UNIVERSITA' DEGLI STUDI DI TORINO

Facoltà di Scienze Matematiche Fisiche Naturali
Corso di Laurea in Matematica

Tesi di Laurea Magistrale

La Trasformata di Winograd nella Teoria dei Codici Correttori

Relatore
Prof. **Umberto Cerruti**
Università di Torino

Controrelatrice
Prof.ssa **Lea Terracini**
Università di Torino

Laureando
Sebastiano Ferraris

Sessione di Luglio 2013
Anno Accademico 2012-2013

Introduzione

"Calcolo esatto per cominciare a conoscere tutte le cose esistenti e tutti i segreti oscuri e misteriosi. "

- Ahmes, 1600 a.C.

Il punto di partenza di questa tesi sono alcune lezioni di Laboratorio di Applicazioni dell'Algebra tenuto nel 2011 ed un pre-print del prof. Umberto Cerruti [8] che ho di utilizzato per approfondire diversi argomenti inerenti la teoria dei codici correttori che hanno suscitato il mio interesse e la mia curiosità. I principali che propongo in questa tesi sono:

capitolo 1 Scomposizione dell'algebra $\mathcal{R}_{r,\mathbb{F}} = {\mathbb{F}[x]}\Big/{(x^r-1)}$ in un prodotto di campi: attraverso lo studio della fattorizzazione di x^r-1 tramite un gruppo isomorfo al gruppo di Galois $Gal(\mathbb{F}(\xi),\mathbb{F}))$ che agisce sul gruppo generato da x in $\mathcal{R}_{r,\mathbb{F}}$ arriviamo a costruire una nuova algebra sui fattori irriducibili del polinomio x^r-1. Ricordando che per $M^{(v)}(x)$ fattore irriducibile di x^r-1, ogni quoziente

$$\mathbb{F}[x]\Big/M^{(v)}(x)$$

è un campo, costruiamo la nuova algebra come il prodotto di tali campi. Questa è ancora isomorfa a $\mathcal{R}_{r,\mathbb{F}}$ e l'isomorfismo così creato viene definito **trasformata di Winograd**.
Sempre nel capitolo 1 dimostriamo una formula basata sul teorema di Burnside per determinare la cardinalità dell'insieme dei fattori irriducibili $M^{(v)}(x)$.

capitolo 2 Studio degli ideali e degli idempotenti di $\mathcal{R}_{r,q}$: a partire da questo punto proseguiamo considerando solo i campi finiti per orientare la direzione sulle applicazioni alla teoria dei codici correttori. Dopo la definizione di alcuni operatori su $\mathcal{R}_{r,q}$ presentiamo uno studio sugli ideali e sugli idempotenti che giocano un ruolo fondamentale nella teoria dei codici correttori. Data la semplicità degli ideali e degli idempotenti di un campo, questo studio non avviene direttamente su $\mathcal{R}_{r,q}$ ma sulla scomposizione in campi ricavata nel capitolo precedente.

capitolo 3 Trasformata di Winograd come trasformazione lineare fra $\mathcal{R}_{r,q}$ ed il prodotto di campi in cui si scompone: approfondiamo la definizione di trasformata di Winograd ricavando la matrice associata alla trasformazione

e la matrice inversa e presentiamo alcune delle sue proprietà più rilevanti per gli scopi della tesi. Per poter definire la matrice di trasformazione in modo più semplice partiamo dalla definizione di una ulteriore algebra, che chiamiamo **algebra dei vettori circolanti concatenati**. Questa continua ad essere un'algebra isomorfa a quelle già proposte e mantiene la struttura di prodotto di campi, ma con una forma più efficace per parlare di matrici di trasformazioni. Anche in questo capitolo i vari sviluppi della teoria sono accompagnati da esempi numerici.

capitoli 4 e 5 Introduzione alla teoria dei codici correttori: in questo interludio presentiamo la teoria dei codici correttori dalle basi, per arrivare a definire i codici lineari, i codici ciclici ed i codici BCH. Nel corso del capitolo utilizziamo la maggior parte dei risultati ottenuti nei capitoli 1 e 2 e prepariamo il terreno per poter presentare le applicazioni della trasformata di Winograd ai codici correttori, scopo della tesi.

capitolo 6 Applicazioni dello studio di $\mathcal{R}_{r,q}$ e della trasformata di Winograd alla teoria dei codici correttori: come prima applicazione vediamo che una scelta di blocchi della trasformata di Winograd definisce una matrice che può essere usata come matrice di controllo o come matrice generatrice di determinati codici correttori. La seconda applicazione è un sistema per codificare e decodificare un messaggio che permette di diminuire la quantità di informazione per inviare un messaggio codificato, individuando in ogni parola alcuni sottovettori che non non contengono informazioni rilevanti.

Originariamente la trasformata di Winograd è stata scoperta come strumento per diminuire la complessità computazionale del prodotto di convoluzione e come alternativa alla trasformata di Fourier discreta.
È stata presentata per la prima volta nell'articolo *On Computing the Discrete Fourier Transform* [31] di Shmuel Winograd.
In questa ricerca ho omesso i collegamenti fra la trasformata di Winograd e la Trasformata di Fourier, ho omesso le implicazioni con la teoria dei codici spettrale e non ho parlato di una applicazione della trasformata di Winograd alla teoria dei codici correttori scoperta da Miller, Truong e Reed presentata nel 1980 con l'articolo *Efficient Program for decoding the* $(255, 223)$ *Reed-Solomon Code over* $GF(2^8)$ [22].
Ho invece considerato la trasformata di Winograd come una trasformazione lineare fra due spazi vettoriali, esaminando due applicazioni ai codici ciclici.
Le fonti principali, oltre al già citato articolo del relatore della tesi, sono *Theory and Practice of Error Control Codes*, di Richard E. Blahut [4] ed *Algebra e teoria dei codici correttori* di Luigia Berardi [2].

Ringraziamenti

Sono stati innumerevoli gli aiuti ed i contributi, diretti o indiretti, grazie ai quali ho potuto procedere con la tesi. I più diretti alla risoluzione di alcuni problemi sono stati sicuramente quelli di Stefano Barbero che ringrazio per la sua perenne disponibilità a risolvere i dubbi degli studenti di Matematica. Di palazzo Campana vorrei anche ringraziare Paolo Martini, Andrea Montabone, Nadir Murru, Simone Garruto, Riccardo Jadanza, Andrea Ricolfi e Michele Voto per l'amicizia e per avere condiviso con me le loro idee.
Ringrazio Giuliano De Rossi ex-collega della tc-web, per i suggerimenti sui diagrammi commutativi e i colleghi della sim-tec che sono stati loro malgrado assillati dalle mie deformazioni matematiche. Ringrazio anche Filippo Ferraris per i suggerimenti sulla stesura (che non ho avuto modo di seguire) e per l'interesse che ha sempre dimostrato nei miei studi. Ringrazio infine, per il sostegno morale Francesco Giovo e Barbara Bosia, per quello immorale Andrea Baglione e per quello immortale Federica Narciso.

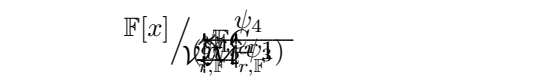
Sebastiano Ferraris, *Villar Dora* 2013.

Indice

Capitolo 0

Quattro algebre isomorfe

Il tema principale della prima parte della tesi riguarda lo studio di alcune strutture algebriche e delle trasformazioni che si possono definire fra loro. In questo capitolo vediamo la struttura algebrica dei polinomi modulo $x^r - 1$, delle matrici circolanti, dei vettori circolanti e delle combinazioni lineari di elementi nel gruppo ciclico C_r definite sul campo $\mathbb{F}$. Lo scopo di questo capitolo è dimostrare che le algebre sopra elencate sono fra di loro isomorfe.



Nel prossimo capitolo amplieremo il diagramma rappresentando $\mathbb{F}[x]/(x^r - 1)$ come prodotto di campi tramite la fattorizzazione di $x^r - 1$.

