^{m(t+1)} \prod_{j}^{t} f(\omega_j^{-1} x) = (-1)^{m(t+1)} \prod_{j}^{t} f(\omega_j x)$$

poiché scorriamo su tutte le possibili j. $\square$

Teorema 3.28. *Sia f un polinomio irriducibile su $\mathbb{F}_q$ di grado n e sia $k \in \boldsymbol{N}$. Allora f fattorizza in d polinomi irriducibili su $\mathbb{F}_{q^k}[x]$ dello stesso grado n/d, dove $d = (k, n)$.*

Dimostrazione. Assumiamo il caso $f(0) \neq 0$, altrimenti è banale. Sia g un fattore irriducibile di f su $\mathbb{F}_{q^k}$. Allora, g ha ordine uguale a quello di una sua qualunque sua radice, che è uguale all'ordine di f. Ancora poiché tale radice è una e-esima radice dell'unità, sta nell'n-esimo campo ciclotomico dove n è l'ordine moltiplicativo di $q \pmod{e}$, e il grado di g è l'ordine moltiplicativo di $q^k \pmod{e}$, ossia e/d con $d = (n, k)$. $\square$

Corollario 3.29. *Un polinomio irriducibile di grado n su $\mathbb{F}_q[x]$ rimane irriducibile in $\mathbb{F}_{q^k}[x]$ se n, k sono coprimi.*

Capitolo 4

Fattorizzazione di polinomi

All'interno di questo capitolo si assume che $f \in \mathbb{F}_q[x]$ sia un polinomio di grado almeno 1 del quale vorremmo conoscere la fattorizzazione, ossia i rispettivi fattori irriducibili con le molteplicità. Ci sono stati menzionati due algoritmi, con applicazioni diverse in termini di efficienza. Il primo è l'algoritmo di Berlekamp, che permette di fattorizzare in campi con caratteristica piccola; il secondo `TODO: nome?` funziona meglio quando la caratteristica è grande.

4.1 L'algoritmo di Berlekamp

Vogliamo anzitutto ridurci al problema di fattorizzazione quando f non ha radici doppie, e per portarci in questo caso calcoliamo $d = (f, f')$.

- se $d = 1$ allora f non ha fattori ripetuti e abbiamo finito;
- se $d = f$ allora $f' = 0$ e f è della forma $f = g(x^p)$ dove p è la caratteristica di $\mathbb{F}_q$;
- se $d \neq 1, f$ allora d è un fattore non banale e possiamo procedere ricorsivamente fattorizzando separattamente $d, f/d$.

Prima o poi si convergerà al primo caso.

Teorema 4.1. *$f \in \mathbb{F}_q[x]$ monico e $h \in \mathbb{F}_q[x]$. $h^q \equiv h \pmod f$. Allora*

$$f = \prod_{c \in \mathbb{F}_q} \gcd(f, h - c) \tag{4.1}$$

Dimostrazione. Ogni fattore sulla destra divide banalmente f per definizione di gcd. D'altro canto, tutti quei fattori sono relativamente primi tra loro, per cui $\prod_c \gcd(f, h - c) \mid f$.

Si nota che $f \mid h^q - h = \prod_c (h - c)$ da cui segue che f divide la parte destra di 4.1. □

Diremo che un polinomio è f-riducibile se genera una fattorizzazione non banale di f.

Se $f = f_1 \cdots f_k$ prodotto di polinomi irriducibuli coprimi tra loro, allora per il teorema cinese dei resti sappiamo che, data una qualunque tupla $(c_1, \ldots, c_k)$ esiste un polinomio h soluzione del sistema

$$\begin{cases} h(x) = c_1 \pmod{f_1} \\ h(x) = c_2 \pmod{f_2} \\ \quad \vdots \\ h(x) = c_k \pmod{f_k} \end{cases}$$

dove $\deg h < \deg f$. Ragionando sul teorema precedente, possiamo notare che tutti i polinomi irriducibili di f dividono un qualche $h - c$ con $c \in \mathbb{F}_q$, ma ancora di più h deve avere proprio le proprietà dell'ipotesi:

$$h^q(x) = c_i^q = c_i = h(x).$$

Rimane capire come trovare un polinomio $h = a_0 + \cdots + a_{n-1}x^{n-1} \in \mathbb{F}_q[x]$ con tali proprietà, e questo possiamo farlo in tempo cubico. Si nota che:

$$h(x) = \sum_i a_i x^i \equiv h^q(x) = \sum_i a_i x^{iq} = \sum_i \sum_j a_i b_j x^i \pmod{f_i},$$

dove

$$x^{iq} = \sum_{j=0}^{n-1} b_{ij} x^j \pmod{f}.$$

Posto allora $n = \deg f$ creo la matrice $n \times n$ $B = (b_{ij})_{ij}$ con $0 \leq i, j \leq n-1$, e i fattori di h soddisfano:

$$(a_0, \ldots, a_{n-2}, a_{n-1})B = (a_0, \ldots, a_{n-2}, a_{n-1})$$

che si risolve subito applicando il teorema spettrale. Calcolo $\ker(B-I)$. La sua dimensione è in numero di fattori irriducibili di tale polinomio. Essa non è mai nulla poiché $h = 1$ è sempre soluzione. Se la dimensione è 1 allora $h = 1$ è l'unica soluzione e quindi non esistono fattori non banali, quindi f è irriducibile. Altrimenti, mi prendo una soluzione qualunque non unitaria, e trovo i fattori non banali $\gcd(f, h-c)$ sui quali applico ricorsivamente l'algoritmo.

Applicazione. Riporto di sotto un'implementazione MAGMA dell'algoritmo:

```
function Berlekamp(f)
    K<x> := Parent(f);
    c := Characteristic(K);
    n := Degree(f);
    M := MatrixRing(K, n);

    gcd := GCD(f, Derivative(f));
    if gcd eq f then
        reduced := Root(f, c);
        return &cat[Berlekamp(K ! reduced): i in [1..c]];
    elif gcd ne 1 then
        return Berlekamp(K ! gcd) cat Berlekamp(K ! (f/gcd));
    else
        x_iq := [x^(c * i) mod f: i in [0 .. n-1]];
        B := M ! [[Coefficient(p, j): j in [0 .. n-1]] : p in x_iq];
        ker := Kernel(B-Id(M));
        if Dimension(ker) eq 1 then
            return [f];
        else
            basis := [K ! Eltseq(h): h in Basis(ker)];
            h := [b: b in basis | b ne 1][1];
            factors := [K ! GCD(f, h - a) : a in GF(c)];
            return &cat[Berlekamp(factor): factor in factors | factor gt 1];
        end if;
    end if;
end function;
```

Capitolo 5

Sequenze Linearmente Ricorrenti

Definizione 5.1 (Sequenza)**.** Una *sequenza* $s_0, s_1, \dots$ di elementi in $\mathbb{F}_q$ è detta *linear recurring seuqence* se

$$s_{n+k} = a_{k-1}s_{n+k-1} + a_{k-2}s_{n+k-2} + \cdots + a_0 s_n + a \qquad \forall n \in \mathbf{N} \tag{5.1}$$

L' ordine di tale sequenza è fissato a k, e la relazione 5.1 è detta *linear recurring sequence releation.* Se la sequenza è tale che $a = 0$ allora è detta *omogenea* ed è del genere che abbiamo visto a crittografia, altrimenti è detta *inomogenea.*

Possiamo implementare questo oggetto con dei circuiti chiamati LFSR.

Definizione 5.2 (Ultimately Recurring Sequence)**.** Sia $s_0, s_1, \dots$ una sequenza. Se esistono interi $r > 0$ e $n_0 \geq 0$ tali che $s_{n+r} = s_n \ \forall n \geq n_0$ allora la sequenza è detta *ultimately periodic.* r è detto *periodo* ed n_0 *preperiodo.* Il più piccolo tra tutti i periodi è detto *periodo minimo* della sequenza.

Lemma 5.3. *Ogni periodo di una* ultimately periodic sequence *è divisibile per il periodo minimo.*

Dimostrazione. È immediato usando la definizione di resto: preso un qualunque periodo r_0 che funziona dopo n_0 passi - i.e. $s_{n+r_0} = s_n \ \forall n \geq n_0$- e il periodo minimo r_1 che funziona dopo n_1 passi, allora $r_0 = r_1 q + t$ per qualche q intero e $0 \leq t < r$ e

$$\forall n \geq \max(n_0, n_1) \qquad s_n = s_{n+r_0} = s_{n+r_1 q+t} = s_{n+t}$$

che implica $t = 0$, altrimenti vi sarebbe un assurdo. □

Definizione 5.4 (Sequenza Periodica)**.** Una sequenza è detta *periodica* se è *ultimately periodic* con preperiodo 0, ossia $s_{n+r} = s_n \quad \forall n \in \mathbf{N}$, dove r è il periodo.

Teorema 5.5. *Sia $\mathbb{F}_q$ un campo finito e k un intero positivo. Allora ugni sequenza di ordine k è ultimately periodic con periodo $r \leq q^k$, e $r \leq q^k - 1$ se la sequenza è omogenea.*

Dimostrazione. Notiamo che esistono q^k k-uple distinte di elementi su $\mathbb{F}_q$. Pertanto devono esistere s_i, s_j tali che $s_i = s_j$ (con $0 \leq i \leq j \leq q^k$). Per induzione si può dimostrare che $s_{n+j-i} = s_n \ \forall n \geq i$:

- se $i = n$ è immediato: $s_{i+j-i} = s_j = s_i$;
- se $s_{n+j-i} = s_n \ \forall i \leq n < n'$ allora

$$s_{n'+j-i} = f(s_{n'+j-i-1}, \dots, s_{n'+j-i-k}) = f(s_{n'-1}, \dots, s_{n'-k}) = s_n,$$

da cui segue che il periodo minimo è $r \leq j - i \leq q^k$. Nel caso in cui la sequenza sia omogenea uno può usare lo stesso argomento usando $q^k - 1$ come maggiorante, visto che lo $\mathbf{0}$-vettore è tale che $f(\mathbf{0}) = \mathbf{0}$. In tal caso, infatti, il periodo minimo sarebbe 1. □

Teorema 5.6. *Sia $s_0, s_1, \dots$ una sequenza linearmente ricorrente su un campo finito $\mathbb{F}_q$. Se $a_0 \neq 0$ allora la sequenza è periodica.*

Dimostrazione. Abbiamo già visto che è ultimately periodic per il teorema precedente; sia r il suo periodo minimo e n_0 il suo preperiodo.

Supponiamo per assurdo di avere $n_0 \geq 1$. Pongo $n = n_0 + r + k - 1$, e usando la 5.1 ottengo (dopo aver raggruppato per a_0):

$$\begin{aligned} s_{n_0+r-1} &= a_0^{-1}(s_{n_0+r-1+k} - a_{k-1}s_{n_0+r+k-2} - \cdots - a_1 s_{n_0+r+k-k} - a) \\ &= a_0^{-1}(s_{n_0-1+k} - a_{k-1}s_{n_0+k-2} - \cdots - a_1 s_{n_0} - a). \qquad \text{[togliamo il periodo]} \end{aligned}$$

D'altra parte:

$$s_{n_0-1} = a_0^{-1}(s_{n_0-1+k} - a_{k-1}s_{n_0-2+k} - \cdots - a_1 s_{n_0} - a$$

Dunque $s_{n_0+r-1} = s_{n_0-1}$ e pertanto il periodo minimo è $n_0 - 1$, che è una contraddizione. □

Usando la definizione di linear recurring sequence releation, esplicitiamo la relazione di una sequenza omogenea

$$s_{n+k} = a_{k-1}s_{n+k-1} + a_{k-2}s_{n+k-2} + \cdots + a_0 s_n \qquad \forall n \in \mathbf{N} \tag{5.2}$$

e cotruiamo per essa una matrice associata $k \times k$ su $\mathbb{F}_q$ definita come

$$A = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \cdots & 0 & a_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & a_{k-1} \end{pmatrix}$$

che appartiene a $GL(k, \mathbb{F}_q)$ sse $a_0 \neq 0$, poiché $\det A = (-1)^{k-1}a_0 \neq 0$.

Lemma 5.7. *Sia $s_0, s_1, \cdots$ una sequenza linearmente ricorrente con relazione come da 5.1. Allora*

$$s_n = s_0 A^n$$

Dimostrazione. Qui bisogna fare attenzione con la notazione. $s_n = (s_n, s_{n+1}, \ldots, s_{n+k-1})$. Da qui si può verificare immediatamente per induzione:

$$(s_n, s_{n+1}, \ldots, s_{n+k-1})A = (s_{n+1}, s_{n+2}, \ldots, s_{n+k}).$$ □

Teorema 5.8. *Sia $s_0, s_1, \ldots$ una sequenza lineare omogenea di ordine k su $\mathbb{F}_q$, tale che la matrice compagna sia nonsingolare (i.e. $a_0 \neq 0$). Allora il periodo minimo divide l'ordine della matrice associata in $GL(k, \mathbb{F}_q)$.*

Dimostrazione. Sia m l'ordine di A, allora $s_{m+n} = s_n A^m = s_n$. □

Corollario immediato di questo è che in una sequenza omogenea il periodo divide l'ordine $\operatorname{ord}(GL(k, \mathbb{F}_q)) = q^{k(k-1)/2}(q-1)(q^2-1)\cdots(q^k-1)$.

Si può notare che il caso omogeneo è strettamente legato al caso inomogeneo, basta infatti prendere due termini consecutivi di una seuqenza inomogenea e sottrarli: otterremo una sequenza omogenea. Alternativamente, possiamo definire una matrice associata come sopra, però di dimensione $(k+1) \times (k+1)$ e associando un altro vettore a $s_n = (1, s_n, \ldots, s_{n+k})$.

5.1 Sequenze Impulso-Risposta

Data una relazione di ricorrenza per una sequenza omogenea come quella in 5.2, possiamo estrarre quella (sequenza) avente periodo minimo massimo. Tale sequenza $d_0, d_1, \ldots$ è identificata unicamente dai suoi primi k valori ed è detta *sequenza impulso-risposta*. Le sequenze impulso-risposta, poiché devono avere il periodo minimo quanto più grande possibile, sono tutte della forma $d_0 = d_1 = \cdots = d_{k-2} = 0$, $d_{k-1} = 1$.

Lemma 5.9. *Sia $d_0, d_1, \ldots$ una sequenza impulso-risposta. Allora, $d_m = d_n \iff A^n = A^m$, dove A è la matice associata.*

Dimostrazione.

