# Appunti sui Codici di Reed Muller

Giovanni Barbarino

# Capitolo 1

# Codici di Reed-Muller

I codici di Reed-Muller sono codici lineari su $\mathbb{F}_q$ legati alle valutazioni dei polinomi sullo spazio affine. Per semplicità supporremo in questo caso di lavorare su $\mathbb{F}_2$. Consideriamo quindi lo spazio vettoriale $(\mathbb{F}_2)^s$ ed enumeriamone gli elementi

$$(\mathbb{F}_2)^s = \{v_1, \dots, v_{2^s}\}$$

Dato un polinomio $f \in \mathbb{F}_2[x_0, \dots, x_{s-1}]$, possiamo associargli il vettore di $\mathbb{F}_2^{2^s}$ costituito da $(f(v_1), \dots, f(v_{2^s}))$.

**Definizione 1.1.** Il codice di Reed-Muller $\mathcal{R}(m, r)$ di lunghezza $2^m$ e ordine $r$ è il sottospazio di $(\mathbb{F}_2)^{2^m}$ dato dai vettori

$$(f(v_1), \dots, f(v_{2^m}))$$

al variare di $f$ tra i polinomi di grado $\leq r$ in $\mathbb{F}_2[x_0, \dots, x_{m-1}]$.

Studiamo la dimensione di un codice di Reed-Muller. Indichiamo i polinomi di grado $\leq r$ con $\mathbb{F}_2[x_0, \dots, x_{m-1}]_{\leq r}$. Consideriamo l'applicazione lineare

$$\begin{array}{rccc} \varphi\colon & \mathbb{F}_2[x_0, \dots, x_{m-1}] & \longrightarrow & \mathbb{F}_2^{2^m} \\ & f & \longmapsto & (f(v_1), \dots, f(v_{2^m})) \end{array}$$

Allora $\mathrm{Ker}(\varphi)$ è l'ideale generato da $(x_i^2 - x_i)$, [1] e dunque il codice ha dimensione uguale al numero di monomi liberi da quadrati di grado $\leq r$:

$$1 + \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{r}$$

*Esempio.* Il codice $\mathcal{R}(m, 0)$ è formato dai vettori $(0, \dots, 0)$ e $(1, \dots, 1)$ e dunque è il codice di ripetizione. Il codice $\mathcal{R}(m, m)$ è invece il codice formato da tutti gli elementi di $\mathbb{F}_2^{2^m}$ perché

$$\sum_{i=0}^{m} \binom{m}{i} = 2^m$$

e dunque la dimensione del codice è uguale alla dimensione dello spazio vettoriale.

[1] per dimostrarlo, la maniera piú semplice é per induzione sul numero di variabili

Cerchiamo ora di capire quale sia la distanza di $\mathcal{R}(m,r)$. Intanto, notiamo che il polinomio $f = x_0x_1\dots x_{r-1}$ è tale che $\varphi(f)$ abbia $2^m - 2^{m-r}$ componenti nulle. Questo mostra che la distanza di un codice di Reed-Muller $\mathcal{R}(m,r)$ è $\leq 2^{m-r}$. Per mostrare l'altra disuguaglianza, utilizziamo il seguente lemma:

**Lemma 1.2.** Sia $f \in \mathbb{F}_2[x_0,\dots,x_{m-1}]$ di grado $s$ (se $f=0$, si può considerare $s=+\infty$). Allora $\#V(f) \leq 2^m - 2^{m-s}$.

*Dimostrazione.* Osserviamo immediatamente che, se $\tilde{s} \leq s$, allora $2^m - 2^{m-\tilde{s}} \leq 2^m - 2^{m-s}$, quindi basta dimostrarlo con $f$ di grado $s$.

Dimostriamo prima il risultato per gli $f$ con monomi squarefree, per induzione sul numero di variabili. Per $m=1$ il risultato è ovvio ($f \in \{0,\,1,\,x,\,x+1\}$). Supponendo la tesi vera per $m$, osserviamo che, per $f$ con monomi squarefree, si può sempre scrivere

$$f(x_0,\dots,x_m) = x_m f_1(x_0,\dots,x_{m-1}) + (1-x_m)f_0(x_0,\dots,x_{m-1}),$$

dove $f_\ell(x_0,\dots,x_{m-1}) = f(x_0,\dots,x_{m-1},\ell)$. Si osserva che, a seconda che $x_m$ sia 0 o 1, le radici di $f$ annullano uno tra $f_0$ e $f_1$, dunque

$$\#V(f) = \#V(f_0) + \#V(f_1)$$

e banalmente $\#V(f_0), \#V(f_1) \leq 2^m$. Se $f_0, f_1 \neq 0$, allora per ipotesi induttiva

$$\#V(f_0), \#V(f_1) \leq 2^m - 2^{m-s} \implies \#V(f) \leq 2^{m+1} - 2^{m+1-s}$$

Se wlog $f_1 = 0$, in particolare $\deg(f_0) \leq s-1$ e per ipotesi induttiva $\#V(f_0) \leq 2^m - 2^{m-s+1}$ (è un polinomio di grado al più $s-1$ in $m$ variabili), da cui $\#V(f) \leq 2^m - 2^{m-s+1} + 2^m = 2^{m+1} - 2^{m+1-s}$.