0.1 Strutture algebriche

Algebra delle matrici circolanti

Una matrice quadrata è detta circolante se, fissata la prima riga, ogni riga successiva è il risultato di una permutazione ciclica di un posto verso destra della riga precedente.

$$C = \begin{pmatrix} a_0 & a_1 & \dots & a_{r-1} \\ a_{r-1} & a_0 & \dots & a_{r-2} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_0 \end{pmatrix}$$

L'insieme delle matrici circolanti $r \times r$ a coefficienti nel campo $\mathbb{F}$ è indicato con $\mathcal{M}^c_{r,\mathbb{F}}$ e la permutazione ciclica è chiamata **shift**.

Definizione 0.1.1. *Sia A matrice $r \times r$ i cui elementi in $\mathbb{F}$ sono indicati con $(A)_{i,j} = a_{i,j}$. A è detta* **matrice circolante** *se $a_{i,j} = a_{k,l}$ quando*

$$j - i \equiv l - k \mod r$$

Sia $\mathbf{a} \in \mathbb{F}^r$ vettore corrispondente alla prima riga di una matrice circolante, allora $\mathbf{a}$ prende il nome di **vettore circolante** *e la matrice circolante completamente determinata da tale vettore viene indicata con $circ(\mathbf{a})$.*

Possiamo considerare le matrici circolanti con la struttura algebrica ereditata dalle matrici quadrate a coefficienti in $\mathbb{F}$, quindi $\mathcal{M}^c_{r,\mathbb{F}}$ ha struttura di spazio vettoriale di dimensione r su $\mathbb{F}$ essendo chiusa per la somma ed il prodotto scalare.

Ciascuna delle strutture algebriche presentate in questo capitolo possiede un elemento particolare detto **shifter**. Lo shifter delle matrici circolanti è definito da s_r come:

$$s_r = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix} \qquad (s_r)_{i,j} = \begin{cases} 1 & i \equiv j+1 \mod r \\ 0 & altrimenti \end{cases}$$

Osserviamo che $s_r = circ((0,1,0,\dots,0))$ e che l'insieme delle potenze di s_r forma un gruppo ciclico di ordine r.
Ad esempio per $r = 3$ si ha

$$s_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad s_3^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad s_3^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Quindi ogni matrice circolante è esprimibile come combinazione lineare degli elementi del gruppo delle potenze di s_r: indicata con M la generica matrice circolante, $M = circ(\mathbf{a})$, dove $\mathbf{a} = (a_0, a_1, \dots, a_{r-1})$, allora

$$\begin{aligned} M &= a_0 s_r^r + a_1 s_r^1 + \dots + a_{r-1} s_r^{r-1} \\ &= a_0 I_r + a_1 s_r^1 + \dots + a_{r-1} s_r^{r-1} \end{aligned}$$

Possiamo vedere ogni matrice circolante come un polinomio a coefficienti in $\mathbb{F}$ la cui indeterminata è s_r

$$\mathcal{M}^c_{r,\mathbb{F}} = \{\sum_{j=0}^{r-1} a_j s_r^j \mid a_j \in \mathbb{F}\}$$

Alla somma ed al prodotto scalare ereditati dalla struttura matriciale, aggiungiamo il prodotto fra matrici circolanti, che risulta ben definito e compatibile con il prodotto scalare dalla possibilità di esprimere ogni matrice circolante come un polinomio ad indeterminate appartenenti ad un gruppo ciclico. Quindi $\mathcal{M}^c_{r,\mathbb{F}}$ ha la struttura di algebra.

Algebra dei vettori circolanti

Dato lo spazio vettoriale r-dimensionale sul campo $\mathbb{F}$, i cui elementi vettoriali hanno convenzionalmente i pedici numerati da 0 ad $r-1$, vogliamo definire un prodotto che ne determini la struttura di algebra. Indichiamo con $(\mathbf{a})_i = a_i$ l'elemento dell'i-esimo posto nel vettore, allora per $\mathbf{a}$, $\mathbf{b}$ vettori, definiamo il vettore $\mathbf{a} \star \mathbf{b}$, i cui elementi sono determinati dalla sommatoria

$$(\mathbf{a} \star \mathbf{b})_i = \sum_{j \in \mathbb{Z}_r} a_j b_{i-j} \qquad \forall i \in \mathbb{Z}_r$$

che chiameremo **prodotto di convoluzione** di **a** per **b**. Possiamo anche definirlo come il prodotto del vettore **a** per la matrice circolante definita da **b**. Ad esempio per $r = 3$

$$\mathbf{a} \star \mathbf{b} = \begin{pmatrix} a_0 & a_1 & a_2 \end{pmatrix} \begin{pmatrix} b_0 & b_1 & b_2 \\ b_2 & b_0 & b_1 \\ b_1 & b_2 & b_0 \end{pmatrix}$$

In questa struttura lo **shifter** è dato dal vettore $(0, 1, 0, \ldots, 0)$, infatti si verifica immediatamente che

$$\mathbf{a} \star (0, 1, 0, \ldots, 0) = (0, 1, 0, \ldots, 0) \star \mathbf{a} = (a_{r-1}, a_0, \ldots, a_{r-2})$$

Possiamo verificare che con il prodotto di convoluzione l'insieme dei vettori non nulli formano un semigruppo, quindi l'insieme $\mathcal{V}^c_{r,\mathbb{F}}$ è un'algebra detta algebra dei vettori circolanti.

Algebra $\mathbb{F}C_r$

Consideriamo la struttura definita sul gruppo ciclico di ordine r generato da g, indicato con $C_r =< g >$, su quale agisce il campo $\mathbb{F}$. Si tratta dell'algebra i cui elementi sono i polinomi di grado inferiore ad r e la cui indeterminata è il generatore del gruppo ciclico.

$$\mathbb{F}C_r = \{a_0 + a_1 g + a_2 g^2 + \cdots + a_{r-1} g^{r-1} \mid a_i \in \mathbb{F}\}$$

Osserviamo che, rappresentando i polinomi come vettori i cui termini sono i coefficienti delle indeterminate ordinati, la struttura introdotta può essere scritta come

$$\mathbb{F}C_r = \{(a_0, a_1, \ldots, a_{r-1}) \mid a_i \in \mathbb{F}\}$$

Lo **shifter** dell'algebra $\mathbb{F}C_r$ è dato dal prodotto per g, infatti

$$\begin{aligned} g(a_0 + a_1 g + a_2 g^2 + \cdots + a_{r-1} g^{r-1}) &= a_0 g + a_1 g^2 + a_2 g^3 + \cdots + a_{r-1} g^r \\ &= a_{r-1} + a_0 g + a_1 g^2 + \cdots + a_{r-2} g^{r-1} \end{aligned}$$

E' noto che la struttura $\mathbb{F}C_r$ con il prodotto scalare, la somma ed il prodotto usuale è un'algebra[1].

Nei prossimi capitoli l'algebra $\mathbb{F}C_r$ sarà indicato con $\mathcal{A}_{r,\mathbb{F}}$ o con $\mathcal{A}$ quando non ci sono ambiguità su r ed $\mathbb{F}$.

Algebra dei polinomi modulo $x^r - 1$

Dato il campo $\mathbb{F}$ definiamo il campo campo quoziente

$$\mathcal{R}_{r,\mathbb{F}} := {}^{\mathbb{F}[x]}\Big/_{x^r - 1} = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_{r-1} x^{r-1} \mid a_j \in \mathbb{F}\}$$

i cui elementi sono polinomi di grado inferiore ad r e la cui indeterminata soddisfa la relazione

$$x^r = 1$$

[1] Ad esempio in [14] pag 408.

Moltiplicando ambo i membri della relazione precedente per x risulta evidente che l'insieme delle potenze dell'indeterminata è isomorfa al gruppo ciclico di ordine r. La struttura è un'algebra per l'usuale somma e prodotto di polinomi modulo $x^r - 1$.