- "$\Longleftarrow$" $A^m = A^n \implies d_m = d_n$ per il Lemma 5.7;
- "$\Longrightarrow$" $d_m = d_n \implies d_{m+t} = d_{n+t}$ $(\forall t \geq 0) \implies d_t A^m = d_t A^n$ $(\forall t \geq 0) \implies A^m = A^n$ poiché sicuramente i primi k termini $d_0, d_1, \ldots, d_{k-1}$ sono linearmente indipendenti. □

Teorema 5.10. *Il minimo periodo di una sequenza lineare omogenea divide il periodo della corrispettiva sequenza impulso-risposta.*

Dimostrazione. Sia $s_0, s_1, \ldots$ la sequenza omogenea e $d_0, d_1, \ldots$ la corrispondente sequenza impulso-risposta. Siano n_0, r rispettivamente preperiodo e periodo della sequenza impulso-risposta. Allora,

$$\begin{aligned} A^{n+r} &= A^n \quad \forall n \geq n_0 && \text{[Lemma 5.9]} \\ s_{n+r} &= s_n \quad \forall n \geq n_0 && \text{[Lemma 5.7]} \\ r &\mid \text{periodo.} && \square \end{aligned}$$

Teorema 5.11. *Se $d_0, d_1, \ldots$ è una sequenza impulso-risposta in cui la relazione lineare (tipo la 5.2) soddisfa $a_0 \neq 0$, allora l'ordine della matrice associata è il periodo minimo della sequenza.*

Dimostrazione. Se r è il periodo minimo di $d_0, d_1, \ldots$ allora è immediato che $r \mid \text{ord}(A)$ per il Lemma 5.8.
D'altra parte $a_0 \neq 0$ quindi la sequenza è periodica, quindi $d_r = d_0 \implies A^r = A^0 = 1 \implies \text{ord}(A) \mid r$. $\square$

Teorema 5.12. *Sia $s_0, s_1, \ldots$ una sequenza lineare di ordine k su $\mathbb{F}_q$ con $a_0 \neq 0$, periodo r, e preperiodo n_0. Se esistono $s_{m_1}, \ldots, s_{m_k}$ con $m_j \geq n_0$ vettori di stato linearmente indipendenti su $\mathbb{F}_q$, allora la sequenza e la corrispondente sequenza impulso-risposta sono periodiche e hanno il medesimo periodo minimo.*

Dimostrazione. Abbiamo per ogni $1 \leq j \leq k$ che $s_{m_j} A^r = s_{m_j+r} = s_{m_j} \implies A^r = \mathbf{1}$ poiché questo succede per tutti gli m_j he sono linearmente indipendenti. Pertanto $s_r = s_0 A^r = s_0$, da cui segue che la sequenza è periodica.
D'altra parte $d_0, d_1, \ldots$ è tale che il suo periodo minimo è un multiplo di r, e contemporaneamente $d_r = d_0 A^r = d_0$, dunque il suo periodo minimo divide r. Ne dobbiamo conseguire che il suo periodo minimo è proprio r. $\square$

Definizione 5.13 (Polinomio Caratteristico)**.** Sia la sequenza omogenea $s_0, s_1, \ldots$ su $\mathbb{F}_q$ definita dalla relazione

$$s_{n+k} = a_{k-1}s_{n-1} + a_{k-2}s_{n-2} + \cdots + a_0 s_n \quad \forall n \in \mathbf{N} \tag{5.3}$$

dove $a_j \in \mathbb{F}_q$. Allora il polinomio

$$f(x) = x^k - a_{k-1}x^{k-1} - a_{k-2}x^{k-2} - \cdots - a_0 \ \in \mathbb{F}_q[x] \tag{5.4}$$

è detto *polinomio caratteristico* della sequenza. Si può notare facilmente che f è il polinomio caratteristico della matrice associata A (in realtà è proprio definito così, ed A è detta *matrice compagna*).

Teorema 5.14. *Sia $s_0, s_1, \ldots$ una sequenza lineare omogenea di ordine k su $\mathbb{F}_q$ con polinomio caratteristico $f(x)$. Se le radici di quest'ultimo $\alpha_1, \ldots, \alpha_k$ sono tutte distinte, allora:*

$$s_n = \sum_{j=1}^{k} \beta_j \alpha_j^n \quad \forall n \in \boldsymbol{N}$$

dove i β_j sono determinati dai primi k elementi.

Dimostrazione. Vogliamo anzitutto assicurarci che le soluzioni per β_j esistono. Abbiamo il sistema di equazioni:

$$\sum_{j=1}^{k} \alpha_j^n \beta_j = s_n \quad \forall 0 \leq n \leq k-1$$

la matrice associata al sistema è di Vandermonde, per cui il determinante è $\prod_{1 \leq i < j \leq k}(\alpha_i - \alpha_j) \neq 0$ per ipotesi sugli α_i. Inoltre, per lo regola di Cramer, tali β_i possono stare nel campo di spezzamento di $f(x)$. A questo punto rimane dimostrare che la formula è corretta:

$$\begin{aligned} & s_{n+k} - a_{k-1}s_{n+k-1} - a_{k-2}s_{n+k-2} - \cdots - a_0 s_n = 0 \\ \impliedby & \sum_{j=1}^{k} \beta_j \alpha_j^{n+k} - \alpha_{k-1} \sum_{j=1}^{k} \beta_j \alpha_j^{n+k-1} - \cdots - a_0 \sum_{j=1}^{k} \beta_j \alpha_j^n = 0 \\ \impliedby & \sum_{j=1}^{k} \beta_j \alpha_j^n f(\alpha_j) = 0 \quad \forall n \in \mathbf{N}. \end{aligned}$$

$\square$

Nota: esiste una formula simile anche se f ha radici con molteplicità al più uguale alla caratteristica di $\mathbb{F}_q$, e viene dimostrata nel 6.23 del Niederreiter.

Teorema 5.15. *Sia $s_0, s_1, \ldots$ una sequenza omogenea di ordine k su $K = \mathbb{F}_q$ il cui polinomio caratteristico è irriducibile. Sia α una radice sull'estensione $F = \mathbb{F}_{q^k}$. Allora $\exists!\theta \in F$.*

$$s_n = \mathrm{Tr}_{F/k}(\theta\alpha^n) \quad \forall n \in \boldsymbol{N}.$$

Dimostrazione. Dato che $(1, \alpha, \ldots, \alpha^{k-1})$ costituiscono una base per F, possiamo definire una mappa lineare $L(\alpha^n) = s_n \; \forall 0 \leq n \leq k-1$. Vogliamo mostrare che tale mappa è proprio $\mathrm{Tr}_{F/K}(\theta \cdot \alpha^n)$, ma questo è immediato per il Teorema 2.17. □

Salto il teorema 6.25.

Teorema 5.16. *L'ordine del polinomio caratteristico è uguale all'ordine della matrice associata, in una sequenza linearmente ricorrente omogenea con $a_0 \neq 0$.*

Dimostrazione. Poiché f è polinomio minimo per A, e d'altra parte A è la matrice compagna di f, esiste e minimo per cui $A^e = 1 \iff f \mid x^e - 1$. Il resto segue banalmente per definizione di ordine. □

Capitolo 6

Costruzione di Funzioni Booleane

Questo capitolo segue il testo "Boolean Functions for Cryptography and Error Correcting Codes".

Le funzioni booleane ricoprono ruoli rilevanti sia in crittografia che in teoria dei codici:

- in teoria dei codici, possiamo interpretare ogni codice 2^n come un insieme di funzioni booleane, gacché ogni funzione booleana è determinata univocamente dalla propria tabella di verità, e quindi ad ogni codice possiamo associare un vettore unico in 2^n.
- in crittografia simmetrica, tutte le trasformazione crittografiche, come le S-box o i generatori di numeri casuali, sono definite come composizione di funzioni booleane nonlineari.

Chiameremo l'iniseme delle *funzioni booleane* $\mathcal{B}_n$ i cui elementi sono le mappe $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$. D'ora in avanti, quando non diversamente specificato, indicheremo il campo dei bit con $\mathbb{F}$.

Definiamo la *funzione booleana elementare*

$$x_i : v \mapsto v_i \quad \text{proiezione sull' } i\text{-esima componente,}$$

da cui posso definire la mappa $X_I = \prod_{i \in I} x_i$, con la convenzione che $X_\emptyset = 1$.

Definizione 6.1 (Funzione Booleana)**.** La forma normale di una funzione booleana a n registri è una funzione:

$$f : \mathbb{F}_{2^n} \to \mathbb{F}_2 : (x_1, \dots, x_n) \mapsto \sum_{S \subset \{1,\dots,n\}} a_S X_S$$

tale forma è spesso detta *Algebraic Normal Form* (ANF). L'insieme di tutte queste possibile funzioni è chiamato $\mathcal{R}_n$, ed è l'insieme di tutti i polinomi square-free di $\mathbb{F}$ in n incognite, i.e.

$$\mathcal{R}_n = \mathbb{F}[x_1, \dots, x_n]/(x_1^2 + x_1, \dots, x_n^2 + x_n) = \{ \sum_{I \in \mathcal{P}(n)} \lambda_I X_I \mid \lambda_I \in \mathbb{F} \}$$

Proposizione 6.2. *Lo spazio dei polinomi square-free ad n variabili ha* 2^{2^n} *elementi.*

Dimostrazione. Intendiamo $\mathcal{R}_n = \langle X_I \rangle_{I \in \mathcal{P}(n)}$ come spazio vettoriale su $\mathbb{F}$, la tesi è immediata.

Una dimostrazione alternativa si può sviluppare banalmente dal teorema successivo. □

Proposizione 6.3. *Tutte le mappe* $\mathbb{F}_2^n \to \mathbb{F}_2$ *sono riducibili ad una ANF. Ossia,* $\mathcal{B}_n = \mathcal{R}_n$*.*

Dimostrazione. Data una mappa $f : \mathbb{F}_2^n \to \mathbb{F}_2$ possiamo isolare gli 1 e ricomporre le ANF per questi termini minimi usando l'interpolazione di Lagrange. Questo implica che lo spazio delle funzioni che manda un polinomio $p \in \mathcal{R}_n$ nella mappa $(x \in \mathbb{F}_2^n) \mapsto p(x)$ che vive su $\mathcal{B}_n$ è surgettiva. Poiché inoltre le due hanno la stessa cardinalità - si dimostra facilmente contando-, la mappa è bigettiva per il lemma dei cassetti.

Alternativamente, posso dimostrare che è surgettiva allo stesso modo, e mostrare che $h \in \mathcal{B}_n \ . \ h \equiv 0 \iff h = 0 \in \mathcal{R}_n$. Poiché la mappa è un morfismo, segue la tesi. □

Definizione 6.4. Data una funzione booleana $f \in \mathcal{B}_n$, possiamo definire la sua *truth table*. Fissato un ordinamento sui 2^n punti di $\mathbb{F}^n$

$$\underline{f} = (f(P_1), \dots, f(P_{2^n})).$$

Con un piccolo abuso di notazione, spesso useremo sottoinsiemi di $\mathcal{B}_n$ per indicare in realtà le loro truth table.

Definizione 6.5 (Funzione Booleana Affine)**.** Una *funzione booleana affine* è una funzione

$$f \in \mathcal{B}_n \ . \ f = \sum_{i=1}^{n} a_i x_i + a_0.$$

Lo spazio delle funzioni affine viene identificato con $\mathcal{A}_n$.

Si può osservare che lo spazio delle funzioni booleane affini su n variabili è ugualle allo spazio dei codici Reed-Muller su n di grado massimo 1, ossie $\mathcal{A}_n = RM(n,1)$ dove $RM(n,m) = \{\, \underline{f} \mid \deg f \leq m, f \in \mathbb{F}[x_1, \ldots, x_n] \,\}$ è appunto il codice Reed-Muller.

Definizione 6.6. Il *peso* di Hamming di una funzione $w(f) = |\,\{\, x \in \mathbb{F}_2^n : f(x) \neq 0 \,\}\,|$ è la cardinalità del supporto del codice. La *distanza* di Hamming $d(f,g) = |\,\{\, x \in \mathbb{F}_2^n : f(x) \neq g(x) \,\}\,|$ è pertanto uguale a $w(f+g)$.

Osservazione 6.7. Comunque presa $f \in \mathcal{B}_n$, $f+1$ ha distanza massima da f. Il peso di $w(X_I) = 2^{n-|I|}$, poiché un punto sta nel supporto sse ha $|I|$ punti fissi, e gli altri comunque presi.

Definizione 6.8 (Non-Linearità)**.** Sia $f \in \mathcal{B}_n$ una funzione booleana, si definisce *non-linearità* la distanza dalle funzioni affini: $N(f) = d(\underline{f}, \mathcal{A}_n) = \min_{\alpha \in \mathcal{A}_n} d(f, \alpha)$. Tale funzione ci indica la lontananza della mappa dalla sua migliore approssimazione affine.

Teorema 6.9. $f \in \mathcal{B}_n \implies N(f) \leq \min\{\, 2^n - w(f), w(f) \,\}$.

Dimostrazione. $\mathbf{0}, \mathbf{1} \in \mathcal{A}_n \implies d(f,0) = w(f) \ \wedge \ d(f,1) = 2^n - w(f)$. □

Osservazione 6.10. In particolare, si può osservare che $\alpha \in \mathcal{A}_n \ \wedge \ \alpha \neq \mathbf{0}, \mathbf{1}$, allora $w(\alpha) = 2^{n-1}$. Infatti, se $a_0 = 0$, allora il supporto di $\alpha = \sum a_i x_i + a_0 = a \cdot x$ è

$$|\operatorname{Supp}(\alpha)| = |\,\{\, x \in \mathbb{F}^n \mid \alpha(x) \neq 0 \,\}\,| = 2^n - |\,\{\, x \in \mathbb{F}^n \mid a \cdot x = 0 \,\}\,| = 2^n - |\,\langle a \rangle^{\perp}\,| = 2^n - 2^{n-1} = 2^{n-1}$$

Analogamente, se $a_0 = 1$, si ha che il supporto di α è $\{\, x \in \mathbb{F}^n \mid \alpha(x) = 1 \,\} = \{\, x \in \mathbb{F}^n \mid a \cdot x = 0 \,\}$ sommando i termini noti, da cui si procede come nel caso precedente.

Definizione 6.11 (Funzione Bilanciata)**.** Sia $f \in \mathcal{B}_n$ una funzione booleana di peso $w(f) = 2^{n-1}$. Allora f è detta *bilanciata*.

Tale classe di funzioni risulta rilevante poiché hanno metà zeri, metà uni, quindi (scegliendo opportunamente la funzione) permettono di produrre uno stream perfettamente casuale che possiamo integrare all'interno di uno stream di merda casuale e fuorviare le crittanalisi.