D'altra parte, se $f$ non fosse a monomi squarefree, chiamando $\tilde{f}$ il polinomio di grado $\tilde{s} \leq s$ ottenuto rendendo squarefree i monomi di $f$, le valutazioni di $f$ e $\tilde{f}$ coincidono, quindi $\#V(f) = \#V(\tilde{f})$. Applicando il risultato a $\tilde{f}$, otteniamo

$$\#V(f) \leq 2^m - 2^{m-\tilde{s}} \leq 2^m - 2^{m-s}. \qquad \square$$

**Interpolazione** Sfruttiamo l'interpolazione polinomiale in più variabili. Sia $c$ il messaggio ricevuto e supponiamo di voler ricostruire il messaggio del mittente.

Per prima cosa, troviamo un polinomio con monomi liberi da quadrati tale che $c$ sia il vettore delle valutazioni sui punti di $(\mathbb{F}_2)^m$. Questo é possibile farlo tramite interpolazione multivariata: dato $v \in (\mathbb{F}_2)^m$ chiamiamo $L_v$ il polinomio

$$L_v(x) = (1+x_0+v_0)(1+x_1+v_1)\dots(1+x_{m-1}+v_{m-1})$$

che ha la proprietá $L_v(x) = 1 \iff x = v$. Da questo, possiamo scrivere il polinomio interpolatore di $c$ come

$$p_c(x) = \sum_{c_v=1} L_v(x)$$

Notiamo che $p_c(x)$ é un polinomio con monomi liberi da quadrati, dunque é l'unico di questo tipo che interpoli $c$. Se il suo grado é minore o uguale ad $r$, allora abbiamo ricostruito il messaggio originale.

Notiamo anche che imporre $p_c(v) = c_v$ é una condizione lineare sui coefficienti di $p_c(x)$, dunque per ricostruire un polinomio del codice, basta imporre tante condizioni quanto é la dimensione del codice.

L'interpolazione ci permette di dire se un messaggio è stato corrotto o meno, ma la decodifica vera e propria è più complicata, e la vedremo nel prossimo capitolo.

**Definizione Ricorsiva** I codici di Reed Muller si possono caratterizzare anche in maniera ricorsiva. Abbiamo giá detto che $\mathcal{R}(m, 0)$ sono i codici di ripetizione, e $\mathcal{R}(m, k) \cong \mathbb{F}_2^m$ per $k \geq m$. Cerchiamo di ricavare $\mathcal{R}(m, k)$ con $m > k$ in relazione a quelli con $k$ minori.

Preso un polinomio $p(x)$ con monomi squarefree su $\mathbb{F}_2[x_0, \ldots, x_{m-1}]$ di grado minore o uguale a $k$, lo spezziamo nei monomi che contengono $x_0$, e quelli che non lo contengono, ossia $p(x) = p_0(x) + x_0 p_1(x)$. Il relativo messaggio del codice $\mathcal{R}(m, k)$ puó essere spezzato in due parti: le valutazioni con $x_0 = 0$ e quelle con $x_1 = 1$, indicandole con $c = (c_0, c_1)$.

- $p_0(x)$ non contiene $x_0$, dunque le sue valutazioni non dipendono dalla prima variabile, dunque saranno del tipo $(\alpha, \alpha)$, ma visto che $p_0(x)$ puó essere visto come un polinomio in $\mathbb{F}_2[x_1, \ldots, x_{m-1}]$ di grado minore o uguale a $k$, allora $\alpha$ appartiene a $\mathcal{R}(m-1, k)$.
- le valutazioni di $x_0 p_1(x)$ saranno del tipo $(0, \beta)$, ma visto che $p_1(x)$ non contiene $x_0$, puó essere visto come un polinomio in $\mathbb{F}_2[x_1, \ldots, x_{m-1}]$ di grado minore o uguale a $k-1$, dunque $\beta$ appartiene a $\mathcal{R}(m-1, k-1)$.