Nella prossima sezione vedremo gli isomorfismi fra le algebre appena introdotte. Quando non ci saranno ambiguità su r ed $\mathbb{F}$, l'algebra dei polinomi modulo $x^r - 1$ sarà indicata con $\mathcal{R}$.

0.2 Isomorfismi

In questa sezione dimostriamo che ogni freccia del diagramma presentato all'inizio del capitolo è un isomorfismo di algebre, cominciando con il seguente lemma la cui verifica è immediata:

Lemma 0.2.1. *Ciascuna delle strutture algebriche presentate nel paragrafo precedente è isomorfa come spazio vettoriale ad* $\mathbb{F}^r$.

Inoltre ogni omomorfismo di algebre ψ_i è completamente determinato dall'immagine dello shifter. Quindi è sufficiente dimostrare che sono omomorfismi per il prodotto affinché risultino isomorfismi di algebre.

Proprietà 0.2.1. *La funzione*

$$\begin{aligned} \psi_1 : \mathcal{V}^c_{r,\mathbb{F}} &\longrightarrow \mathcal{M}^c_{r,\mathbb{F}} \\ (0,1,0,\ldots,0) &\longmapsto circ(0,1,0,\ldots,0) = s_r \end{aligned}$$

è un isomorfismo di algebre.

Dimostrazione. Siano $\mathbf{a}$ e $\mathbf{b}$ elementi di $\mathcal{V}^c_{r,\mathbb{F}}$. L'immagine del loro prodotto tramite ψ_1 è definita come

$$\begin{aligned} \psi_1(\mathbf{a} \star \mathbf{b}) &= \psi_1(\sum_{j\in\mathbb{Z}_r} a_j b_{-j}, \sum_{j\in\mathbb{Z}_r} a_j b_{1-j}, \ldots, \sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j}) \\ &= circ(\sum_{j\in\mathbb{Z}_r} a_j b_{-j}, \sum_{j\in\mathbb{Z}_r} a_j b_{1-j}, \ldots, \sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j}) \end{aligned}$$

Da cui segue che

$$\begin{aligned} \psi_1(\mathbf{a} \star \mathbf{b}) &= \begin{pmatrix} \sum_{j\in\mathbb{Z}_r} a_j b_{-j} & \sum_{j\in\mathbb{Z}_r} a_j b_{1-j} & \ldots & \sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j} \\ \sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j} & \sum_{j\in\mathbb{Z}_r} a_j b_{-j} & \ldots & \sum_{j\in\mathbb{Z}_r} a_j b_{r-2-j} \\ \vdots & & & \vdots \\ \sum_{j\in\mathbb{Z}_r} a_j b_{1-j} & \sum_{j\in\mathbb{Z}_r} a_j b_{2-j} & \ldots & \sum_{j\in\mathbb{Z}_r} a_j b_{-j} \end{pmatrix} \\ &= \begin{pmatrix} a_0 & a_1 & \ldots & a_{r-1} \\ a_{r-1} & a_0 & \ldots & a_{r-2} \\ \vdots & & & \vdots \\ a_1 & a_2 & \ldots & a_0 \end{pmatrix} \begin{pmatrix} b_0 & b_1 & \ldots & b_{r-1} \\ b_{r-1} & b_0 & \ldots & b_{r-2} \\ \vdots & & & \vdots \\ b_1 & b_2 & \ldots & b_0 \end{pmatrix} \\ &= circ(\mathbf{a})circ(\mathbf{b}) \\ &= \psi_1(\mathbf{a})\psi_1(\mathbf{b}) \end{aligned}$$

Quindi le algebre $\mathcal{V}^c_{r,\mathbb{F}}$ e $\mathcal{M}^c_{r,\mathbb{F}}$ sono isomorfe.

□

Proprietà 0.2.2. *La funzione*

$$\psi_2 : {}^{\mathbb{F}[x]}\Big/_{(x^r-1)} \longrightarrow \mathcal{V}^c_{r,\mathbb{F}}$$
$$x \longmapsto (0,1,0,\ldots,0)$$

è un isomorfismo di algebre.

Dimostrazione. Siano $a(x) = \sum_{j\in\mathbb{Z}_r} a_j x^j$ e $b(x) = \sum_{j\in\mathbb{Z}_r} b_j x^j$ elementi di $\mathcal{R}_{r,\mathbb{F}}$. Esaminiamo l'immagine del loro prodotto

$$\psi_2(a(x)b(x)) = \psi_2(a_0(\sum_{j\in\mathbb{Z}_r} b_j x^j) + a_1 x(\sum_{j\in\mathbb{Z}_r} b_j x^j) + \ldots a_{r-1}x^{r-1}(\sum_{j\in\mathbb{Z}_r} b_j x^j))$$

come visto nella sezione precedente, moltiplicare un polinomio di $\mathcal{R}_{r,\mathbb{F}}$ per x^k equivale ad effettuare uno shift sui suoi coefficienti di k posti verso destra. Quindi

$$\begin{aligned}\psi_2(a(x)b(x)) &= \psi_2(a_0(\sum_{j\in\mathbb{Z}_r} b_j x^j) + a_1(\sum_{j\in\mathbb{Z}_r} b_{j+1} x^j) + \ldots a_{r-1}(\sum_{j\in\mathbb{Z}_r} b_{j+r-1} x^j))\\ &= \psi_2(\sum_{j\in\mathbb{Z}_r} a_j b_{-j} + (\sum_{j\in\mathbb{Z}_r} a_j b_{1-j})x + \cdots + (\sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j})x^{r-1})\\ &= \psi_2(a(x)) \star \psi_2(b(x))\end{aligned}$$

Quindi le algebre $\mathcal{R}_{r,\mathbb{F}}$ e $\mathcal{V}^c_{r,\mathbb{F}}$ sono isomorfe. □

Ad esempio per $r = 3$, rappresentando i polinomi di $\mathcal{R}_{3,\mathbb{F}}$ come vettori

$$a(x) = a_0 + a_1 x + a_2 x^2 = (a_0, a_1, a_2)$$
$$b(x) = b_0 + b_1 x + b_2 x^2 = (b_0, b_1, b_2)$$

si può considerare il prodotto convolutivo di $\psi_2(a(x))$ per $\psi_2(b(x))$ come

$$\psi_2(a(x)) \star \psi_2(b(x)) = \begin{pmatrix} a_0 & a_1 & a_2 \end{pmatrix} \begin{pmatrix} b_0 & b_1 & b_2 \\ b_2 & b_0 & b_1 \\ b_1 & b_2 & b_0 \end{pmatrix}$$
$$= (a_0b_0 + a_1b_2 + a_2b_1, a_0b_1 + a_1b_0 + a_2b_2, a_0b_2 + a_1b_1 + a_2b_0)$$
$$= \psi_2(a_0b_0 + a_1b_2 + a_2b_1 + (a_0b_1 + a_1b_0 + a_2b_2)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2)$$
$$= \psi_2((a_0 + a_1x + a_2x^2)(b_0 + b_1x + b_2x^2))$$
$$= \psi_2(a(x)b(x))$$

Proseguiamo nel dimostrare che ψ_3 e ψ_4 sono isomorfismi di algebre.

Proprietà 0.2.3. *La funzione*

$$\psi_3 : \mathbb{F}C_r \longrightarrow \mathcal{M}^c_{r,\mathbb{F}}$$
$$g \longmapsto s_r$$

è un isomorfismo di algebre.