Definiamo ora la funzione

$$\delta_a(x) = \begin{cases} 1 & \text{se } x = a, \\ 0 & \text{altrimenti.} \end{cases}$$

Si nota banalmente che $w(\delta_a) = 1$; altrettanto facilmente si può dimostrare che $N(\delta_a(x)) = 1$ poiché δ_a dista 1 da $\mathbf{0}$, e non può essere affine (se per assurdo lo fosse, avrebbe peso 2^{n-1}).

Definizione 6.12 (Trasformata di Fourier)**.** Si dice *trasformata (discreta) di Fourier* di una funzione booleana $f \in \mathcal{B}_n$ in $a \in \mathbb{F}^n$ la mappa:

$$F_f(a) = \sum_{x \in \mathbb{F}^n} (-1)^{x \cdot a} f(x) \in \mathbf{Z}$$

Nota che qui viene fatto un piccolo abuso: l'esponente $x \cdot a$ e $f(x)$ vengono infatti considerati come interi.

Per $f \in \mathcal{B}_n$ definiamo anche l'operatore *segno* di f, $\hat{f}(x) = (-1)^{f(x)} \in \mathbf{Z}$.

Definizione 6.13 (Trasformata di Walsh)**.** La *Ttrasformata di Walsh* è banalmente la trasformata di Fourier sul segno:

$$\mathcal{W}_f(a) = \sum_{x \in \mathbb{F}^n} (-1)^{x \cdot a} \hat{f}(x) = \sum_{x \in \mathbb{F}^n} (-1)^{x \cdot a + f(x)} \in \mathbf{Z}.$$

Osservazione 6.14. Una funzione è bilanciata sse $\mathcal{W}_f(0) = 0$. Questo è banalmente vero, perché

$$\mathcal{W}_f(0) = \sum_x (-1)^{f(x)} = w(f) - w(f+1) = w(f) - w(f) = 0.$$

Teorema 6.15. *La trasformata di Walsh gode delle seguenti proprietà:*

(i)
$$F_f(a) = 2^{n-1}\delta_0(a) - \frac{\mathcal{W}_f(a)}{2};$$

(ii)
$$\mathcal{W}_f(\mathbf{0}) = 2^n - 2w(f);$$

(iii)
$$F_f(\mathbf{0}) = w(f);$$

(iv)
$$\mathcal{W}_f(a) = 2^n - 2w(f + a\cdot x) = w(f + a\cdot x + 1) - w(f + a\cdot x).$$

Dimostrazione. Per (i), possiamo riordinare la disuguaglianza di sopra come $\mathcal{W}_f(a) = 2^n\delta_0(a) - 2F_f(a)$. Da qui procedo per dimostrazione diretta:

$$\begin{aligned}
\mathcal{W}_f(a) + 2F_f(a) &= [2^n\delta_0(a)] \\
&= \sum_{x\in\mathbb{F}^n}(-1)^{x\cdot a + f(x)} + 2\sum(-1)^{x\cdot a}f(x) \\
\boxed{\text{per } a = 0} \qquad &= \sum_{f(x)=0} 1 + \sum_{f(x)=1}(-1) + 2\sum_{f(x)=1} 1 = 2^n = 2^n\delta_0(a) \\
\boxed{\text{per } a \neq 0} \qquad &= 2\left[\sum_{\substack{f(x)=1\\ a\cdot x=0}} 1 + \sum_{\substack{f(x)=1\\ a\cdot x=1}}(-1)\right] + \left[\sum_{\substack{f(x)=0\\ a\cdot x=0}} 1 + \sum_{\substack{f(x)=0\\ a\cdot x=1}}(-1) + \sum_{\substack{f(x)=1\\ a\cdot x=0}}(-1) + \sum_{\substack{f(x)=1\\ a\cdot x=1}} 1\right] \\
&= \sum_{\substack{f(x)=0\\ a\cdot x=0}} 1 + \sum_{\substack{f(x)=0\\ a\cdot x=1}}(-1) + \sum_{\substack{f(x)=1\\ a\cdot x=0}} 1 + \sum_{\substack{f(x)=1\\ a\cdot x=1}}(-1) \\
&= \sum_{a\cdot x=0}(1) + \sum_{a\cdot x=1}(-1) = 0 = \delta_0(a)2^n. \qquad \text{[lemma dei cassetti generalizzato su } x\mapsto a\cdot x\text{]}
\end{aligned}$$

Per (ii), si vede immediatamente applicando la definizione:

$$\begin{aligned}
\mathcal{W}_f(\mathbf{0}) &= \sum_{x\in\mathbb{F}^n}(-1)^{f(x)} \\
&= \sum_{f(x)=1}(-1) + \sum_{f(x)=0} 1 \\
&= -w(f) + (2^n - w(f)) = 2^n - 2w(f).
\end{aligned}$$

La (iii) è immediata usando (i) e (ii):

$$F_f(\mathbf{0}) = 2^{n-1}\delta_0(\mathbf{0}) - \frac{\mathcal{W}_f(\mathbf{0})}{2} = 2^{n-1} - 2^{n-1} + w(f) = w(f).$$

Per la (iv): $\mathcal{W}_f(a) = \sum_{x\in\mathbb{F}^n}(-1)^{a\cdot x + f(x)} = \mathcal{W}_{f+a\cdot x}(0) = 2^n - 2w(f + a\cdot x)$ da (iii). Nota bene che per $a\cdot x$ intendiamo la mappa $x\mapsto a\cdot x$. Per l'ultimo passaggio basta notare che $w(f + a\cdot x + 1) = 2^n - w(f + a\cdot x)$. □

Proposizione 6.16.

$$N(f) = \min\left\{ 2^{n-1} - \frac{\mathcal{W}_f(a)}{2}, 2^{n-1} + \frac{\mathcal{W}_f(a)}{2} \right\}_{a\in\mathbb{F}^n} = 2^{n-1} - \max_{a\in\mathbb{F}^n}\frac{\mathcal{W}_f(a)}{2}$$

Dimostrazione. È immediato una volta che notiamo che $N(f) = d(f, \mathcal{A}_n) = \min\{\, d(f, a\cdot x), d(f, a\cdot x + 1)\,\}$, e consideriamo l'affermazione (iv) del Teorema 6.15. □

Teorema 6.17 (Parseval Relation)**.** *Abbiamo che, per ogni funzione booleana $f\in\mathcal{B}_n$,*

$$\sum_{a\in\mathbb{F}^n}\mathcal{W}_f^2(a) = 2^{2n}$$

Dimostrazione. La dimostrazione è diretta:

$$\begin{aligned}
\sum_{a\in\mathbb{F}^n} \mathcal{W}_f^2(a) &= \sum_a \left[\sum_x (-1)^{a\cdot x+f(x)}\right]\cdot\left[\sum_y (-1)^{a\cdot y+f(y)}\right] && \text{[definizione di trasformata di Walsh]}\\
&= \sum_a \sum_{x,y} (-1)^{f(x)+f(y)+a\cdot(x+y)} && \text{[per linearità del prodotto scalare]}\\
&= \sum_{x,y} (-1)^{f(x)+f(y)} \left[\sum_a (-1)^{a\cdot(x+y)}\right] = \sum_{x,y} (-1)^{f(x)+f(y)} \mathcal{W}_{x+y}(0)\\
&= \sum_{x,y} (-1)^{f(x)+f(y)} (2^n - 2w(x+y))\\
&= \left[\sum_{\substack{x,y\\x=y}} (-1)^{f(x)+f(x)}\cdot 2^n\right] + \left[\sum_{\substack{x,y\\x\neq y}} (-1)^{f(x)+f(y)}\cdot 0\right] && \text{[distinguiamo i casi } x=y \text{ e } x\neq y]\\
&= \sum_{x=y} 2^n = 2^{2n}. && \square
\end{aligned}$$

Proposizione 6.18. *Data una funzione booleana $f \in \mathcal{B}_n$:*

$$\max_{a\in\mathbb{F}^n} |\mathcal{W}_f(a)| \geq 2^{n/2}$$

Dimostrazione. Se considero una somma di termini positivi $\{c_i\}_i$, $\sum_i c_i = C$, è facile notare che $\max_i c_i \geq C/n$. Nel nostro caso, sappiamo che $\sum_a \mathcal{W}_f^2(a) = \sum_a |\mathcal{W}_f(a)|^2 = 2^{2n}$. Pertanto, $\max_a |\mathcal{W}_f(a)|^2 \geq 2^{2n}/2^n = 2^n \implies \max_a |\mathcal{W}_f(a)| \geq 2^{n/2}$. □

Corollario 6.19 (Covering Radius Bound)**.** *Unendo la Proposizione di cui sopra alla Proposizione 6.16 otteniamo il* covering radius bound*:*

$$N(f) = 2^{n-1} - \max_a \frac{|\mathcal{W}_f(a)|}{2} \leq 2^{n-1} - 2^{n/2-1}.$$

Definizione 6.20 (Funzione bent)**.** Se $f \in \mathcal{B}$ è tale che $\mathcal{W}_f(a) = \pm 2^{n/2}\ \forall a$ allora f viene detta *bent*. Equivalmentemente, possiamo dire che f è bent sse $N(f) = 2^{n-1} - 2^{n/2-1}$.

Nota bene che una condizione necessaria per avere funzioni bent è che n sia pari.

Definizione 6.21. Chiamiamo *spettro* di f il multi-set di tutti i valori possibili della trasformata di Walsh. Chiamiamo *spettro esteso* di f il multi-set di tutti i possibili valori, in modulo, della trasformata di Walsh. Chiamiamo *supporto di Walsh* i valori dello spettro non nulli.

Teorema 6.22. *Tutte le funzioni della forma:*

$$f = x_1x_2 + x_3x_4 + \cdots + x_{n-1}x_n.$$

sono bent*.*

Dimostrazione. Per induzione:

- $n = 2$ banale: distinguendo i possibili valori di x $\{(00),(01),(10),(11)\}$ si ottiene

  $$\mathcal{W}_f(a) = 1 + (-1)^{a_2} + (-1)^{a_1} + (-1)^{a_1+a_2+1} = \pm 2.$$

  (Nota: tutti questi condi si fanno velocemente a mano, e si ottiene il meno solo per $a = (11)$);
- $n = 4$ la funzione $f = x_1x_2 + x_3x_4$ ha trasformata di Walsh:

  $$\begin{aligned}
  \mathcal{W}_f((a_1,a_2,a_3,a_4)) &= \left[\sum_{x_1,x_2} (-1)^{a_1x_1+a_2x_2+x_1x_2}\right]\left[\sum_{x_3,x_4} (-1)^{a_3x_3+a_4x_4+x_3x_4}\right]\\
  &= (\pm 2^{n/4})(\pm 2^{n/4}) = \pm 2^{n/2};
  \end{aligned}$$

- si procede come nel caso $n = 4$, notando che la somma degli esponenti è il prodotto delle Walsh. □

Un grosso risultato negativo delle funzioni *bent* è che esse non sono bilanciate:

Teorema 6.23. *Se $f \in \mathcal{B}_n$ è bent, allora non è bilanciata. Tuttavia, $D_a(f)$ è bilanciata $\forall a \in (\mathbb{F}^n)^\times$, dove $D_a(f) = f(x+a) + f(x)$.*

Dimostrazione. Se per assurdo f bent fosse bilanciata, $\mathcal{W}_f(0) = \pm 2^{n/2}$ per definizione di bent, e $\mathcal{W}_f(0) = 0$ per l'Osservazione 6.14. Questo è un assurdo. `TODO: il secondo punto si dimostra sfruttando una proprietà delle funzioni bent, esplicata nella sezione 6 del carlet.` □

Definizione 6.24 (Funzione semi-bent)**.** Una funzione $f \in \mathcal{B}_n$ è detta semi-bent sse

$$\mathcal{W}_f(a) = \begin{cases} 2^{\frac{n+1}{2}} \text{ oppure } 0 & \text{per } n \text{ dispari} \\ 2^{n/2} \text{ oppure } 0 & \text{per } n \text{ pari.} \end{cases}$$

In questo caso, possiamo notare che si guadagna negli attacchi di correlazione poiché per n dispari $N(f) = 2^{n-1} - \max_a \frac{|\mathcal{W}_f(a)|}{2} = 2^{n-1} + 2^{(n-1)/2}$.

Definizione 6.25 (Grado)**.** $f \in \mathcal{B}_n$, ha *grado algebrico* il grado della sua ANF, i.e.:

$$\deg f = \deg \sum_{I \subset \mathcal{P}(n)} a_I x_I$$

Osservazione 6.26. Se $f \in \mathcal{B}_n$ è bilanciata, e $n \geq 2$, allora il suo grado è superiormente limitato da $n-1$.

Dimostrazione. Se per assurdo f fosse bilanciata e avesse grado n, allora sarebbe della forma

$$f = x_1 x_2 \cdots x_n + g \quad \text{con } \deg g \leq n-1.$$

e avrebbe peso:

$$w(f) = w(x_1 \cdots x_n) + w(g) - 2|\{\, x_1 \cdots x_n = 1,\ g(x) = 1 \,\}| = 2^{n-1}.$$

Affinché f abbia peso pari, a questo punto, notiamo che dev'esser $w(g)$ dispari, in quanto $w(x_1 \cdots x_n) = 1$. A meno di un cambio di variabili possiamo suppore che $w(g)$ non contenga una certa variabile x_i; ma allora $\forall x = (x_1 \cdots x_n) \quad g(x) = 1 \implies g(x + (0 \cdots 1 \cdots 0)) = 1$, e quindi il suo peso risulta esser pari. Abbiamo un assurdo. □

6.1 Funzioni Booleane Vettoriali

Definizione 6.27. Una funzione $F : \mathbb{F}^m \to \mathbb{F}^n$ è detta *funzione booleana vettoriale*, o V.B.F., o ancora (m,n)-V.B.F. quando voglio specificare dimensione di dominio e codominio.

Posso definire una ANF anche qui, ed esprimere $F = \sum_{I \in \mathcal{P}(m)} a_I x_I$ dove $a_I \in \mathbb{F}^n$ (mentre nel caso precedente le avevamo in $\mathbb{F}$), e conseguentemente posso esprimere F come vettore di funzioni booleane $(f_1, f_2, \ldots, f_m)$.

Un altro tipo di rappresentazione si può fare quando $n = m$. In tal caso, esiste un polinomio $\mathbb{F}_{2^n} \simeq \mathbb{F}^n$ che lo identifica.