Questo ci dice che le valutazioni di $p(x)$ sono $(\alpha, \alpha + \beta)$, con $\alpha \in \mathcal{R}(m-1, k)$ e $\beta \in \mathcal{R}(m-1, k-1)$. Questo ci dá una definizione equivalente:

**Definizione 1.3.** I codici di Reed Muller $\mathcal{R}(m, k)$ sono definiti come

- $\mathcal{R}(m, 0)$ sono i codici di ripetizione di lunghezza $2^m$.
- se $k \geq m$, allora $\mathcal{R}(m, k)$ é isomorfo a $\mathbb{F}_2^k$.
- $\mathcal{R}(m, k)$ con $k < m$ é ricavato ricorsivamente come

$$\{v \in \mathbb{F}_2^m \,:\, v = (\alpha, \alpha + \beta),\ \alpha \in \mathcal{R}(m-1, k),\ \beta \in \mathcal{R}(m-1, k-1)\}$$

Da questa definizione possiamo dare una dimostrazione simile che la distanza di $\mathcal{R}(m-1, k)$ sia $2^{m-k}$, sempre per induzione sul numero di variabili. Infatti notiamo che $\mathcal{R}(m, 0)$ ha distanza $2^m$, mentre $\mathcal{R}(k, k)$ ha distanza 1. Dato che $\mathcal{R}(m, k)$ contiene tutte le stringhe del tipo $(\alpha, \alpha)$, con $\alpha \in \mathcal{R}(m-1, k)$ allora la distanza é al massimo

$$dist(\mathcal{R}(m, k)) \leq 2 \cdot dist(\mathcal{R}(m-1, k)) = 2^{m-k}$$

Dato che $\mathcal{R}(m-1, k-1) \subseteq \mathcal{R}(m-1, k)$, ogni elemento del codice $\mathcal{R}(m, k)$ é composto da due stringhe di $\mathcal{R}(m-1, k)$, da cui la distanza può calare solo se una delle stringhe $\alpha$ o $\alpha + \beta$ è nulla, ossia se

$$c = (0, \beta) \qquad \text{oppure} \qquad c = (\beta, 0)$$

ma in entrambi i casi, $\beta \in \mathcal{R}(m-1, k-1)$, dunque il suo peso è comunque maggiore o uguale alla distanza del codice, che è $2^{(m-1)-(k-1)} = 2^{m-k}$.

# Capitolo 2

# Decodifica

Scriviamo ora la decodifica per un codice di Reed-Muller. Per farlo, descriviamo un metodo di decodifica più generale.

Se $v$ è un vettore ortogonale ad un codice binario $C$, lo chiamiamo *test di parità*. Notiamo che, dato un qualsiasi vettore $x$,

$$x \in C \implies x^t v = 0$$

dunque diciamo che il test di parità fallisce su $x$ se $x^t v = 1$.

**Definizione 2.1.** Sia $C$ un codice, e $\{v_1, \dots, v_r\}$ dei test di parità. Questi si dicono *concentrati sulla coordinata $i$* se $v_j^{(i)} = 1$ per ogni vettore, mentre ogni altra coordinata ha al massimo un vettore che non si annulla su di essa.

**Lemma 2.2.** Sia $C$ un codice, e $\{v_1, \dots, v_r\}$ concentrati sulla coordinata $i$. Se $\widetilde{c} = c + e$ con $c \in C$ ed $e$ un errore di peso al massimo $\lfloor \frac{r}{2} \rfloor$, allora $e_i = 1$ se e solo più della metà dei test di parità falliscono.

*Dimostrazione.* Se $e_i = 0$, allora i bit di errore fanno fallire al massimo la metà dei test di parità. Viceversa, se $e_i = 1$, i test che non falliscono sono meno della metà, per lo stesso motivo. □

**Corollario 2.3.** Se un codice $C$ possiede set di $r$ test di parità concentrati su ogni coordinata, allora riesce a correggere $\lfloor \frac{r}{2} \rfloor$ errori

*Osservazione* 2.4. Per i codici ciclici, basta trovare $r$ test di parità concentrati su una sola coordinata.

**Lemma 2.5.** In un codice $C$ di lunghezza $n$, se è possibile trovare set di $r$ test di parità concentrati su ogni coordinata, allora

$$r \leq \frac{n-1}{d^\perp - 1}$$

dove $d^\perp$ è la distanza del codice duale a $C$.

*Dimostrazione.* Notiamo che se abbiamo un set di $r$ test di parità $\{v_1, \dots, v_r\}$ concentrati su una coordinata, allora necessariamente la somma dei pesi di $v_i$ meno 1 deve fare al massimo $n-1$, ma dato che i $v_i$ stanno nel codice duale, la tesi segue. □

Notiamo che questo bound non è molto buono, dunque cerchiamo un metodo per migliorarlo.