Dimostrazione. Siano a, b elementi del algebra $\mathbb{F}C_r$ della forma

$$a = a_0 + a_1 g + a_2 g^2 + \cdots + a_{r-1} g^{r-1}$$
$$b = b_0 + b_1 g + b_2 g^2 + \cdots + b_{r-1} g^{r-1}$$

Allora l'immagine del loro prodotto tramite ψ_3 è data da

$$\begin{aligned}
\psi_3(ab) &= \psi_3(\sum_{j\in\mathbb{Z}_r} a_j b_{-j} + (\sum_{j\in\mathbb{Z}_r} a_j b_{1-j})g + \cdots + (\sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j})g^{r-1}) \\
&= circ(\sum_{j\in\mathbb{Z}_r} a_j b_{-j} + (\sum_{j\in\mathbb{Z}_r} a_j b_{1-j})g + \cdots + (\sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j})g^{r-1}) \\
&= circ(a_0, a_1, a_2, \ldots, a_{r-1}) circ(b_0, b_1, b_2, \ldots, b_{r-1}) \\
&= \psi_3(a)\psi_3(b)
\end{aligned}$$

Quindi le algebre $\mathbb{F}C_r$ e $\mathcal{M}^c_{r,\mathbb{F}}$ sono isomorfe. □

Proprietà 0.2.4. *La funzione*

$$\begin{aligned}
\psi_4 : {}^{\mathbb{F}[x]}\Big/_{(x^r-1)} &\longrightarrow \mathbb{F}C_r \\
x &\longmapsto g
\end{aligned}$$

è un isomorfismo di algebre.

Dimostrazione. Siano $a(x) = \sum_{j\in\mathbb{Z}_r} a_j x^j$ e $b(x) = \sum_{j\in\mathbb{Z}_r} b_j x^j$ elementi di $\mathcal{R}_{r,\mathbb{F}}$. Esaminiamo allora l'immagine del loro prodotto tramite ψ_4:

$$\begin{aligned}
\psi_4(a(x)b(x)) &= \psi_4(a_0(\sum_{j\in\mathbb{Z}_r} b_j x^j) + a_1 x(\sum_{j\in\mathbb{Z}_r} b_j x^j) + \ldots a_{r-1} x^{r-1}(\sum_{j\in\mathbb{Z}_r} b_j x^j)) \\
&= \psi_4(a_0(\sum_{j\in\mathbb{Z}_r} b_j x^j) + a_1(\sum_{j\in\mathbb{Z}_r} b_{j+1} x^j) + \ldots a_{r-1}(\sum_{j\in\mathbb{Z}_r} b_{j+r-1} x^j)) \\
&= \sum_{j\in\mathbb{Z}_r} a_j b_{-j} + (\sum_{j\in\mathbb{Z}_r} a_j b_{1-j})g + \cdots + (\sum_{j\in\mathbb{Z}_r} a_j b_{r-1-j})g^{r-1} \\
&= \psi_4(a(x))\psi_4(b(x))
\end{aligned}$$

□

Abbiamo quindi quattro algebre isomorfe ciascuna delle quali possiede un elemento particolare, che genera il gruppo su cui le algebre sono definite e il cui prodotto con uno degli elementi della sua algebra è uno shift dei coefficienti sugli elementi della base.

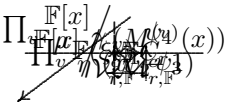
Strutture	$\mathcal{V}^c_{r,\mathbb{F}}$	$\mathcal{M}^c_{r,\mathbb{F}}$	$\mathbb{F}C_r$	$\mathcal{R}_{r,\mathbb{F}}$
Shifter	$(0, 1, 0, \ldots, 0)$	s_n	g	x

Per r ed $\mathbb{F}$ generici $\mathcal{R}_{r,\mathbb{F}}$ non è un campo, quindi non tutti i suoi elementi possiedono inverso. È possibile scomporre tale struttura in un prodotto di campi? Affrontiamo il problema nel prossimo capitolo.

Capitolo 1

Fattorizzazione di $x^r - 1$

Possiamo ampliare il diagramma del capitolo precedente con altre due strutture algebriche definite dalla fattorizzazione dell'algebra $\mathcal{R}_{r,\mathbb{F}}$ in un prodotto di campi:



Per questo scopo è necessario fattorizzare $x^r - 1$ nel prodotto di polinomi irriducibili su $\mathbb{F}$. Proponiamo inizialmente tre esempi sui quali sarà articolata la teoria:

Esempio 1.0.1. *Per $\mathbb{F} = \mathbb{Q}$ ed $r = 6$ allora*

$$\mathcal{R}_{6,\mathbb{Q}} := {}^{\mathbb{Q}[x]}\Big/_{(x^6-1)}$$

ed $x^6 - 1$ si scompone[1] *nel prodotto dei polinomi ciclotomici irriducibili:*

$$x^6 - 1 = \prod_{d|r} \Phi_d(x) = \Phi_1(x)\Phi_2(x)\Phi_3(x)\Phi_6(x)$$
$$x^6 - 1 = (x-1)(x+1)(x^2+x+1)(x^2-x-1)$$

In conseguenza del fatto che ci sono radici seste dell'unità distinte aventi lo stesso polinomio minimo, il polinomio $x^6 - 1$ non si scompone esclusivamente in polinomi di primo grado. Sia ξ_6 radice primitiva sesta:
$\xi_6^0 = 1$ è radice di $x - 1$.
$\xi_6^3 = -1$ è radice di $x + 1$.
ξ_6^2, ξ_6^4 sono radici di $x^2 + x + 1$.
ξ_6^1, ξ_6^5 sono radici di $x^2 - x - 1$.
Osserivamo che le radici di uno stesso polinomio ciclotomico espresse come potenze della radice primitiva sesta ξ_6 hanno esponenti raggruppati nelle orbite dell'azione di $\mathbb{Z}_6^\star$ su $\mathbb{Z}_6$:
$O(0) = \{0\}$.
$O(1) = \{1, 5\}$.
$O(2) = \{2, 4\}$.

[1] Vedere *ricerca dei polinomi minimi sui campi finiti* in appendice.

$O(3) = \{3\}$.
Questo fatto sarà dimostrato in generale nel corso del capitolo, così come si dimostrerà che esiste un isomorfismo fra $\mathcal{R}_{6,\mathbb{Q}}$ ed il prodotto dei quozienti sui singoli polinomi:

$$\mathbb{Q}[x] \Big/ (x^6 - 1) = \mathbb{Q}[x] \Big/ \prod_{d|r} \Phi_d(x) \cong \prod_{d|r} \mathbb{Q}[x] \Big/ \Phi_d(x)$$

Dato che i polinomi ciclotomici sono irriducibili, i fattori in cui si scompone $\mathcal{R}_{6,\mathbb{Q}}$ sono campi.

Esempio 1.0.2. *Se scegliamo invece il campo finito, $\mathbb{F} = \mathbb{Z}_2 = GF(2)$ campo di Galois di ordine 2, ed $r = 9$ allora*

$$\mathcal{R}_{9,\mathbb{Z}_2} := \mathbb{Z}_2[x] \Big/ (x^9 - 1)$$

ed $x^9 - 1$ si scompone nel prodotto di polinomi irriducibili:

$$\begin{aligned} x^9 - 1 &= M^{(0)}(x) M^{(1)}(x) M^{(3)}(x) \\ &= (x-1)(x^6 + x^3 + 1)(x^2 + x + 1) \end{aligned}$$

Sia ξ_9 radice primitiva nona:
$\xi_9^0 = 1$ è radice di $M^{(0)}(x) = x - 1$.
$\xi_9^1, \xi_9^2, \xi_6^4, \xi_9^8, \xi_9^7, \xi_9^5$ sono radici di $M^{(1)}(x) = x^6 + x^3 + 1$.
ξ_9^3, ξ_9^6 sono radici di $M^{(3)}(x) = x^2 + x + 1$.
In questo caso, gli esponenti di ξ_9 che soddisfano lo stesso polinomio ciclotomico sono raggruppate nelle orbite dell'azione di $\mathbb{Z}_9^\star$ su $\mathbb{Z}_9$:
$O(0) = \{0\}$.
$O(1) = \{1, 2, 4, 8, 7, 5\}$.
$O(3) = \{3, 6\}$.
e vale l'isomorfismo

$$\mathbb{Z}_2[x] \Big/ (x^9 - 1) \cong \prod_{v=0,1,3} \mathbb{Q}[x] \Big/ M^{(v)}(x)$$

I polinomi $M^{(v)}(x)$ della fattorizzazione dell'esempio precedente coincidono con i polinomi ciclotomici definiti dalla fattorizzazione di $x^9 - 1$ nel campo dei razionali. Possiamo scrivere impropriamente

$$M^{(0)}(x) = \Phi_1(x) \qquad M^{(1)}(x) = \Phi_9(x) \qquad M^{(3)}(x) = \Phi_3(x)$$

quindi abbiamo la stessa fattorizzazione che si avrebbe avuto scegliendo $\mathbb{F} = \mathbb{Q}$ anziché $\mathbb{F} = \mathbb{Z}_2$.
Questo non accade in generale; ad esempio per la fattorizzazione $x^7 - 1$, dove i polinomi della fattorizzazione sul campo finito sono di più di quelli nel caso della fattorizzazione sui razionali.