Esempio 6.28. Consideriamo la *patched inversion* $F : \mathbb{F}^2 \to \mathbb{F}^2 : x \mapsto x^{-1}$ e $0 \mapsto 0$ definita nel campo $\mathbb{F}^2 = \{\, 0,\ 1,\ \alpha,\ \alpha^2 = \alpha + 1 \,\}$. La mappa $F = (f_1, f_2)$ si comporta come nalla tabella seguente.

$\mathbb{F}_4$	$\mathbb{F}^2$	x^{-1}
0	(00)	(00)
1	(01)	(01)
α	(10)	(11)
α^2	(11)	(10)

Nel caso in cui volessimo trovare $f_1 = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$ basterebbe notare:

$$\begin{aligned} f(0) &= a_0 = 0, \\ f(1) &= a_2 = 0, \\ f(\alpha) &= a_1 = 1, \\ f(\alpha^2) &= 1 + a_{12} = 1. \end{aligned}$$

Nota: la *patched inversion* può esser riscritta come $x \mapsto x^{q-2}$. Questo risultato è immediato usando Eulero-Fermat.

Definizione 6.29 (Componente)**.** Sia $F : \mathbb{F}^m \to \mathbb{F}^n$ una V.B.F.. Si dice *componente* di F ogni funzione $v \cdot F$ con $v \in (\mathbb{F}^n)^\times$. Usando la notazione vettoriale, possiamo notare che $v \cdot F = \sum_i^n v_i f_i$.

Definizione 6.30 (Grado)**.** Il grado di una V.B.F. $F : \mathbb{F}^m \to \mathbb{F}^n$ è $\deg F = \max_{v \in (\mathbb{F}^n)^\times} \deg(v \cdot F)$.

Definizione 6.31 (Balanced Vectorial Boolean Function)**.** Una funzione F (m,n)-V.B.F. è detta *bilanciata* sse $\forall u_1, u_2 \in \mathbb{F}^m$ si ha che $|F^{-1}(u)| = |F^{-1}(v)| = 2^{m-n}$.

Proposizione 6.32. *Si ha che $F = (f_i)_i$ (m,n)-V.B.F. è bilanciata sse le f_i sono tutte bilanciate.*

Dimostrazione. " $\implies$ ". Si consideri la classe di equivalenza $\sim_i$ sull'immagine di F, per cui un vettore $u \sim v \iff u_i = v_i$. Notiamo che le classi di equivalenza $[u_i]$ di $F/\sim_i$ sono le stesse di F/f_i poiché stiamo quozientando sull'immagine di f_i. Studiamo ora le controimmagini:

$$\left|F^{-1}([u_i])\right| = 2^{m-n} \cdot 2^{n-1} = 2^{m-1} \implies f_i \text{ è bilanciata.}$$

" $\impliedby$ ". Considero il prodotto cartesiano delle immagini. Per ogni vettore $u \in (\mathrm{Im}(f_1) \times \mathrm{Im}(f_2) \times \cdots \times \mathrm{Im}(f_n))$ ho che $f_i^{-1}(u) = 2^{m-1}$. Considero $F = (f_i)_i$, allora

$$\left|F^{-1}(u)\right| = \frac{2^{m-1}}{2^{n-1}} = 2^{m-n}$$

poiché, fissata una f_i, abbiamo 2^{m-1} controimmagini, delle quali la metà sono zero, la metà uno, per $f_j \neq f_i$ poiché f_j bilanciata (`TODO: questa parte è imprecisa. Il carlet nella sezione 2.3 compone` f `con` δ_b `e ne tira fuori una dimostrazione. Comunque non è stata vista a lezione.`). □

Proposizione 6.33. *Nel caso $n = m$ abbiamo che $F : \mathbb{F}^n \to \mathbb{F}^n$ è una permutazione sse è bilanciata.*

Dimostrazione. Poiché $n = m$, usando la definizione di *bilanciata*, abbiamo che $\left|F^{-1}(u)\right| = 2^0 = 1$, quindi è bigettiva per il lemma dei cassetti, quindi una permutazione. L'altro verso è banale: se è una permutazione, la controimagine è unica, quindi è bilanciata. □

Sia Sym il grupppo simmetrico, ricordiamo che $\mathrm{Sym}(\mathbb{F}^2) = \mathrm{AGL}(\mathbb{F}^2)$, ossia tutte le trasformazioni lineari.

Infatti, presa una permutazione, essa è bilanciata, quindi ha grado non massimo.

Sfruttando questo fatto, nell'Esempio 6.28, avremmo potuto risparmiarci la ricerca del coefficiente di x_1x_2 sapendo che F è una permutazione, quindi i singoli f_i sono bilanciati, quindi lo è F, quindi non può avere grado massimo.

Definizione 6.34 (Trasformata di Walsh)**.** La *trasformata di Walsh* per F (m,n)-V.B.F. è

$$\mathcal{W}_F(u,v) = \sum_{x \in \mathbb{F}^n} (-1)^{u \cdot x + v \cdot F(x)} = \mathcal{W}_{v \cdot F}(u).$$

Definizione 6.35. In modo analogo al grado possiamo definire la nonlinearità $N(f) = \min_{v \neq 0} N(vF)$.

Osservazione 6.36. $N(F) \leq N(f_i) \quad \forall i$.

Corollario 6.37 (Covering Radius Bound)**.**

$$N(F) \leq 2^{m-1} - 2^{m/2-1} \tag{6.1}$$

Dimostrazione. Immadiato per l'osservazione precedente, una volta che applichiamo la trasformata di Walsh:

$$N(F) = 2^{m-1} - \frac{1}{2} \max_{u,v \neq 0} |\mathcal{W}_F(u,v)| \leq 2^{m-1} - \frac{1}{2} 2^{m/2}.$$

□

Definizione 6.38. Una V.B.F. $F : \mathbb{F}^m \to \mathbb{F}^n$ si dice *bent* se vale l'uguaglianza nella 6.1.

Come nella sezione precedente quindi, m deve esser pari anche in questo caso.

Teorema 6.39. *Se m è pari e $n \leq m/2$ allora esiste una funzione F (m,n)-V.B.F. bent.*

Teorema 6.40 (Vectorial Parseval Relation)**.** *Abbiamo che, per ogni* (m,n)-V.B.F. F*,*

$$\sum_{\substack{u,v\\v\neq 0}} \mathcal{W}_f^2(u,v) = (2^n-1)(2^{2m}).$$

Dimostrazione. Si vede facilmente:

$$\sum_{\substack{u,v\\v\neq 0}} \mathcal{W}_F^2(u,v) = \sum_{v\neq 0} \underbrace{\left(\sum_u \mathcal{W}_{v\cdot F}^2(u)\right)}_{2^{2m} \text{ per Parseval}} = (2^{m-1})2^{2m}.$$

□

Teorema 6.41 (Sidelnikov-Chabaud-Vandermay Bound)**.** *Siano* $m,n>0$ *con* $n \geq m-1$ *e* F *una* (m,n)-V.B.F.*. Allora:*

$$N(F) \leq 2^{m-1} - \frac{1}{2}\sqrt{3\cdot 2^m - 2 - 2\frac{(2^m-1)(2^{m-1}-1)}{2^n-1}}.$$

Dimostrazione. Sappiamo già che

$$N(f) = 2^{m-1} - \frac{1}{2}\max_{u,v\neq 0}\{\ |\mathcal{W}_F(u,v)|\ \}_{u,v}.$$

Volendo trovare un maggiorante di questa quantità, possiamo sfruttare un fatto più generale:

$$\bar{a} = \max_i a_i \geq \frac{\sum_i a_i^2}{\sum_i a_i}.$$

Infatti, se fosse vero l'opposto si avrebbe che $a_i\bar{a} < a_i^2$, ma questo è un assurdo poiché $\bar{a}$ è il massimo valore. Abbiamo quindi che

$$\max_{\substack{v\neq 0\\u}} [\mathcal{W}_F(u,v)]^2 \geq \frac{\sum_{\substack{v\neq 0\\u}} (\mathcal{W}_F(u,v))^4}{\sum_{\substack{v\neq 0\\u}} (\mathcal{W}_F(u,v))^2}. \tag{6.2}$$

Il denominatore ci è noto per Parseval, ed esso è uguale a $(2^n-1)2^{2m}$. Consdero ora il numeratore:

$$\begin{aligned}
&\sum_{u,v}\left[\sum_x (-1)^{v\cdot F(x)+u\cdot x}\right]^4\\
=&\sum_{u,v}\left[\left(\sum_x (-1)^{v\cdot F(x)+u\cdot x}\right)\left(\sum_y (-1)^{v\cdot F(y)+u\cdot y}\right)\left(\sum_z (-1)^{v\cdot F(z)+u\cdot z}\right)\left(\sum_t (-1)^{v\cdot F(t)+u\cdot t}\right)\right]\\
=&\sum_{u,v}\sum_{x,y,z,t} (-1)^{v\cdot(F(x)+F(y)+F(z)+F(t))+u\cdot(x+y+z+t)} && \text{[proprietà della sommatoria]}\\
=&\sum_{x,y,z,t}\left[\sum_{v\neq 0}(-1)^{v\cdot(F(x)+F(y)+F(z)+F(t))}\right]\left(\sum_u (-1)^{u\cdot(x+y+z+t)}\right)\\
=&2^{n+m}\cdot|\{\,(x,y,z,t)\mid x+y+z+t=0\ \wedge\ F(x)+F(y)+F(z)+F(t)=0\,\}|
\end{aligned}$$

infatti se incontriamo una quadrupla che non soddisfa i criteri di quelli'insieme, il prodotto si annulla (questo lo si può notare considerando anche solo che la funzione $x+y+z+t$ è bilanciata). Il problema quindi si riduce a cercare le quadruple tali che $x+y+z=t \implies F(x)+F(y)+F(z)=F(x+y+z)$. Nota che questo è sempre vero se F è lineare. Studiamo il primo fattore, quello tra parentesi quadre. Definiamo

$$S = \{\,(x,y,z)\in\mathbb{F}^3 \mid F(x)+F(y)+F(z)=F(x+y+z)\,\}.$$

La sua cardinalità può esser minorata contando i casi ovvi in cui due delle tre incognite sono uguali, i.e.

$$|S| \geq |\{\,(x,y,z)\mid x=y\ \vee\ x=z\ \vee\ y=z\,\}|. \tag{6.3}$$

In tale insieme distinguiamo i casi: (x,x,z) (x,y,x) (x,y,y) (x,x,x), e i primi tre sono in stesso numero. Quindi la sua cardinalità è $3\cdot|\{\,(x,x,y)\,\}| - 2\cdot|\{\,(x,x,x)\,\}| = 3\cdot 2^{2m} - 2\cdot 2^m$.

Studiamo ora il secondo fattore, quello tra parentesi tonde: $\sum_u \mathcal{W}_F^4(u,0) = \sum_u [\sum_x (-1)^{u\cdot x}]^4$. Tale sommatoria è 0 se $u \neq 0$ poiché stiamo sommando su tutti gli elementi del campo. Se $u = 0$ invece abbiamo 1, quindi la sommatoria è uguale a 2^{4m}. Inglobando tutte questa stime nell'equazione iniziale otteniamo che:

$$\max_{\substack{u \\ v\neq 0}} \mathcal{W}_F^2(u,v) \geq \frac{2^{n+m}[3\cdot 2^{2m} - 2\cdot 2^m] - 2^{4m}}{(2^n-1)2^{2m}} = \frac{2^n[3\cdot 2^m - 2] - 2^{2m}}{(2^n-1)}.$$

Se si confronta con l'espressione sotto radice, si noterà che i due sono uguali. □

Definizione 6.42 (AB v.B.f.)**.** Data una funzione F (n,n)-v.B.f., essa è detta *almost bent* (AB) se vale l'uguale nel bound CSV, ossia

$$N(F) = 2^{n-1} - 2^{(n-1)/2}$$

Si può notare che per una funzione vettoriale booleana, si ha il massimo posssibile per n dispari, contrariamente al caso in cui n sarebbe dovuto esser pari per una funzione booleana. Inoltre, questo ha un significato morale più profondo: quando abbiamo una funzione lineare, tener conto della linearità è molto semplice; quando abbiamo a che fare invece con un vettore di funzioni booleane riceventi lo stesso input, è più difficile cercare la non linearità.

Definizione 6.43 (Derivata)**.** Come nel caso scalare, definiamo la *derivata* $D_a(F) = F(x) + F(x+a)$ per un qualche $a \neq 0$.

Proposizione 6.44. *Una $F = (f_i)_i$ (m,n)-v.B.f. è* bent *se e soltanto se tutte le $D_a(f_i)$ sono bilanciate.*

Dimostrazione. Si ricordano i seguenti fatti:

(i) $f \in \mathcal{B}_n$ è bent $\iff$ $D_a(f)$ è bilanciata, per ogni $a \neq 0$;

(ii) $F = (f_i)_i$ è bilanciata $\iff$ f_i è bilanciata, per ogni i;

(iii) F è bent $\iff$ vF è bent, per ogni $v \neq 0$;

Il verso " $\implies$ " si giustifica dicendo che se F è bent, allora vF lo è per (iii), allora lo sono tutte le sue componenti, allora $D_a(f_i)$ è bilanciata, per ogni i ed ogni $a \neq 0$.
Il verso " $\impliedby$ "`TODO: ??` □

Definizione 6.45 (δ-uniforme differenziabilità)**.** Sia F una (m,n)-v.B.f., $\delta > 0$ un intero positivo. F è detta *δ-uniformemente differenziabile* se $\forall a \in \mathbb{F}^m \,.\, a \neq 0,\ \forall b \in \mathbb{F}^m$

$$\Delta_{a,b}(F) = |D_a(F)^{-1}(b)| = |\,\{\, x \in \mathbb{F}^n \mid D_a(F)(x) = F(x) + F(x+a) = b \,\}\,| \leq \delta$$

Possiamo notare immediatamente che non esistono funzioni 1-uniformemente differenziabili, poiché se $\exists x \,.\, D_a(F)(x) = b \implies D_a(F)(x+a) = b$

Definizione 6.46 (APN)**.** Una v.B.f. F è detta *APN (Almost Perfectly Nonlinear)* se è differentiably 2-uniform.

Osservazione 6.47. Se la v.B.f. F è lineare, allora

$$\left|D_a(F)^{-1}(b)\right| = \begin{cases} 2^n & \text{se } b = F(a) \\ 0 & \text{altrimenti.} \end{cases}$$

Quindi una funzione lineare ragiunge il massimo valore di differenziabilità δ-uniforme, e una funzione è tanto più nonlineare quanto più vi è distante.