**Definizione 2.6.** Sia $C$ un codice di lunghezza $n$, $S$ un sottoinsieme degli indici da 1 a $n$, e $\{v_1, \dots, v_r\}$ dei test di parità. Questi si dicono concentrati su $S$ se $v_j^{(i)} = 1$ per ogni vettore e per ogni $i \in S$, mentre ogni altra coordinata ha al massimo un vettore che non si annulla su di essa.

Notiamo che, con la stessa dimostrazione, vale un risultato simile al lemma sopra

**Lemma 2.7.** Sia $C$ un codice, e $\{v_1, \dots, v_r\}$ concentrati sul $S$. Se $\widetilde{c} = c + e$ con $c \in C$ ed $e$ un errore di peso al massimo $\lfloor \frac{r}{2} \rfloor$, allora $\sum_{i \in S} e_i = 1$ se e solo più della metà dei test di parità falliscono.

Grazie a questi nuovi strumenti, possiamo attuare una decodifica di *maggioranza logica ad L step*. Descriviamo questo metodo:

- Dato $C$ un codice binario, e $c \in C$, sia $\widetilde{c} = c + e$ un messaggio corrotto, con al massimo $t$ errori
- Con $2t$ test di parità concentrati su $S$, possiamo ricavare la somma degli errori su $S$.
- Preso adesso un $R$ un sottoinsieme degli indici, se troviamo $2t$ insiemi $S_i$ calcolati al passo precedente, la cui intersezione sia $R$, e tali che $S_i/R$ siano disgiunti, possiamo ricavare la somma degli errori su $R$ nella stessa maniera sopra, ossia a maggioranza.
- Possiamo ripetere il processo prendendo $2t$ insiemi $R_i$ per calcolare la somma degli errori su un insieme più piccolo, e così via fino a calcolare l'errore su un singolo indice.

Vediamo dunque come si applica questo algoritmo ai codici di Reed Muller. Abbiamo bisogno prima di un po' di lemmi preparatori.

**Lemma 2.8.** Dato $V \subseteq \mathbb{F}_2^m$ un sottospazio affine di dimensione $m - r$, allora il polinomio con monomi squarefree che si annulla esattamente sul suo complementare ha grado $r$.

*Dimostrazione.* Il sottospazio $V$ è definito da $r$ equazioni lineari affini $f_1, \dots f_r$, e il prodotto $(1 - f_1) \dots (1 - f_r)$ è un polinomio di grado $\leq r$ che si annulla esattamente sul complementare di $V$. Se però avesse grado minore di $r$, allora apparterrebbe a $\mathcal{R}(m, r-1)$, che ha distanza maggiore di $2^{m-r}$. □

**Lemma 2.9.** Visti come sottospazi di $(\mathbb{F}_2)^{2^m}$, l'ortogonale del codice $\mathcal{R}(m, k)$ è $\mathcal{R}(m, m-k-1)$

*Dimostrazione.* Notiamo che su ogni campo, la somma delle dimensioni di un sottospazio e del suo ortogonale è la dimensione dello spazio ambiente. Dato che

$$\dim \mathcal{R}(m, k) + \dim \mathcal{R}(m, m-k-1) = 2^m$$

allora basta far vedere che prese due stringhe, una dalla base di ogni codice, queste sono ortogonali. Siano dunque $c_1 \in \mathcal{R}(m, k)$ e $c_2 \in \mathcal{R}(m, m-k-1)$

due stringhe di base, dove la base è composta dai monomi squarefree. Chiamati $p_1(x)$ e $p_2(x)$ i rispettivi monomi, allora il prodotto componente per componente di $c_1$ e $c_2$ (che chiameremo $c$) sarà associato al monomio $p(x) = p_1(x)p_2(x)$ di grado $< m$. Riducendolo ad un monomio squarefree, avrà grado $r < m$, ma dato che è ancora è un monomio, allora il peso di $c$ sarà $2^{m-r}$, che è pari, e pertanto $c_1$ e $c_2$ sono ortogonali. □

Questi due lemmi ci dicono che un qualsiasi spazio affine di dimensione $r+1$ fornisce un vettore ortogonale a $\mathcal{R}(m, r)$.

Prendiamo dunque $V$ spazio affine di dimensione $s \leq r$. esistono esattamente $\geq 2^{m-r} - 1$ spazi affini $W_i$ di dimensione $s + 1$ che contengono $V$, e l'intersezione di due di essi è esattamente $V$. Inoltre, per motivi di dimensione, $W_i/V$ sono disgiunti. Nel caso $s = r$, abbiamo trovato un set di $2^{m-r} - 1$ test di parità concentrati su $V$, mentre nel caso $s < r$ possiamo iterativamente trovare comunque abbastanza sottospazi concentrati su $V$. Ciò vuol dire che possiamo applicare la decodifica a logica maggioritaria a $(t + 1)$ step (perchè dobbiamo arrivare al singolo punto, che ha dimensione zero).