Esempio 1.0.3. *Sia $\mathbb{F} = \mathbb{Q}$ ed $r = 7$ allora*

$$\begin{aligned} x^7 - 1 &= \Phi_1(x)\Phi_7(x) \\ &= (x-1)(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1) \end{aligned}$$

Sia ξ_7 radice primitiva settima:
$\xi_7^0 = 1$ *è radice di* $x - 1$.
ξ_7^j *è radice di* $\Phi_7(x)$ *per* $j = 1, \ldots, 6$.
Come prima gli esponenti delle radici primitive settime che soddisfano lo stesso polinomio ciclotomico sono raggruppate nelle orbite dell'azione di $\mathbb{Z}_7^\star$ su $\mathbb{Z}_7$:
$O(0) = \{0\}$.
$O(1) = \{1, 2, 3, 4, 5, 6\}$.
Mentre se fattorizziamo $x^7 - 1$ sul campo $\mathbb{F} = \mathbb{Z}_2$, otteniamo il prodotto

$$\begin{aligned} x^7 - 1 &= M^{(0)}(x)M^{(1)}(x)M^{(3)}(x) \\ &= (x-1)(x^3 + x + 1)(x^3 + x^2 + 1) \end{aligned}$$

Le radici primitive settime si distribuiscono nel modo seguente:
$\xi_7^0 = 1$ *è radice di* $M^{(0)}(x) = x - 1$.
$\xi_7^1, \xi_7^2, \xi_7^4$ *sono radici di* $M^{(1)}(x) = x^3 + x + 1$.
$\xi_7^3, \xi_7^5, \xi_6^1$ *sono radici di* $M^{(3)}(x) = x^3 + x^2 + 1$.
A differenza del caso sui razionali gli esponenti delle radici primitive settime che soddisfano lo stesso polinomio irriducibile della fattorizzazione sono raggruppate nelle orbite dell'azione di un **sottogruppo** *di $\mathbb{Z}_7^\star$ (precisamente quello costituito dagli elementi che coincidono con gli elementi dell'orbita O(1)), su $\mathbb{Z}_7$:*
$O(0) = \{0\}$.
$O(1) = \{1, 2, 4\}$.
$O(3) = \{3, 5, 6\}$.
Quindi $\mathcal{R}_{7,\mathbb{Q}}$ si scompone nel prodotto di due campi,

$$\mathbb{Q}[x] \Big/ (x^7 - 1) \cong \mathbb{Q}[x] \Big/ \Phi_1(x) \times \mathbb{Q}[x] \Big/ \Phi_7(x)$$

mentre $\mathcal{R}_{7,\mathbb{Z}_2}$ si scompone in tre campi.

$$\mathbb{Z}_2[x] \Big/ (x^7 - 1) \cong \mathbb{Z}_2[x] \Big/ M^{(0)}(x) \times \mathbb{Z}_2[x] \Big/ M^{(1)}(x) \times \mathbb{Z}_2[x] \Big/ M^{(3)}(x)$$

È interessante notare come le algebre $\mathcal{R}_{r,\mathbb{F}}$ si decompongono in campi, in modo simile a come gli interi si fattorizzano in numeri primi. A determinare la decomposizione in campi è la fattorizzazione del polinomio $x^r - 1$ alla quale è dedicato questo capitolo.

1.1 Classi ciclotomiche

Scopo del paragrafo è scomporre il polinomio $x^r - 1$ in un prodotto di polinomi minimi. Esaminiamo il caso generale per poi proseguire nel caso in cui $x^r - 1$ è definito sul campo dei razionali e sui campi finiti. Ricordiamo inizialmente la definizione di polinomio minimo e dimostriamo la sua irriducibilità mediante il seguente lemma[2].

Lemma 1.1.1. *Sia $\mathbb{F}$ campo perfetto, ξ radice (non necessariamente primitiva!) r-esima dell'unità, $\mathbb{F}(\xi)$ estensione di $\mathbb{F}$ e sia v_ξ l'omomorfismo di valutazione:*

$$\begin{aligned} v_\xi : \mathbb{F}[x] &\longrightarrow \mathbb{F}[\xi] \\ f(x) &\longmapsto f(\xi) \end{aligned}$$

Allora valgono le seguenti proprietà:

[2]Variante del teorema 6.1.16 pag. 314 [6].

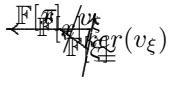
1. *$ker(v_\xi) = (p(x))$ dove $p(x)$ è un polinomio irriducibile monico detto* **polinomio minimo** *di ξ su $\mathbb{F}$.*
2. *Sia $f(x) \in \mathbb{F}[x]$ allora $f(\xi) = 0 \iff p(x) \mid f(x)$.*
3. *$\mathbb{F}[\xi] = \mathbb{F}(\xi)$ ed inoltre vale l'isomorfismo*

$$\mathbb{F}(\xi) \cong {}^{\mathbb{F}[x]}\Big/_{(p(x))}$$

4. *Se $deg(p(x)) = m$ allora $[\mathbb{F}(\xi) : \mathbb{F}] = m$ ed $\mathbb{F}(\xi)$ ha come base $\{1, \xi, \xi^2, \ldots, \xi^{m-1}\}$.*

Dimostrazione. Dimostriamo i 4 punti separatamente:

1. Dal teorema fondamentale degli isomorfismi di anelli segue che



   Quindi

$$\mathbb{F}[\xi] \cong {}^{\mathbb{F}[x]}\Big/_{ker(v_\xi)}$$

   Ma $\mathbb{F}[\xi]$ è sottoanello del campo $\mathbb{F}(\xi)$ ed è quindi dominio di integrità. Quindi lo è ${}^{\mathbb{F}[x]}\Big/_{ker(v_\xi)}$ e da una nota proprietà $ker(v_\xi)$ è un ideale primo. Quindi abbiamo che è generato da un polinomio irriducibile. Esiste allora un polinomio monico $p(x)$ (o riconducibile ad un polinomio monico dividendo per il coefficiente direttivo) tale che $ker(v_\xi) = (p(x))$.

2. Sia $f(\xi) = 0$ allora $f(x) \in ker(v_\xi)$ e viceversa. Essendo $ker(v_\xi) = (p(x))$ dal punto precedente, allora segue la tesi.

3. Dall'irriducibilità di $p(x)$ si ha che ${}^{\mathbb{F}[x]}\Big/_{(p(x))}$ è un campo e dal teorema fondamentale degli isomorfismi di anelli segue che anche $\mathbb{F}[\xi]$ è un campo. Quindi $\mathbb{F}[\xi] = \mathbb{F}(\xi)$ ed

$$\mathbb{F}(\xi) \cong {}^{\mathbb{F}[x]}\Big/_{(p(x))}$$

4. Sia m il grado di $p(x)$, allora si dimostra che $\Xi = \{1, \xi, \xi^2, \ldots, \xi^{m-1}\}$ è una base di $\mathbb{F}[\xi] = \mathbb{F}(\xi)$.
   Linearmente indipendenti: sia

$$\sum_{j=0}^{m-1} a_j \xi^j = 0 \qquad\qquad a_j \in \mathbb{F}$$

   e sia $f(x) = \sum_{j=0}^{m-1} a_j \xi^j$ polinomio associato alla combinazione lineare precedente. Dato che $f(x)$ si annulla in ξ allora appartiene al nucleo dell'omomorfismo di valutazione $ker(v_\xi)$ ed è un quindi un multiplo di $p(x)$. Dato che la differenza fra il grado di $p(x)$ ed il grado di $f(x)$ è 1 allora $f(x)$ deve essere necessariamente il polinomio nullo.