Osservazione 6.48. Ogni funzione Almost-Bent è APN. `TODO: credo questo si dimostri usando l'SCV bound, ma comunque non è stato visto a lezione.`

Ci domandiamo se, tra le (n,n)-v.B.f., le funzioni APN esistano sempre. Abbiamo il seguente risultato:

Teorema 6.49. *La patched inversion è una funzione booleana APN se $n \geq 3$ è dispari.*

Dimostrazione. Posto $f : x \mapsto x^{-1}$ si ha che

$$\hat{f}_a = f(x+a) + f(x) = \frac{1}{x+a} + \frac{1}{x} = \frac{a}{x^2+ax} = b.$$

Possiamo notare che se $x \neq 0$ (poiché anche $a \neq 0$ per definizione di derivata), allora possiamo ridurci alla risoluzione di $x^2 + ax + a/b = 0$, equzione di secondo grado che dunque ha al più due suluzioni. Rimane poi verificare i casi $x = 0$ e $x = a$ per cui:

$$\hat{f}_a(a) = \hat{f}_a(0) = f(a) = b \implies b = a^{-1},$$

che nel caso dispari viene un valore diverso da tutti gli altri, nel caso pari si ripete uno dei precedenti, quindi la funzione non è APN. □

Vi sono anche degli altri risultati nel caso n pari, ma meno importanti.

- per $n \geq 3$ dispari, possiamo sempre definite la patched inversion;
- per $n = 4$, è stato dimostrato che non esistono permutazioni APN;
- per $n = 6$, è stata trovata un'unica permutazione a meno di una relazione di equivalenza che introdurremo in seguito, detta CCZ.;
- $n \geq 8$ pari, è ancora un problema aperto.

È stata inoltre data una definizione più debole di δ-differenziabilità.

Definizione 6.50. Una (n, m)-V.B.F. è detta *debolmente δ-uniformemente differenziabile* se $\forall a \neq 0$

$$|\mathrm{Im}(D_a(F))| > \frac{2^{n-1}}{\delta}.$$

Proposizione 6.51. *Se $F : \mathbb{F}^m \to \mathbb{F}^n$ è δ-uniformemente differenziabile, allora è debolmente δ-uniformemente differenziabile.*

Dimostrazione. Abbiamo che $a \neq 0$. Considero:

$$\sum_{b \in \mathrm{Im}(D_a(F))} \left|D_a(F)^{-1}(b)\right| = 2^n.$$

Infatti, tutte le quantità dentro la sommatori sono disgiunte (per definizione di funzione) e maggiorate da δ. Segue che:

$$\begin{aligned} 2^n = \sum |D_a(F)^{-1}(b)| \leq \sum \delta = \delta \cdot |\mathrm{Im}(D_a(F))| \\ \implies |\mathrm{Im}\, D_a(F)| > \frac{2^n}{\delta} > \frac{2^{n-1}}{\delta}. \end{aligned}$$

□

Definizione 6.52. Se F è debolmente 2-uniformemente differenziabile, allora è detta *debolmente APN*.

Abbiamo che $\forall n$ esiste una permutazione $F : \mathbb{F}^n \to \mathbb{F}^n$ debolmente APN. Essa è la patched inversion:

Proposizione 6.53. *La* patched inversion *è una permutazione debolmente APN.*

Dimostrazione. Abbiamo $F : \mathbb{F}^n \to \mathbb{F}^n : z \mapsto z^{-1}$ e $0 \mapsto 0$. Allora, dato $a \neq 0$, $D_a(F) = z^{-1} + (z + a)^{-1} = b \iff b = (z + a + z)/(z \cdot (a + z)))$ (sommando i termini come fossero frazioni, restringendoci al caso $z \neq 0, a$) $\iff b = a^{-1}(z^2 + az)$ che è un'equazione di secondo grado, la quale ammette al più due soluzioni. Quindi $|\{\, x \in \mathbb{F}_2^m \mid x \neq 0,\ x \neq a,\ D_a(x) = b \,\}|$ ha cardinalità al più 2. D'altro canto, non può esser meno di due. □

Definizione 6.54. Una mappa $F : \mathbb{F}^n \to \mathbb{F}^n$ con $F(0) = 0$ è detta *l-anti-invariant* se per ogni sottospazio U di $\mathbb{F}^n$, si ha che $F(U) = U$ (U è F-invariante) $\implies \dim U < n - l$.

Diremo che la stessa mappa è *fortemente l-anti-invariant* se $\forall U, V < \mathbb{F}^n$ abbiamo che $F(U) = W \implies \dim U = \dim V < n - l$.

`TODO: a questo punto negli appunti leggo che possiamo chedere al massimo` $l = n - 2$`, poiché per` $l = n - 1$ `succede qualcosa di paradossale, ma non riesco a capire cosa.`

Lemma 6.55. *Sia $E \leq \mathbb{F}^n$ un sottospazio su cui $l \in \mathcal{B}_n$ risulta lineare. Allora $\sum_{x \in E} (-1)^{l(x)} = 0$.*

Dimostrazione. La dimostrazione è facile: sia m la dimensione di E, possiamo scrivere x come somma di $\lambda_i \cdot e_i$ dove $\lambda_i \in \mathbb{F}$ e gli e_i sono una base per $\mathbb{F}^m$. Per linearità, $l(x) = \lambda_1 l(e_i) + \cdots + \lambda_m l(e_m)$. Ora, se $l(e_i) = 0$, allora possiamo considerare la somma senza quel termine. Allora

$$\sum_{x \in E} (-1)^{\lambda_1} \cdots (-1)^{\lambda_k} = \sum_{x \in E} (-1)^{\lambda_1 + \cdots + \lambda_k}$$

la cui somma dei termini è per metà volte pari, per metà dispari. □

Lemma 6.56. *Sia $\mathbf{1}_E$ la funzione caratteristica, i.e. tale che x viene mandato in 1 s e $x \in E$, 0 altrimenti. Allora $F_{\mathbf{1}_E}(a) = |E| \cdot \mathbf{1}_{E^\perp}(a)$.*

Dimostrazione. Dalla definizione di trasformata di Fourier, possiamo notare:

$$F_{\mathbf{1}_E}(a) = \sum_{x \in \mathbb{F}^n} \mathbf{1}_E(x)(-1)^{a \cdot x} = \sum_{x \in E} (-1)^{a \cdot x}.$$

Ora, se $a \in E^\perp \implies a \cdot x = 0 \ \forall x \in E$; se invece $a \notin E^\perp \implies a \cdot$x è lineare su E. Pertanto, se $a \in E^\perp$, allora $F_{\mathbf{1}_E} = |E|$, altrimenti la somma è zero per il teorema precedente. □

Lemma 6.57 (Formula di Inversione di Poisson)**.** *Dato un sottospazio $E \leq \mathbb{F}^n$ e una funzione booleana $\varphi \in \mathcal{B}_n$, $\sum_{u \in E} F_\varphi(u) = |E| \sum_{x \in E^\perp} \varphi(x)$.*

Dimostrazione. Per dimostrazione diretta:

$$\sum_{u \in E} \left[\sum_{x \in \mathbb{F}^n} \varphi(x)(-1)^{u \cdot x} \right] = \sum_{x \in \mathbb{F}^n} \varphi(x) \sum_{u \in E} (-1)^{u \cdot x} = |E| \sum_{x \in E^\perp} \varphi(x).$$ □

Lemma 6.58. *Sia $f \in \mathcal{B}_n$, e sia k un intero tale che $2^k \mid \mathcal{W}_f(a) \ \forall a$. Allora* $\deg f \leq n - k + 1$.

Dimostrazione. Sia $d = \deg f$, e supponiamo per assurdo che $d > n - k + 1 \implies k > n - d + 1$. Allora $f = x_{i_1} x_{i_2} \cdots x_{i_d} + g$. Consideriamo

$$E = \{\, x \in \mathbb{F}^n \mid v_j = 0 \ \forall i = 1..d \,\} \qquad \text{[È uno spazio vettoriale]}$$
$$E^\perp = \{\, x \in \mathbb{F}^n \mid v_j = 0 \ \forall i \neq 1..d \,\} \simeq \mathbb{F}^d.$$

Pertanto, $f_{E^\perp} = x_{i_1} \cdots x_{i_d} + g'$, dove $\deg g' < d$ (nota bene: prima non potevo dirlo).

Noto che il peso di $f_{E^\perp}$ è dispari: infatti, il monomione ha peso $2^{d-|I|} = 2^{d-d} = 1$ e g ha peso pari, in quanto somma di monomi cha hanno peso pari (il loro grado è minore di d).

Considerando la funzione segno $\hat{f}$ di f, e usando la formula di inversione di Poisson,

$$\sum_{u \in E} \mathcal{W}_f(u) = \sum_{u \in E} F_{\hat{f}}(u) = |E| \sum_{u \in E^\perp} \hat{f}(x) = |E|(2^d - 2w(f_{E^\perp})).$$

Ricordando che il peso di $f_{E^\perp}$ è dispari, abbiamo che $4 \nmid 2w(f_{E^\perp})$, e moltiplicando da entrambe le parti per $|E| = 2^{n-2}$ ottengo

$$2^{n-d+2} \nmid \sum_{u \in E} \mathcal{W}_f(u)$$

ma per ipotesi avevamo supposto che con $k \geq 2$, $2^{n-d+k} \mid \sum_{u \in E} \mathcal{W}_f(u)$. □

Il lemma di sopra è importante perché afferma che la nonlinearità della funzione è tantopiù grande quanto più k è grande, e d'altra parte, quando il grado è molto basso.

Teorema 6.59. *Se una funzione booleana $f \in \mathcal{B}_n$ è bent $\implies$* $\deg f \leq n/2 + 1$. [1]

Dimostrazione. Poiché f è bent, allora $\mathcal{W}_f(a) = \pm 2^{n/2} \ \forall a$. Allora mi basta scegliere $k = n/2$, e usando il lemma precedente ottengo:

$$\deg f = n - k + 1 = n - n/2 + 1 = n/2 + 1.$$ □

[1] In realtà è vero anche $\deg f \leq n/2$, ma trattiamo codesto caso per la facilità della dimostrazione.

Definizione 6.60 (Affine Equivalenza)**.** Due funzioni $f, g \in \mathcal{B}_n$ sono dette *affinemente equivalenti* se sono uguali a meno di un cambio di variabili, i.e. $\exists \alpha \in AGL(\mathbb{F}^n)\ .\ f \circ \alpha(x) = f(Ax+b) = g(x)$. Similmente, due funzioni vettoriali booleane sono dete *affinemente equivalenti* se $\exists \alpha, \beta \in \mathrm{AGL}(\mathbb{F}^n)$ tali che $\beta \circ F \circ \alpha(x) = G(x)$. Tutte le proprietà invarianti rispetto a questa relazione sono dette *invarianti affini.*

Proposizione 6.61. *Il grado di una funzione booleana (vettoriale o meno) è un'invariante affine. La nonlinearità di una funzione booleana (vettoriale o meno) è un'invariante affine.*

Dimostrazione. Dimostramo entrambi i casi solo per le funzioni booleane.

Persuadere che il grado di una funzione booleana è un'invariante affine è facile:

$$f(x) = g(Ax+b) = \sum_J a_J(Ax^J + b) = \sum_J = (a_J A)x^J + b\sum_J a_J.$$

Per mostrare che la nonlinearità è una invariante affine, considero $f, g \in \mathcal{B}_n\ .\ f(x) = g(Ax+b)$, come prima. Vogliamo mostrare che lo spettro di f e g è lo stesso.

$$\begin{aligned}
\mathcal{W}_f(\bar{a}) &= \sum_{x\in\mathbb{F}^n} (-1)^{f(x)+\bar{a}\cdot x} \\
\mathcal{W}_g(a) &= \sum_{x\in\mathbb{F}^n} (-1)^{f(Ax+b)+a\cdot x} && [\text{sostutuisco } y = Ax+b \implies x = A^{-1}(y+b)] \\
&= \sum_y (-1)^{f(y)+\bar{a}\cdot y} = \mathcal{W}_f(\bar{a}).
\end{aligned}$$

□

Definizione 6.62. Sia F una (m,n)-v.B.f., allora

$$\hat{n}(F) = \max_{\substack{a\neq 0\\ a\in\mathbb{F}^m}} \left|\{\, v \in (\mathbb{F}^n)^\times \mid v \cdot D_a(F) \text{ è costante}\,\}\right|.$$

Teorema 6.63. *Sia F una (n,n)-v.B.f. debolmente APN. Allora $\hat{n}(F) \leq 1$.*

Dimostrazione. Supponiamo per assurdo che $\hat{n}(F) \geq 2$, ovvero $\exists a \neq 0$, $v_1 \neq 0$, $v_2 \neq 0$ tali che $v_1 \cdot D_a(F)$ e $v_2 \cdot D_a(F)$ sono costanti. Senza perder di generalità, posso supporre $v_1 = (1,0,\dots,0)$ e $v_2 = (0,1,0,\dots,0)$ vettori della base canonica - tutt'al più, posso trovare un'affinità che li lega, e l'essere debolmente APN è un'invariante affine. Allora, $D_a(F) = (f_1, f_2, \dots, f_n)$, dove $v_1 \cdot D_a F = f_1 = c_1$ è una funzione costante, ed allo stesso modo f_2 è costante. Allora $\mathrm{Im}(D_a(F)) \subset \{\,(c_1, c_2, \alpha_3, \dots, \alpha_n) \mid \alpha_i \in \mathbb{F}\,\}$, quindi ha cardinalità $|\mathrm{Im}(D_a(F))| \leq 2^{n-2}$. Tuttavia, essendo debolmente APN, $|\mathrm{Im}(D_a(F))| > 2^{n-2}$. Abbiamo un assurdo. □

Corollario 6.64. *F v.B.f. $APN \implies \hat{n}(F) \leq 1$.*

Teorema 6.65. *Sia $F : \mathbb{F}^4 \to \mathbb{F}^4$ una v.B.f.. Allora, se $\hat{n}(F) = 0$, F è debolmente APN.*

Dimostrazione. Sia $F^4 = \{\, x_1, x_2, \dots, x_{16}\,\}$, e $a \neq 0 \in \mathbb{F}^4$. Considero la forma vettoriale della derivata $D_a(F) = (f_1, \cdots, f_4)$. Considero la matrice $M = (m_{ij})_{ij}$:

$$M = \begin{pmatrix} f_1(x_1) & \cdots & f_1(x_{16}) \\ \vdots & \ddots & \vdots \\ f_4(x_1) & \cdots & f_4(x_{16}) \end{pmatrix} \qquad \text{dove quindi } m_{ij} = f_i(x_j)$$

e noto che $|\mathrm{Im}(D_a(F))| > 2^{n-2} = 4 \iff \exists$ almeno 5 colonne distinte in M.

Questo si dimostra per assurdo. Se M contiene $n \leq 4$ colonne distinte, allora sia M' la matrice composta da queste n colonne. Possiamo distinguere i casi

- se $\mathrm{rk}(M') = 4$, allora esistono $\alpha_1, \dots, \alpha_4$ non tutti nulli tali che $(1,1,1,1) = \alpha_1 M'_1 + \cdots + \alpha_4 M'_4$, dove M'_i è la i-esima riga di M (questo per definizione di rango). Ma allora,

$$\sum_{i=1}^4 \alpha_i M_i = (1,1,1,1) = (\alpha_1 \cdots \alpha_4) \cdot D_a F \implies \hat{n}(F) \geq 1,$$

  `TODO: perché?` ma questo è un assurdo poiché avevamo suppposto che $\hat{n}(F) = 0$.

- se $\mathrm{rk}(M') < 4$, consideriamo ancora una volta M' con il massimo numero di colonne distinte e linearmente indipendenti. Allora $\exists \alpha_1, \cdots, \alpha_4$ non tutti nulli, tali che $\sum \alpha_i M'_i = (0 \cdots 0)$. Come prima, questo implica che $(\alpha_1, \cdots, \alpha_4) D_a(F) = (0,0,0,0) \implies \hat{n}(F) \geq 1$ che è un assurdo. □

Esercizio 6.66. Si può dimostrare che $F : \mathbb{F}^3 \to \mathbb{F}^3$ V.B.F. tale che $\hat{n}(F) = 0$ implica che F è debolmente APN. `TODO: la dimostrazione di questo fatto dovrebbe essere analoga al teorema precedente.`

Osservazione 6.67. Per $n \geq 4$ non si sa nulla, fatta eccezione per il caso in cui $n = 2m$, con m dispari. Per l'appunto, in generale:

$$\hat{n}(F) = 0 \text{ non implica } F \text{ debolmente APN.}$$

Vi è infatti un controesempio in merito.

Considero $t \geq 0$. $\gcd(d, 2^n - 1) = 1$, dove $d = 2^{2t+1} - 2^{2t} + 1$, e $F(x) = x^d$. Tale funzione è debolmente 2-uniformemente differenziabile, poiché $\gcd(2^t, n) = 2$, quindi ogni derivata è una funzione 4 in 1 `TODO: perché?` e inoltre $\hat{n}(F) = 0$.

Proposizione 6.68. *Sia $F : \mathbb{F}^n \to \mathbb{F}^n$ una permutazione. Allora $N(F) = N(F^{-1})$.*

Dimostrazione. Sapendo che la non-linearità dipende solo dal massimo delle trasformate di Walsh:

$$\mathcal{W}_F(u,v) = \sum_x (-1)^{vF(x)+ux} = \sum_y (-1)^{vy+uF^{-1}(y)} = \mathcal{W}_{F^{-1}}(v,u). \qquad \square$$

Proposizione 6.69. *Sia $F : \mathbb{F}^4 \to \mathbb{F}^4$ una permutazione tale che $F(0) = 0$. (Nota: quest'ultima condizione non è una restrizione, visto che possiamo comunque effettuare una traslazione). Se F è 4-differenziabile e 2-fortemente anti-invariante, allora F è debolmente APN.*

Dimostrazione. Ricordiamo che F è debolmente APN se per ogni $a \neq 0$, $|\operatorname{Im} D_a F| > 2^{n-1}/2 = 4$. Per assurdo, ipotizziamo invece che $\exists a \neq 0$. $|\operatorname{Im} D_a F| \leq 4$. Allora, si avrebbe che $\forall b \in \operatorname{Im}(D_a F) \quad |D_a F^{-1}(b)| = 4$. Infatti, per ogni numero più piccolo otterrei alla fine meno di 16 elementi, poiché $\bigcup_b D_a F^{-1}(b) = \mathbb{F}^4$.

Consideriamo le preimmagini di $F(a)$ secondo $D_a F$. Abbiamo che

- $D_a F(0) = F(0) + F(0+a) = F(a)$;
- $D_a F(a) = F(a) + F(a+a) = F(a)$;
- poiché devono esserci esattamente 4 preimmagini, deve esistere anche un $x \neq 0,\ x \neq a$. $D_a F(x) = F(a)$;
- $D_a F(x+a) = D_a F(x) = F(a)$.

Consideriamo a questo punto lo spazio vettoriale $U = \{\, 0, a, x, x+a \,\}$. Notiamo che esso ha dimensione $\dim U = 2$, e la sua immagine secondo F ha ancora dimensione due. Tuttavia, per esser 2-fortemente anti-invariante, dovrebbe esser che $\dim U < n - 2 = 2$. Abbiamo un assurdo. □

Definizione 6.70. Con la notazione

$$n_i(F) = |\, \{\, v \mid \deg vF = i \,\} \,|$$

indicheremo il numero di componenti di grado i.

Proposizione 6.71. *Sia $F : \mathbb{F}^4 \to \mathbb{F}^4$ una permutazione. Se F è debolmente APN, allora il suo grado è 3, e $n_3(F) \in \{\, 14, 15 \,\}$.*

Dimostrazione. `TODO: dimostrare` □

Definizione 6.72 (Plateaud)**.** Una funzione booleana ad n registri f è detta *plateaud* se esiste un λ per cui $\mathcal{W}_f(a) \in \{\, 0, \pm\lambda \,\}$, per ogni $a \in \mathbb{F}^n$; λ è detto *coefficiente di plateaud*.

Ricordiamo che una funzione è almost bent se vale l'uguaglianza nel SCV bound, quindi vale l'uguaglianza nelle equazioni 6.2 e 6.3. Formulando meglio:

Proposizione 6.73. *Una funzione vettoriale booleana è almost bent sse vale l'uguaglianza nelle equazioni 6.2 e 6.3.*

Escursione in Teoria dei Codici. Sia $F : \mathbb{F}^n \to \mathbb{F}^n$ una funzione booleana vettoriale tale che $F(0) = 0$, e sia α un elemento primitivo di $\mathbb{F}^n$. Considero la matrice:

$$H = \begin{bmatrix} 1 & \alpha & \cdots & \alpha^{2n-2} \\ F(1) & F(\alpha) & \cdots & F(\alpha^{2n-2}) \end{bmatrix} \quad \text{matrice di dimensione } 2 \times (2^n - 1)$$

Sia $\mathcal{C}_F$ il codice associato ad H, matrice del controllo parità. Allora, è possibile dimostrare che $3 \leq d(\mathcal{C}_F) < 6$, la funzione F è APN sse $d(\mathcal{C}_F) = 5$, e F è almost-bent sse la distibuzione dei pesi del codice ortogonqale $\mathcal{C}_F^{\perp}$ è $\{\, 0, 2^{-1} \pm 2^{(n+1)/2} - 1, 2^{n-1} \,\}$. `TODO: dimostrare questi fatti.`

Definizione 6.74 (Extended Affine Equivalence)**.** Due funzione vettoriali booleane $F, G \in \mathbb{F}^n \to \mathbb{F}^n$ sono dette EA (Extended Affine equivalent), in simboli

$$F \sim_{EA} G \iff \exists A, B, C \in \mathrm{AGL}(\mathbb{F}^n) \;.\; G(x) = AF(B(x)) + C(x).$$

Osservazione 6.75. Notiamo immediatamente che affine equivalenza implica affine equivalenza estesa (basta infatti porre $C = \mathbf{0}$).

La proprietà di uniforme δ-differenziabilità è EA-invariante, poiché, date le F, G come da definizione di sopra, e tenendo conto del fatto che tale proprietà è affinemente invariante

$$D_a(G)(x) = F(x) + C(x) + F(x + a) + C(x) + C(a) = D_a(f) + C(a)$$

dove $C(a)$ è banalmente costante. Lo stesso ragionamente si può fare per la debole δ-uniforme differenziabilità.

La nonlinearità è un'altra invariante affine. Anche in questo caso, date infatti F, G come sopra, e tenuto conto che la nonlinearità è una affine invarianza, abbiamo che

$$nl(G) = d(G, \mathcal{A}_n) = d(F + C, \mathcal{A}_n) = d(F, \mathcal{A}_n + C) = nl(F).$$

Definizione 6.76. Due funzioni $F, G \in \mathbb{F}^n \to \mathbb{F}^n$ osno dette CCZ-eqivalenti se le mappe, intese come coppia, sono invarianti a meno di una trasformazione affine in $\mathrm{AGL}(\mathbb{F}^{2n})$, o meglio, se esiste un $\Lambda \in \mathrm{AGL}(\mathbb{F}^{2n})$ tale che

$$\{\, (x, F(x)) \mid x \in \mathbb{F}^n \,\} = \{\, \Lambda(x, G(x)) \mid x \in \mathbb{F}^n \,\}.$$

Osservazione 6.77. La lineare equivalenza implica CCZ-equivalenza.

Dimostrazione. Date F, G funzioni vettoriali booleane tali che $G(x) = AF(Bx + b) + a$, e dato

$$\Lambda = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{pmatrix} x \\ F(x) \end{pmatrix} + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{6.4}$$

vogliamo fare in modo che $(A_1x + A_2F(x) + a_1, A_3x + A_4F(x) + a_2)$ sia uguale a $(y, G(y))$. Ma allora basta porre

$$\Lambda = \begin{bmatrix} B^{-1} & 0 \\ 0 & A \end{bmatrix} + \begin{bmatrix} B^{-1}b \\ a \end{bmatrix}$$

poiché

$$\Lambda \begin{bmatrix} Bx + b \\ F(Bx + b) \end{bmatrix} = \begin{bmatrix} x \\ AF(Bx + b) + a \end{bmatrix} = \begin{bmatrix} x \\ G(x) \end{bmatrix}. \qquad \square$$

Insomma $\mathrm{AGL}_{\sim} \subset EA_{\sim} \subset CCZ_{\sim}$

Proposizione 6.78. *Data una* (n, n)-V.B.F. *F, essa è CCZ-equivalente alla sua inversa F^{-1}.*

Dimostrazione. Basta porre $\Lambda = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}$. $\square$

Proposizione 6.79. *Tutte le proposizioni che agiscono sullo spettro di una (n, n) funzione vettoriale booleana F sono CCZ-invarianti.*

Dimostrazione. Per mostrare questo fatto, ci basta mostrare che tutte le funzioni $F \sim_{CCZ} G$ sono tali che $\{ \mid \mathcal{W}_F(u,v)| \} = \{ \mid \mathcal{W}_G(u,v)| \}$. Si potrebbe scrivere banalmente che per linearità posso ri-raggruppare le $F(x)$ e gli x, ma per esser prolissi consideriamo Λ come nell'Equazione 6.4, allora

$$\begin{aligned}\mathcal{W}_G(u,v) &= \sum_x (-1)^{vG(x)+ux} \\ &= \sum_x (-1)^{vA_1x+vA_2F(x)+va_2+uA_1x+uA_2F(x)+ua_1} \\ &= \sum_x (-1)^{\bar{v}F(x)+\bar{u}x+c} \\ &= \pm\mathcal{W}_F(u,v)\end{aligned}$$

□

Proposizione 6.80. *La δ-uniforme differenziabilità è una CCZ-invarianza.*

Dimostrazione. Ricordiamo che quando applicavo alla coppia $(x, F(x))$ l'afifnità λ, ottenevo qualcosa della forma $(F_1(x), F_2(x)) = (y, G(y))$, dove la mappa F_1 era invertibile. Pertanto $(y, F_2 \circ F_1^{-1}(y)) = (y, G(y))$. Quindi posso studiare la δ-uniforme differenziabilità semplicemente studiando il sistema

$$\begin{cases} x + y = a \\ G(x) + G(y) = F_2 \circ F_1^{-1}(x) + F_2 \circ F_1^{-1}(y) = b \end{cases}$$

Poiché F_1 è una permutazione, posso studiare, senza perdere di generalità, il sistema in cui $x = F_1(x)$ e lo stesso per y:

$$\begin{cases} x + y = a \\ F_2(x) + F_2(y) = b \end{cases} \qquad \iff \qquad \begin{bmatrix} F_1(x) \\ F_2(x) \end{bmatrix} + \begin{bmatrix} F_1(y) \\ F_2(y) \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

e se applico $\Lambda^{-1} : (y, G(y)) \mapsto (x, F(x))$ ad entrambi i membri, sfruttando la sua linearità, ottengo il sistema

$$\begin{cases} x + y = a' \\ F(x) + F(y) = b' \end{cases}$$

che ha lo stesso numero di soluzioni. □

Osservazione 6.81. La debole δ-uniforme differenziabilità non è CCZ-invariante.

Il corso si è concluso con lo studio di funzioni particolari, tipo le funzioni almost-bent monomiali della forma x^d, che si dividono in Gold Functions, kasami, Welch, Nilw; certe APN (escludendo i casi di sopra) di cui ricordo solo il nome di Dobertin.

Capitolo 7

Sull'imprimitività di alcuni block cipher

Una prolissa rielaborazione di un vecchio articolo: `http://arxiv.org/abs/0806.4135`.

7.1 Azioni di gruppo

Definizione 7.1 (Azione)**.** Sia A un insieme, e G un gruppo. Un' *azione di gruppo,* o *G-azione*, è una mappa $\varphi : G \times A \to A : (g, a) \mapsto g \cdot a$ dove $\cdot$ è definita tale che:

- $\mathbf{1} \cdot a = a$;
- $g \cdot h \cdot a = (g \cdot h) \cdot a = g \cdot (h \cdot a)$.

Definizione 7.2 (Orbita)**.** Un' *orbita* è una relazione di equivalenza tra gli oggetti di A, tale che

$$x \sim y \iff \exists g \in g \;.\; y = g \cdot x.$$

Le classi di equivalenza vengono dette *orbite*. L'orbita di un elemento $x \in A$ è quindi data da $\mathcal{O}(x) = \{\, g \cdot x \mid g \in G \,\}$.

Esiste un lemma sul numero delle obite, detto *Lemma di Burnside*, il quale afferma che il numero di orbite di un'azione è pari a:

$$\frac{1}{|G|} \sum_{g \in G} |\operatorname{fix}(g)|,$$

dove $\operatorname{fix}(g) = \{\, s \in A \mid g \cdot s = s \,\}$.