Generatori: sia $a \in \mathbb{F}(\xi)$ elemento generico allora esiste un polinomio $f(x)$ che calcolato in ξ risulta valere a:

$$a = f(\xi) = v_\xi(f(x))$$

Considerando la divisione di $f(x)$ per $p(x)$ si ottiene

$$f(x) = p(x)q(x) + r(x) \qquad 0 \leq deg(r(x)) \leq deg(p(x))$$

da cui, calcolando $f(x)$ in ξ:

$$f(\xi) = 0 + r(\xi) = a$$

e quindi $r(\xi) = a$ è una combinazione lineare ad elementi in $\mathbb{F}$ di Ξ.

□

Osservazione 1.1.1. *Se ξ è una radice* primitiva *r-esima dell'unità, allora possiamo applicare il lemma 1.1.1 precedente ad ξ^t con $t \in \mathbb{Z}_r^\star$ ed ottenere così l'estensione $\mathbb{F}(\xi^t)$ del campo $\mathbb{F}$ con base su $\mathbb{F}$ data da $\{1, \xi^t, \xi^{2t}, \ldots, \xi^{m_t - 1}\}$ per m_t grado del polinomio minimo di ξ^t.*

Nella fattorizzazione di $x^r - 1$ possono esistere delle radici distinte aventi lo stesso polinomio minimo, come affermato nel caso dei campi finiti dal seguente lemma:

Lemma 1.1.2. *Sia β elemento del campo finito $\mathbb{F}_q$ di caratteristica p e di ordine $p^n = q$, allora β e β^p hanno lo stesso polinomio minimo su $\mathbb{F}$.*

Dimostrazione. Sia $a(x) = \sum_{j=0}^{m} a_j x^j$ il polinomio minimo di β allora

$$a(\beta^p) = \sum_{j=0}^{m} a_j (\beta^p)^j = \sum_{j=0}^{m} a_j^p (\beta^j)^p = (\sum_{j=0}^{m} a_j \beta^j)^p = 0$$

□

Anche per i campi a caratteristica zero accade che radici distinte dell'unità abbiano lo stesso polinomio minimo. Per esempio in $\mathbb{Q}$ tutte le radici primitive r-esime di ordine d, dove $d \mid r$, appartengono allo stesso polinomio ciclotomico $\Phi_d(x)$ irriducibile[3]. Le radici distinte di un polinomio che hanno lo stesso polinomio minimo nella sua fattorizzazione sono dette **radici coniugate**. Esiste un modo per determinare quali sono le radici coniugate di un polinomio se osserviamo che sono legate fra di loro dagli automorfismi del gruppo di Galois dell'estensione che le contiene[4].

Teorema 1.1.1. *Sia $\mathbb{F}$ campo perfetto, ξ radice primitiva r-esima dell'unità, $Gal(\mathbb{F}(\xi), \mathbb{F})$, gruppo di Galois dell'estensione $\mathbb{F}(\xi)$ su $\mathbb{F}$, allora*

1. *Per ogni $\varphi \in Gal(\mathbb{F}(\xi), \mathbb{F})$, ξ e $\varphi(\xi)$ hanno lo stesso polinomio minimo.*
2. *Per ogni $t \in \mathbb{Z}_r$ e per ogni $\varphi \in Gal(\mathbb{F}(\xi), \mathbb{F})$, ξ^t e $\varphi(\xi^t)$ hanno lo stesso polinomio minimo.*

[3][6] pag. 133 per una presentazione generale. [19] pag. 40 per l'irriducibilità. Per ricavare il d-esimo polinomio ciclotomico usando la formula di inversione di Moebius [27] pag. 198.

[4]Variante della proposizione 7.2.3 [6] pag. 349.

Dimostrazione. È sufficiente dimostrare il secondo punto, essendo il primo un suo caso particolare per $t = 1$.
Sia $a(x) = \sum_{j=0}^{m} a_j x^j$ polinomio minimo di ξ^t allora

$$a(\varphi(\xi^t)) = \sum_{j=0}^{m} a_j(\varphi(\xi^t))^j = \sum_{j=0}^{m} \varphi(a_j)\varphi(\xi^t)^j = \sum_{j=0}^{m} \varphi(a_j)\varphi((\xi^t)^j)$$
$$= \sum_{j=0}^{m} \varphi(a_j(\xi^t)^j) = \varphi(\sum_{j=0}^{m} a_j(\xi^t)^j) = \varphi(0) = 0$$

□

Il teorema precedente ci fornisce tutte le informazioni utili per calcolare i polinomi della fattorizzazione $x^r - 1$, svelando il rapporto che intercorre fra due radici coniugate e due radici non coniugate.
Indicando con $E^{(r)} = \{\xi^j\}_{j=0}^{r-1}$ il gruppo delle radici r-esime dell'unità, riformuliamo nei prossimi due paragrafi, sul campo dei razionali e sui campi finiti, il teorema precedente identificando il generico elemento ξ^j di $E^{(r)}$ con il suo esponente j e ridefinendo $Gal(\mathbb{F}(\xi), \mathbb{F})$ come gruppo che agisce su $\mathbb{Z}_r$ invece che su $E^{(r)}$.

1.1.1 Il gruppo $Gal(\mathbb{Q}(\xi), \mathbb{Q}))$

Sia $Gal(\mathbb{Q}(\xi), \mathbb{Q}))$ per ξ radice primitiva r-esima dell'unità. L'estensione $[\mathbb{Q}(\xi), \mathbb{Q}]$ ha grado $\varphi(r)$ in conseguenza del fatto[5] che il grado dell'r-esimo polinomio ciclotomico ha come radici tutte le potenze ξ^k per $(r, k) = 1$. Dal teorema 1.1.1 ogni elemento di $Gal(\mathbb{Q}(\xi), \mathbb{Q}))$ manda ξ in un'altra radice di $\Phi_r(x)$:
per ogni $\varphi \in Gal(\mathbb{Q}(\xi), \mathbb{Q}))$ esiste k intero, $(k, r) = 1$ tale che $\varphi(\xi) = \xi^k$.
Viceversa per ogni k intero, $(k, r) = 1$ allora l'applicazione φ_k che manda ξ in ξ^k è un automorfismo.
Esiste quindi una corrispondenza fra il gruppo degli elementi invertibili di $\mathbb{Z}_n$ e gli elementi di $Gal(\mathbb{Q}(\xi), \mathbb{Q}))$.

Teorema 1.1.2. *Sia ξ radice primitiva r-esima dell'unità, allora la corrispondenza*

$$\psi : \mathbb{Z}_r^\star \longrightarrow Gal(\mathbb{Q}(\xi), \mathbb{Q}))$$
$$k \longmapsto \varphi_k : \mathbb{Q}(\xi) \longrightarrow \mathbb{Q}(\xi)$$
$$\xi \longmapsto \varphi_k(\xi) = \xi^k$$

è un isomorfismo di gruppi.