Definizione 7.3 (Stabilizzatore)**.** Lo stabilizzatore di una G-azione su A, rispetto ad un $\alpha \in A$, è l'insieme degli elementi che fissano α:

$$G_\alpha = \{\, g \in G \mid \alpha \cdot g = \alpha \,\}$$

Definizione 7.4 (Azione Transitiva)**.** Un'azione di gruppo è detta *transitiva* se esiste un'unica orbita, i.e.:

$$\forall x, y \in A \quad \exists g \in G \;.\; y = g \cdot x$$

Abbiamo in realtà già studiato le *azioni regolari*, che sono azioni transitive (vedi sopra) e semiregolari (ossia $G_x = \{\, 1 \,\}$). Ricordiamo la dirreferenza tra i due:

- *regular action:* $\forall x, y \in A \; \exists! g \in G \;.\; x \overset{g}{\mapsto} y$;
- *transitive action:* $\forall x, y \in A \; \exists g \in G \;.\; x \overset{g}{\mapsto} y$.

L'azione regolare è molto limitante poiché, se G è regolare, allora $|G| = |A|$. I gruppi transitivi invece possono esser molto più grandi.

7.2 Sicurezza dei Cifrari a blocchi

Per il resto di questa sezione, considereremo $\mathcal{C}$ un cifrario a blocchi, in cui lo spazio del plaintext $\mathcal{M}$ è uguale allo spazio del ciphertext, sullo spazio vettoriale $V = \mathbb{F}_{2^n}$, dove $n = ms$ - $s \geq 2$ sarà il blocco usato dall'S-box, ad esempio AES-128 ha $n = 8 \cdot 16$.

Sia $k \in \mathcal{K}$ una chiave appartenente allo spazio delle chiavi; essa induce una permutazione $\tau_k \in \operatorname{Sym}(V)$.

Definizione 7.5 (Translation-Based Cipher)**.** Un cifrario $\mathcal{C}$ è detto *translation based* se è la composizione di un numero finito di round, tale che:

- ogni round può esser scritto come $\gamma\lambda\delta_k$, dove:
  - γ è l'S-box, una *round-dependent bricklayer transformation.*
    Sarebbe a dire che, scrivendo V come somma diretta $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ (in cui $V_i \in \mathbb{F}^m$), $\gamma = (\gamma_1 \cdots \gamma_s)$ con $\gamma_i \in \mathrm{Sym}(V)$ è tale che $\forall v \in V$,
    $$v\gamma = v_1\gamma_1 \oplus v_2\gamma_2 \oplus \cdots \oplus v_s\gamma_s.$$
    Per semplicità, d'ora in avanti chiameremo i V_i *bricks*, e V *wall.*
  - λ è il *mixing layer* una mappa lineare affine da V in V, dipendente dal round, indipendente dalla chiave;
  - $\delta_k \in \tau(V)$ è la somma con la chiave del round corrente, proiettata da $\mathcal{K}$ nello spazio del messaggio.
- per almeno un round, si ha simultaneamente che la proiezione $\mathcal{K} \to V : k \mapsto \bar{k}$ è surgettiva, e λ risulta esser un *MixingLayer proprio*, ossia non esiste una somma diretta nonbanale (i.e. $V, \{0\}$) dei V_i invariante sotto λ (non manda un wall in un wall). Tale round è detto *proprio.*

Ci sono dei sistemi che fanno azioni diverse sulla chiave. Ad esempio, il sistema GOST usato in Russia e Ucraina usa la somma modulare in $\mathbf{Z}_{2^m}$.

Denotiamo con $\Gamma = \Gamma(\mathcal{C})$ il sottospazio del gruppo simmetrico generato da tutti i possibili τ_k. Informalmente, possiamo dire che $\Gamma = \langle \gamma \circ \lambda \circ \delta_k \rangle$ identifica un singolo round. Per poter ottenere una generalizzazione che possa rappresentare cifrari tipo AES, IDEA, DES, SERPENT, si considera il gruppo $\Gamma_\infty = \Gamma_\infty(\mathcal{C})$, il sottogruppo di $\mathrm{Sym}(V)$ generato dal prodotto libero dei Γ_i - in pratica stiamo facendo un numero arbitrario di round.

Osservazione 7.6. $\Gamma \leq \Gamma_\infty$.

Ora, sia V un insieme, e sia $\mathcal{B} = \{B_i\}_i$ una sua partizione, detta *block system*, composta da *blocks.* D'ora in avanti considereremo il caso non banale in cui la partizione non sia l'insieme dei singoletti, o V stesso.

Definizione 7.7 (Imprimitive Block System)**.** $\mathcal{B}$ è un *imprimitive block system* se i blocchi sono rigidi, ossia se $\forall s \in S \;.\; s(B_i) = B_j$. Equivalentemente, possiamo dire che i blocchi sono rigidi se l'immagine di un blocco è un'altro blocco soltanto, o, se preso un punto in un blocco $x \in B_i \;.\; s(x) = y \in B_j \implies s(B_i) = B_j$

Definizione 7.8. $S \subset V$ è detto *primitivo*, se è transitivo e non imprimitivo.

Esempio 7.9. Se Γ_∞ è imprimitivo, allora vi è un sistema di blocchi per Γ_∞, diciamo $\{B_i\}$, tal che:

$$\forall g \in \Gamma_\infty \;\; \forall i \in I \;\; \exists j = j(i) \;.\; g(B_i) = B_j.$$

Osservazione 7.10. Poiché gli elementi di Γ_∞ sono permutazioni, comunque preso $g \in \Gamma_\infty$, $|B_i| = |g(B_i)| \implies |B_i| = |V| \,/\, |I|$.

Osservazione 7.11. Dato S gruppo di permutazioni, si può osservare che:

(i) se $S \subset T$ è regolare, allora T è transitivo;

(ii) se $S \subset T$ è primitivo, allora T è primitivo;

(iii) se $S \subset T$, e T è imprimitivo, allota S è imprimitivo.

Riportiamo ora un risultato comune per la teoria dei gruppi, che sfrutteremo in seguito:

Teorema 7.12. *Sia* $\{B_i\}$ *un sistema di blocchi imprimitivo per* V*. Allora*

$$\exists W < V \;.\; \{B_i\} = \{W + a \mid a \in V\}$$

Attacchi a Cifrari Imprimitivi. Ora, supponiamo $S = \{\varphi_k\}_{k \in K}$ gruppo di permutazioni. Se S è imprimitivo, allora esiste un block system $\mathcal{B} = \{B_i\}_i$, e posso attacarlo con un attacco detto *primitive attack*, di tipo chosen plaintext.

Sia noto il block system $\mathcal{B}$, e sia I l'insieme degli indici dei vari B_i. Fisso x_i, e precomputo $x_i \mapsto \varphi_k(x_i) = y_i \in B_j$. In questo modo mi posso costruire una mappa $I \to I : i \mapsto j$ di indici. A questo punto, identificato un cifrato y appartenente a un dato blocco, posso fare ricerca esaustiva all'interno di quel blocco, per cercare la preimmagine x la cui immagine sia proprio y. La complessità di questo algoritmo sono le $|I|$ precomputazioni iniziali (cifrature), più la ricerca esaustiva su $|B_i|$ (decifrature).

Segue che il costo dell'attacco è $O(|I| + |V|/|I|)$. Possiamo notare che quindi tanti più sono i blocchi, quanto più l'attacco risulta forte. Più pragmaticamente, se abbiamo $V = \mathbb{F}_2^n$ e $I = 2$, allora avremo un costo di $O(2 + 2^{n-1})$, che è praticamente irrilevante. Raggiungiamo il costo minimo ivece quando $|I| \sim |V|/|I|$, ossia se $|I| = 2k \implies O(2^{k+1})$.

Osservazione 7.13. Nota bene che l'attacco agisce sullo spazio del messaggio, non sullo spazio del cifrato. Ad esempio, il cifrario AES cifra sempre su blocchi da 128 bits, e usa keysize differenti (risp. 128, 256 per le varianti AES-128, AES-256). Questo genere di attacco in un caso simile risulterebbe devastante, perché comprometterebbe AES-256 con la stessa complessità di AES-128.

Proposizione 7.14. *$\Gamma_\infty = \langle \tau_k, \bar{\gamma} \circ \lambda' \rangle$, dove $\bar{\gamma}$ è una permutazione costruita su γ che manda $0 \mapsto 0$, e λ' è una trasformazione lineare non affine.*

Dimostrazione. Notiamo anzitutto che, in un cifrario translation-based, $\Gamma_\infty = \langle \{ \gamma \circ \lambda \circ \delta_k \}_k \rangle$, e che $\delta_0 \tau(V) \implies \gamma \circ \lambda \in \Gamma_\infty$. Quindi $\Gamma_\infty = \langle \tau(V), \gamma \circ \lambda \rangle$.

Poiché $\lambda \in AGL(V) \implies \lambda(x) = Ax + b \implies (x)\gamma \circ \lambda \circ \delta_k = A(\gamma(x)) + (b + k)$. Poiché $b + k$ può esser vista come un'unica traslazione delle chiave, senza perder di generalità possiamo quindi supporre che λ sia lineare non affine: $\exists \lambda'(x) = Ax \; . \; \Gamma_\infty = \langle \tau, \gamma \circ \lambda' \rangle$.

Poiché γ è invertibile, chiamo $t = \gamma(0)$, e definisco la funzione $\bar{\gamma} : x \mapsto \gamma(x) + t$. È banale mostrare che $\bar{\gamma}(0) = 0$. Possiamo notare che $\gamma \circ \lambda \circ \delta_k = \lambda(\gamma(x)) + k = \lambda(\gamma(x) + t + t) + k = \lambda(\bar{\gamma}(x)) + (\lambda(t) + k)$. Anche questa volta, possiamo assumere $\lambda(t) + k$ come una possibile traslazione del derivato della chiave.

Unendo tutti queste considerazioni, segue la tesi. □

Proposizione 7.15. *Γ_∞ è imprimitivo sse $\exists U < V$ non banale tale che $\forall u \in U, \; \forall v \in V$ si ha che $(u + v)\gamma\lambda + (v)\gamma\lambda \in U$, ossia equivalentemente $D_U(\gamma)(v) \in U\lambda^{-1}$.*

Dimostrazione. Abbiamo già detto che Γ_∞ è imprimitiva se esiste un sistema di blocchi $\{ B_i \}_i$ tale che

$$\begin{aligned}
&\exists U < V \; . \; \{ B_i \} = \{ U + v \}_{v \in V} \\
\iff &\forall \gamma\lambda \in \Gamma_\infty \quad (U + v)\gamma\lambda = U + v' = U + (v)\gamma\lambda \qquad [\text{perché } U\gamma\lambda = U, \; \gamma(0 + v) = \gamma(v)] \\
\iff &\forall u \in U, \; v \in V \quad (u + v)\gamma\lambda + v\gamma\lambda = D_u(\gamma\lambda(v)) \in U \\
\iff &(u + v)\gamma + v\gamma = D_u(\gamma)(v) \in \lambda^{-1}U.
\end{aligned}$$

□

Teorema 7.16 (Caranti-dalla Volta-Sala)**.** *Dato $\mathcal{C}$ translation-based cipher con $G = \Gamma_h(\mathcal{C})$ un round proprio e $1 \leq r < m/2$. Se un qualunque brick di γ è*

(1) weakly 2^r-uniform e

(2) strongly r-anti-invariant

allora G è primitivo, e dunque Γ_∞ è primitivo.

Dimostrazione. Supponiamo, per assurdo, che G sia imprimitivo. Allora esiste un sistema di blocchi, e allora, per il Teorema 7.12, esiste $U < V \; . \; \{ B_i \} = \{ v + U \}_{v \in V}$. Poiché U è un blocco, abbiamo che $U \; \gamma\lambda = U + v = U$ poiché $0\gamma\lambda = 0 \in U$. Pertanto

$$U\gamma\lambda = U.$$

Consideriamo ora l'insieme I di tutti gli i tali che $\pi_i(U) \neq 0$, dove $\pi_i : V \to V_i : v \mapsto v_i$ è la proiezione canonica. Si ha che:

- o $U \cap V_i = V_i \; \forall i \in I$;
- o $\exists \iota \in I \; . \; U \cap V_\iota \neq V_\iota$.

Nel primo caso, $U = \bigoplus_i V_i$, ma allora è un wall, e $U\gamma = U$ è ancora un wall, poiché γ permuta i brick. Quindi $U\lambda = U$, ma questo è un assurdo con quanto assunto su λ proprio nelle ipotesi.

Nel secondo caso, sia dato $W = U\gamma = U\lambda^{-1}$. Osserviamo, mettendoci nel singolo brick, che

$$(U \cap V_\iota)\gamma_\iota = W \cap V_\iota.$$

Per la Proposizione 7.15 che $\mathrm{Im}(D_u(\gamma_\iota)) \subset W \cap V_\iota \; \forall u \in U \cap V_\iota$. Poiché γ_ι è debolmente 2^r-uniformemente differenziabile,

$$\frac{2^{n-1}}{2^r} < 2^{n-r} \leq |\,\mathrm{Im}(D_u(\gamma_\iota))| \leq |W \cap V_\iota|,$$

da cui segue che $\dim(W \cap V_\iota) \geq n - r$. Ma questo è assurdo poiché γ_ι è strongly r-anti-invariant, e dunque $(U \cap V_i)\gamma_\iota = (W \cap V_\iota) \implies \dim(W \cap V_\iota) < n - r$. □

Corollario 7.17. *È stato dimostrato che* $\Gamma_\infty(\texttt{AES})$, $\Gamma_\infty(\texttt{SERPENT})$, $\Gamma_\infty(\texttt{PRESENT})$, *sono primitivi.*

`TODO: definire il concetto di S-box fortemente propria.`

Esempio 7.18. `AES` non è fortemente proprio. Infatti, l'inversa della prima colonna viene mandata nella prima colonna.

`SERPENT` è fortemente proprio, poiché manda un wall in un non wall. `PRESENT` fa pena.

Capitolo 8

Trapdoors nei Cifrari a Blocchi

Una prolissa rielaborazione di un vecchio articolo: `http://arxiv.org/abs/1411.7681`

Abbiamo studiato $\Gamma_\infty(\mathcal{C}) < \mathrm{Sym}(V)$, dove $V = \mathbb{F}^m$. Vorremmo che questo spazio fosse quanto più grande possibile, idenalmente lo spazio stesso, che ha cardidnalità

$$|\mathrm{Sym}(\mathbb{F}^m)| = 2^m!.$$

Tuttavia questo non è possibile, poiché $\gamma\lambda \subset \mathrm{Alt}(V)$, $\tau(V) \subset \mathrm{Alt}(V)$ (nota: $\mathrm{Alt}(V)$ è lo spazio delle permutazioni pari), quindi il massimo a cui possiamo ambire è che $\Gamma_\infty(\mathcal{C}) = \mathrm{Alt}(V) < \mathrm{Sym}(V)$, dove:

$$|\mathrm{Alt}(V)| = \frac{|\mathrm{Sym}(V)|}{2}.$$

Ora, possiamo dividere $\mathrm{Alt}(V)$ in due sottogruppi massimali, i sottogruppi *primitivi* e quelli *imprimitivi*. Abbiamo studiato i secondi nel capitolo precedente. In questo invece, ci concentreremo sui primi.

Il paper infatti studia come è possibile mettere una *hidden sum*, ossia una trapdoor proveniente da un'alternativa struttura di spazio vettoriale, all'interno di un gruppo primitivo.

Il sottogruppo primitivo è già stato classificato da un certo Ashbacher in 9 gruppi. Incrociando queste famiglie con quelle che contengono le permutazioni, e con delle verifiche sull'ordine, è possibile arrivare a dire che l'unica possibilità per infilare una trapdoor è avere un gruppo isomorfo ad AGL:

Teorema 8.1. *Se $G \sim \mathrm{AGL}(V,+)$, $\tau(V) < G$, allora G è un coniugato di $\mathrm{AGL}(V,+)$, ossia $G = g\,\mathrm{AGL}(V,+)g^{-1}$, e possiamo dunque scriverlo come gruppo affine munito di una operazione binaria $\circ$, diversa da quella normale.*

Definizione 8.2. f è detta *anti-crooked* sse $\mathrm{Im}(\hat{f}_a)$ non è un sottospazio affine di $\mathbb{F}^n$.

Teorema 8.3. *Si può definire una nuova operazione binaria $\cdot$ per dare struttura di anello (detto* di Jacobson radicale*) a $(V,+,\cdot)$ tale che*

- $x + y = x \circ y \circ xy$;
- $x \circ y = x + y + xy$

Teorema 8.4. *Con le condizioni che mi danno la primitività, se inoltre ogni γ_i è anti-crooked, allora $\Gamma_\infty \not\subset AGL(V,\circ)$.*

Dimostrazione. Supponiamo per assurdo che $\Gamma_\infty(\mathcal{C}) \subset \mathrm{AGL}(V,\circ)$. Ricordiamo che avevamo definito $\rho = \gamma\lambda \in \mathrm{AGL}(V,\circ)$ con $0\rho = 0$ poiché λ è lineare. Segue quindi che $(x \circ y)\rho = x\rho \circ y\rho \ \forall x, y$.

Consideriamo $U = \{\, z \in V \mid xz = 0 \ \forall x \in V \,\} \neq \emptyset$ (questo fatto è dimostrabile, ma lo daremo per buono). Allora, fisso un $y \in U$ (`TODO: è detto spazio isotopo?`), e posso notare che:

$$\begin{aligned}(x+y)\rho &= (x \circ y \circ xy)\rho \\ &= (x \circ y)\rho && \text{[per costruzione su } y\text{]} \\ &= x\rho \circ y\rho = x\rho + y\rho + x\rho \cdot y\rho.\end{aligned}$$

Considero a questo punto xV, che è un sottospazio rispetto ad entrambe le operazioni, $+$ e $\circ$, noto che

$$\mathrm{Im}(\hat{\rho}_y) = \{\, (x+y)\rho + x\rho \mid x \in V \,\} = \{\, y\gamma + y\gamma \cdot x\gamma \mid x \in V \,\} = y\gamma + y\gamma \cdot V\rho$$

`TODO: spiegare meglio` coset di un sottospazio rispetto ad entrambe le operazioni.

Restringendoci quindi alla proiezione sul singolo brick $V_i = \mathbb{F}^m$, abbiamo che:

$$\mathrm{Im}(\hat{\gamma}_i)_y = \{\, (x + y_i)\gamma + x\gamma \mid x \in V_i \,\} \quad \text{è un sottospazio affine di } \mathbb{F}^m$$

ma questo è un assurdo con l'ipotesi su γ di essere anti-crooked. □

Corollario 8.5. $\Gamma_\infty(\texttt{AES})$ *ha* γ_i *anti-crooked per ogni* i*, quindi soddisfa le ipotesi del teorema.* $\Gamma_\infty(\texttt{SERPENT})$ *usa* 8 *s-Box, di bui alcune sono anti-crooked, altre no.* [1] $\Gamma_\infty(\texttt{PRESENT})$ *non ha i* γ_i *anti-crooked.*

[1] È importante ricordare qui che comunque stiamo studiando delle estensioni dei cifrari, in questo e nel caso successivo quindi non possiamo affermare nulla.

Capitolo 9

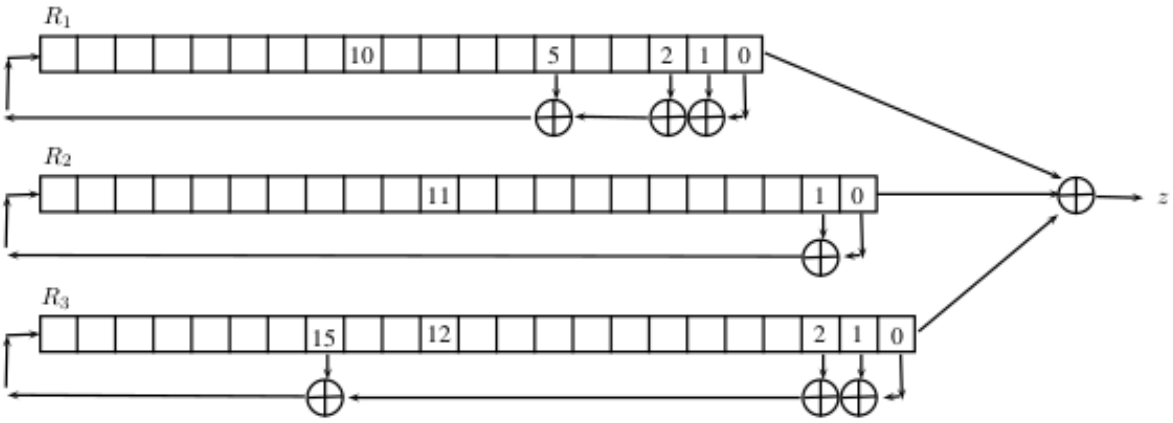
Esempi di Cifrari

9.1 A5/1

$A5/1$ è uno stream cipher usato in telefonia mobile, che agisce su blocchi di 228 bits, sfruttando 3 registri definiti nel modo seguente:

Register	Polynomial	Taps	Clocks
R_1	$x^{19} + x^5 + x^2 + x^1 + 1$	$0, 1, 2, 5$	10
R_2	$x^{22} + x + 1$	$0, 1$	11
R_3	$x^{23} + x^{15} + x^2 + x + 1$	$0, 1, 2, 15$	12

Nota: tali polinomi sono tutti i polinomi primitivi di $\mathbb{F}^{19}$, $\mathbb{F}^{22}$, $\mathbb{F}^{23}$.



Key-Loading. I registri sono inizialmente inizializzati a 0. La chiave K di 64 bit e l'IV di 22 bits vengono quindi caricati (in questo ordine esatto) facendo dei clock standard su ogni registro, ma sommando alla funzione `update()` l'i-esimo bit della chiave (e successivamente dell'IV). A seguito di questo, vi è una fase detta di *warm-up* in cui vengono fatti 100 *clock irregolari* (i.e. l'i-esimo registro effettua un clock sse il suo bit di clock è quello della funzione `maj()`). In questo modo, ogni registro ha probabilità 3/4 di effettuare un clock.

Update. Una volta che questi passi sono completi, vengono generati 228 bit, detti di *keystream*, considerando il bit più significativo restituito dalla funzione `update()`.

Il cifrato viene ottenuto facendo lo `xor` bit -a-bit del messaggio con la chiave. Infine, nel caso in cui vi fossero altri dati da trasmettere, viene effettuato un *re-seeding* e si ripete la procedura.

9.2 E0

E0 viene usato nello standard bluetooth. Si tratta di uno stream cipher con 4 registri lineari e uno nonlineare, in cui i feedback polynomials sono primitivi e i seguenti: Poiché i 4 clock sono regolari e i polinomi primitivi, abbiamo

Register	Polynomial
R_1	$x^{25} + x^{20} + x^{12} + x^8 + 1$
R_2	$x^{31} + x^{24} + x^{16} + x^{12} + 1$
R_3	$x^{33} + x^{28} + x^{24} + x^4 + 1$
R_4	$x^{39} + x^{36} + x^{28} + x^4 + 1$

che il periodo è almeno

$$(2^{39} - 1)(2^{33} - 1)(2^{31} - 1)(2^{25} - 1).$$

L'altro registro, quello nonlineare, è caratterizzato da quattro celle (c_0, c_1, c_2, c_3), e contrariamente ai precedenti mi viene ottenuto facendo lo shift a sinistra di due bit, e considerando la notazione binara dell'intero

$$z = \left\lfloor \frac{s_1 + s_2 + s_3 + s_4 + 2c_3 + c_2}{2} \right\rfloor \quad \text{dove } s_i \text{ è l'output dell'}i\text{-esimo registro}$$

che poiché $0 \leq z < 4$, può esser espresso mediante due soli bit. Nota che qui la non lienarità veiene data dal fatto che la somma di vettori su $\mathbb{F}^n$ come interi è veloce e *nonlienare* rispetto a $\mathbb{F}^n$.

9.3 AES

`AES` è un cifrario a blocchi che agisce su:

- uno spazio dei messaggi $\mathcal{M} = \mathbb{F}^n$, dove $n = 128$;
- uno spazio delle chiavi $\mathcal{K} = \mathbb{F}^l$, dove $l = 128,\ 192,\ 256$;
- un numero di round N pari a $10, 12, 14$ per le rispettive grandezze delle chiavi;
- le s-Box agiscono su 8 bit.

Per semplicità esplicativa, terremo conto solo di `AES-128`.

Osservazione 9.1. Solo relativamente di recente è stato provato, con un attacco che operava su uno spazio di 2^{126} bit, che `AES` non è un *cifrario ideale* (i.e. l'attacco migliore è pari al bruteforce sullo spazio delle chiavi).

I round funzionano nel modo seguente:

R_0 : `AddRoundKey()` (: δ_k), in cui la chiave sommata è quella fornita dall'utente;

R_i $\forall 1 \leq i \leq 9$: si compone delle funzioni `SubBytes()` (: γ), `ShiftRows()` e `MixColumns()` (: λ), `AddRoundKey()` (: δ_k).

R_{10} : è detto *round atipico*, e la trasformazione λ si compone della sola funzione `ShiftRows()`, e viene denotata con $\bar{\lambda}$;

Abbiamo quindi che ogni round è la composizione di $\gamma\lambda\delta_k$, dove però il primo round è tale che $\gamma = \lambda = \mathbf{1}_v$ e l'ultimo round è atipico con $\bar{\lambda}$.

La funzione $\delta_{k^{(i)}} =$ `AddRoundKey()`, in ogni round i, estrae una chiave $k^{(i)}$ mediante il `KeySchedule()` - che divide la chiave in tre blocchi e su di essi applica alcuni `xor` e una funzione nonlineare - ed effettua una somma con essa.

La funzione $\gamma =$ `SubBytes`() applica ad `ogni byte` la stessa trasformazione $\gamma_i : \mathbb{F}_{256} \to \mathbb{F}_{256}$, dove γ_i è la composizione di un'inversione (la *patched inversion*) con una trasformazione affine.

La funzione λ è la composizione di duetrasformazioni lineari:

- la funzione `ShiftRows()` consiste in un increased shift delle righe:

$$\begin{matrix} \mathbf{s_0} & s_4 & s_8 & s_{12} \\ s_1 & \mathbf{s_5} & s_9 & s_{13} \\ s_2 & s_6 & \mathbf{s_{10}} & s_{14} \\ s_3 & s_7 & s_{11} & \mathbf{s_{15}} \end{matrix} \quad \mapsto \quad \begin{matrix} \mathbf{s_0} & s_4 & s_8 & s_{12} \\ \mathbf{s_5} & s_9 & s_{13} & s_1 \\ \mathbf{s_{10}} & s_{14} & s_2 & s_6 \\ \mathbf{s_{15}} & s_3 & s_7 & s_{11} \end{matrix}$$

- la funzione `MixColumns()` moltiplica ogni colonna degli stati per una stessa matrice:

$$\begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 01 \\ 03 & 01 & 02 & 01 \\ 03 & 01 & 01 & 02 \end{bmatrix}$$

nota: è la prima riga shiftata incrementalmente

9.4 Serpent

`SERPENT` è un cifrario a blocchi che agisce su:

- uno spazio dei messaggi $\mathcal{M} = \mathbb{F}^n$, dove $n = 128$;
- unospazio delle chiavi $\mathcal{K} = \mathbb{F}^l$, dove $l = 128, 192, 256$ (inizialmente 128 ma sviluppato perché fosse variabile);
- un numero di round N pari a 32;
- le s-Box agiscono su 4 bit, e ve ne sono 8 differenti;

Osservazione 9.2. `SERPENT` potrebbe essere un cifrario ideale perché non è ancora stato trovato un attacco alla versione a 32 rounds che avesse complessità minore di $O(2^{128})$: i migliori attacchi sono di tipo known-plaintext e agiscono su 10 e 11 round, con rispettivo costo di 2^{89} e 2^{107}.

La cifratura consiste un:

- una permutazione iniziale $\pi : V \to V$ costruita non tanto per la sicurezza quanto per facilitare l'implementazione
- una traslazione con la prima chiave di round;
- $N - 1$ round con la stessa struttura, composizione di $\gamma\lambda\delta_k$, dove γ è una *parallel s-Box*, λ è il *linear mixing layer*, e δ_k è la *traslazione della chiave*;
- l'ultimo round 32-esimo, consta di $\gamma\lambda\delta_k$ con $\lambda = \mathbf{1}_V$, similmente ad `AES`;
- l'inversa della permutazione iniziale π^{-1}.

La funzione $\gamma \in \mathrm{Sym}(V)$ è tale che

$$v\gamma = v_1\gamma_1 \cdot \cdots \cdot v_{32}\gamma_{32} \qquad \text{dove} v_i \in \mathbb{F}^4 \text{ e } \gamma_i \in \mathrm{Sym}(\mathbb{F}^4).$$

Su ogni blocco applichiamo la stessa s-Box, che agisce su gruppi di 4-bit, e al round i-esimo, applichiamo la $i \pmod 8$ s-Box.

La trasformazione lineare λ agisce sull'output delle s-Box, rimappandolo su 4 parole da 32 bits.