Dimostrazione. ψ è biunivoca considerando che la cardinalità dei due insiemi è $\varphi(r)$ e dato che tutte le ξ^k sono distinte per $(r, k) = 1$.
Si verifica che ψ è un omomorfismo di gruppi: siano i, j elementi di $\mathbb{Z}_r^\star$ tali che $ij \equiv k \mod r$ allora $\varphi_i\varphi_j = \varphi_k$, infatti

$$\varphi_i\varphi_j(\xi) = \varphi_i(\xi^j) = \xi^{ji} = \xi^k = \varphi_k(\xi)$$

[5] [6] pag. 133.

segue quindi

$$\mathbb{Z}_r^\star \cong Gal(\mathbb{Q}(\xi), \mathbb{Q}))$$

□

Osserviamo che gli elementi φ_k del gruppo di Galois analizzati nel precedente teorema, oltre ad agire sul sottogruppo moltiplicativo delle radici primitive r-esime dell'unità, possono anche agire su ogni altra radice r-esima. Sia ξ^l radice di ordine l, allora $\varphi_k(\xi^l)$ manda ξ^l in un'altra radice dello stesso polinomio minimo di ξ^l.
È dunque possibile concentrare l'attenzione solo sugli esponenti degli elementi di $E^{(r)} = \{\xi^j\}_{j=0}^{r-1} \cong \mathbb{Z}_r$, rappresentando l'azione delle φ_k di $Gal(\mathbb{Q}(\xi), \mathbb{Q}))$ come l'azione del gruppo $\mathbb{Z}_r^\star$ su $\mathbb{Z}_r$ in virtù del teorema precedente.
Riassumendo, l'automorfismo

$$\begin{aligned} \varphi_k : E^{(r)} &\longrightarrow E^{(r)} \\ \xi^l &\longmapsto \varphi_k(\xi^l) = \xi^{lk} \end{aligned}$$

definisce l'azione di gruppi

$$\begin{aligned} Gal(\mathbb{Q}(\xi), \mathbb{Q})) \times E^{(r)} &\longrightarrow E^{(r)} \\ (\varphi_k, \xi^l) &\longmapsto \varphi_k(\xi^l) = \xi^{lk} \end{aligned}$$

che grazie all'isomorfismo di gruppi appena ricavato ed alla corrispondenza fra gli elementi di $E^{(r)}$ e gli esponenti di ξ che lo definiscono, mantiene le stesse caratteristiche di

$$\begin{aligned} \mathbb{Z}_r^\star \times \mathbb{Z}_r &\longrightarrow \mathbb{Z}_r \\ (k, l) &\longmapsto kl \end{aligned}$$

Risulta quindi essere dimostrato il teorema:

Teorema 1.1.3. *Due elementi l_1 ed l_2 di $\mathbb{Z}_r$ sono nella stessa orbita della azione*

$$\begin{aligned} \mathbb{Z}_r^\star \times \mathbb{Z}_r &\longrightarrow \mathbb{Z}_r \\ (k, l) &\longmapsto kl \end{aligned}$$

se e solo se ξ^{l_1} e ξ^{l_2} hanno lo stesso polinomio minimo su $\mathbb{Q}$.

Definizione 1.1.1. *Le orbite dell'azione definita nel teorema precedente, ciascuna delle quali definisce l'insieme di potenze di ξ che soddisfano lo stesso polinomio minimo, sono dette* **classi ciclotomiche**. *Due elementi appartenenti ad una stessa classe ciclotomica sono detti* **coniugati**.

Negli esempi proposti all'inizio del capitolo, abbiamo osservato che il teorema 1.1.3 non può essere generalizzato per un campo qualsiasi, mantenendo l'isomorfismo $Gal(\mathbb{F}(\xi), \mathbb{F}) \cong \mathbb{Z}_r^\star$ il quale si verifica nel caso specifico dei razionali. Infatti passando da $\mathbb{Q}$ al campo $\mathbb{Z}_2$ nell'esempio 1.0.3, il gruppo $Gal(\mathbb{F}(\xi), \mathbb{F})$, da essere isomorfo a $\mathbb{Z}_r^\star$ si restringe ad un suo sottogruppo.
Purtroppo neppure tutti i campi a caratteristica zero hanno corrispondente gruppo di Galois isomorfo a $\mathbb{Z}_r^\star$, come si verifica esplorando un caso sul campo dei reali.

Esempio 1.1.1. *Sia $\mathbb{F} = \mathbb{R}$, ed r generico:*

$$\mathcal{R}_{r,\mathbb{R}} := {}^{\mathbb{R}[x]}\Big/_{(x^r-1)}$$

Sia ξ radice primitiva di $x^r - 1$, allora $Gal(\mathbb{R}(\xi), \mathbb{R})) = \{\varphi_1, \varphi_{-1}\}$ e quindi $Gal(\mathbb{R}(\xi), \mathbb{R})) \cong \mathbb{Z}_2$.

1.1.2 Il gruppo $Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$

Nel caso finito, la prima informazione utile che si può ottenere è il campo di spezzamento del polinomio $x^r - 1$ che sarà estensione di $\mathbb{F}_q$, ma prima di cominciare la sua ricerca dobbiamo fare una

Osservazione 1.1.2. *Sia $\mathbb{F}_q = GF(q)$ campo di Galois di caratteristica p e di ordine $q = p^n$. Se r e q non sono primi fra loro, cioè se esiste un $d = (p, r) \neq 1$ allora $r = \rho p^n$ con $(\rho, p) = 1$. Il polinomio $x^r - 1$ può essere scomposto come*

$$x^r - 1 = x^{\rho p^n} - 1^{p^n} = (x^\rho - 1)^{p^n}$$

quindi le radici di $x^r - 1$ coincidono con le radici di $x^\rho - 1$ con molteplicità p^n. Quindi per evitare casi ridondanti, da questo punto in poi della tesi p ed r saranno sempre considerati coprimi.

Il prossimo lemma determina il campo di spezzamento di $x^r - 1$ ricordando che il **periodo** di un elemento x di un gruppo X è il più piccolo intero positivo t tale che $x^t = 1_X$.

Lemma 1.1.3. *Sia $\mathbb{K}$ campo di spezzamento di $x^r - 1$ in $\mathbb{F}_q$, allora il grado dell'estensione di $\mathbb{K}$ su $\mathbb{F}_q$ coincide con il periodo di q in $\mathbb{Z}_r^\star$*

Dimostrazione. Dato che $\mathbb{K}$ estende $\mathbb{F}_q \supseteq \mathbb{F}_p$ la sua caratteristica è p. Sia $m = [\mathbb{K} : \mathbb{F}_q]$, cioè

$$\mathbb{F}_p \subseteq \mathbb{F}_{p^n} = \mathbb{F}_q \subseteq \mathbb{F}_{q^m} = \mathbb{K}$$

Allora, indicando con $|\cdot|$ la cardinalità di un insieme, si osserva che

1. $|\mathbb{K}^\star| = q^m - 1$, dato che $|\mathbb{K}| = q^m$ e $\mathbb{K}$ è un campo.
2. $E^{(r)} := \{\xi \in \mathbb{K}^\star \mid \xi^r = 1\}$ gruppo delle radici r-esime dell'unità è sottogruppo moltiplicativo di $\mathbb{K}^\star$ ha cardinalità r.

Dato che l'ordine dei sottogruppi deve dividere l'ordine del gruppo che lo contiene [6] allora

$$\begin{aligned} |E^{(r)}| \mid |\mathbb{K}^\star| \\ r \mid q^m - 1 \end{aligned}$$

Quindi $q^m \equiv 1 \mod r$: il periodo di q in $\mathbb{Z}_r^\star$, indicato con $per_{\mathbb{Z}_r^\star}(q)$ è uguale ad m, grado dell'estensione $[\mathbb{K} : \mathbb{F}_q]$. □

[6] Ad esempio [13] pag 43, [6] pag. 240.

Esempio 1.1.2. *Riprendendo l'esempio 1.0.3 , dove $r = 7$ e $\mathbb{F} = \mathbb{Z}_2$, il campo di spezzamento di $x^7 - 1$ è $\mathbb{F}_{2^3}$, essendo 3 il periodo di 2 in $\mathbb{Z}_7^\star$. Notiamo che il grado dell'estensione è pari alla cardinalità dell'orbita dell'azione che contiene 1.*

Sia ξ radice primitiva r-esima dell'unità sul campo finito $\mathbb{F}_q$, allora la corrispondenza

$$\begin{aligned} \psi : \mathbb{Z}_r^\star &\longrightarrow Gal(\mathbb{F}_q(\xi), \mathbb{F}_q) \\ k &\longmapsto \varphi_k : \mathbb{F}_q(\xi) \longrightarrow \mathbb{F}_q(\xi) \\ &\qquad\qquad \xi^l \longmapsto \varphi_k(\xi^l) = \xi^{kl} \end{aligned}$$

come si può verificare, non è una corrispondenza biunivoca per ogni valore di r e di q [7] . Però è possibile trovare un sottogruppo di $\mathbb{Z}_r^\star$ da sostituire al dominio in modo che ψ diventi un isomorfismo.

Teorema 1.1.4. *Sia ξ radice primitiva r-esima dell'unità sul campo finito $\mathbb{F}_q$, allora esiste un sottogruppo di $\mathbb{Z}_r^\star$ indicato con G tale che la corrispondenza*

$$\begin{aligned} \psi : G &\longrightarrow Gal(\mathbb{F}_q(\xi), \mathbb{F}_q) \\ k &\longmapsto \varphi_k \end{aligned}$$

è un isomorfismo di gruppi.

Dimostrazione. Parte 1: analisi di $Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$. Sia $\varphi \in Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$,

$$\begin{aligned} \varphi : \mathbb{F}_q(\xi) &\longrightarrow \mathbb{F}_q(\xi) \\ \xi^l &\longmapsto \varphi(\xi^l) \end{aligned}$$

Dal teorema 1.1.1 ξ^l e $\varphi(\xi^l)$ devono avere lo stesso polinomio minimo, quindi ogni elemento di $Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$ permuta le radici dei polinomi minimi. Dal lemma 1.1.3 precedente, il campo di spezzamento $\mathbb{F}_q(\xi)$ di $x^r - 1$ su $\mathbb{F}_q$ possiede q^m elementi per $m = per_{\mathbb{Z}_r^\star}(q)$. Inoltre[8]

$$|Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)| = [\mathbb{F}_q(\xi) : \mathbb{F}_q] = per_{\mathbb{Z}_r^\star}(q) = m$$

Quindi il polinomio minimo di ξ è di grado $m-1$ ed il sottogruppo di $\mathbb{Z}_r^\star$ cercato, da mettere in corrispondenza con un sottogruppo di $Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$ è di ordine m.
Parte 2: esistenza ed unicità di G. Dato che il periodo di un elemento di un gruppo deve dividere l'ordine del gruppo in cui è contenuto, esiste (ed è unico) un sottogruppo di $\mathbb{Z}_r^\star$ di ordine $per_{\mathbb{Z}_r^\star}(q) = m$.
Parte 3: ϕ è un isomorfismo. Le informazioni note a questo punto della dimostrazione sono:

- Esiste G sottogruppo di $\mathbb{Z}_r^\star$.
- $|G| = m$ periodo di q in $\mathbb{Z}_r^\star$.
- Se $g \in G$ allora $\varphi_g(\xi^t) = \xi^{gt}$ è ancora una radice del polinomio minimo di ξ^t per $t \in \mathbb{Z}_r$.

[7] Si consideri $q = 2$ ed $r = 7, 15$.
[8] [6] teorema 7.3.5 pag. 361.

Si verifica che

$$\begin{aligned} \psi : G &\longrightarrow Gal(\mathbb{F}_q(\xi), \mathbb{F}_q) \\ g &\longmapsto \varphi_g : \mathbb{F}_q(\xi) \longrightarrow \mathbb{F}_q(\xi) \\ &\qquad\qquad \xi^t \longmapsto \varphi_g(\xi^t) = \xi^{gt} \end{aligned}$$

è un isomorfismo:
L'iniettività segue dal fatto che se $\psi(g_1) = \psi(g_2)$ allora $\xi^{g_1} = \xi^{g_2}$, cioè $g_1 = g_2$ modulo r e quindi $g_1 = g_2$.
La suriettività è conseguenza del fatto che la cardinalità del dominio è uguale a quella del codominio e dell'iniettività.
È inoltre un omomorfismo di gruppi: siano $g_1, g_2 \in G$ tali che

$$g_1 g_2 \equiv k \mod r, \quad (k \in G)$$

Allora

$$\varphi_{g_1}\varphi_{g_2}(\eta) = \varphi_{g_1}(\eta^{g_2}) = \eta^{g_1 g_2} = \eta^k = \varphi_k(\eta)$$

□

Riportiamo in un corollario una delle conseguenze della dimostrazione precedente:

Corollario 1.1.1. *La cardinalità di G coincide con $m = per_{\mathbb{Z}_r^\star}(q)$.*

Analogamente a quanto visto nel caso razionale, segue che l'azione definita da

$$\begin{aligned} Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)) \times E^{(r)} &\longrightarrow E^{(r)} \\ (\varphi_k, \xi^l) &\longmapsto \varphi_k(\xi^l) = \xi^{lk} \end{aligned}$$

può essere vista come l'azione

$$\begin{aligned} G \times \mathbb{Z}_r &\longrightarrow \mathbb{Z}_r \\ (g, l) &\longmapsto gl \end{aligned}$$

semplicemente concentrandosi sugli esponenti delle radici r-esime dell'unità che determinano univocamente gli elementi di $E^{(r)}$.

Si può quindi formulare un analogo del teorema 1.1.3 sui campi finiti, la cui dimostrazione è conseguenza di quanto detto precedentemente:

Teorema 1.1.5. *Sia G definito come sopra. Due elementi l_1 ed l_2 di $\mathbb{Z}_r$ sono nella stessa orbita della azione*

$$\begin{aligned} G \times \mathbb{Z}_r &\longrightarrow \mathbb{Z}_r \\ (g, l) &\longmapsto gl \end{aligned}$$

se e solo se ξ^{l_1} e ξ^{l_2} hanno lo stesso polinomio minimo su $\mathbb{F}$.

In generale non è banale trovare il gruppo $G \cong Gal(\mathbb{F}(\xi), \mathbb{F})$ come sottogruppo di $\mathbb{Z}_r^\star$, ma da quanto visto in precedenza, sappiamo che

$$\begin{aligned} \mathbb{F} = \mathbb{Q} \quad &\Longrightarrow \quad G = \mathbb{Z}_r^\star \\ \mathbb{F} = \mathbb{F}_q \quad &\Longrightarrow \quad G \trianglelefteq \mathbb{Z}_r^\star \qquad |G| = m = per_{\mathbb{Z}_r^\star}(q) \end{aligned}$$

1.1.3 La fattorizzazione di $x^r - 1$

Questo paragrafo utilizza quanto visto fino ad ora con lo scopo di fattorizzare [9] il polinomio $x^r - 1$, noto $G \trianglelefteq \mathbb{Z}_r^\star$ isomorfo a $Gal(\mathbb{F}(\xi), \mathbb{F})$. I fattori di $x^r - 1$ sono polinomi minimi di radici r-esime coniugate i cui esponenti sono quindi nella stessa orbita dell'azione di G su $\mathbb{Z}_r$. Fattorizzare $x^r - 1$ equivale a stabilire tali orbite.
Completiamo la definizione di orbite già data in 1.1.1, nel caso dei razionali, poi dimostriamo il teorema di fattorizzazione di $x^r - 1$.

Definizione 1.1.2. *Sia $t \in \mathbb{Z}$, $Gal(\mathbb{F}(\xi), \mathbb{F}) \cong G \trianglelefteq \mathbb{Z}_r^\star$, allora si definisce $(r, \mathbb{F})$-**orbita** di t l'insieme*

$$O_{r,\mathbb{F}}(t) = O(t) = \{gt \mod r \mid g \in G\} \subseteq \mathbb{Z}_r$$

Le $(r, \mathbb{F})$-orbite sono talvolta dette classi ciclotomiche o impropriamente laterali ciclotomici. Il più piccolo elemento di ogni orbita è detto **etichetta** *dell'orbita. Indichiamo con $\mathscr{L}_{r,\mathbb{F}} = \mathscr{L}$ l'insieme delle etichette, e indichiamo con $l_{r,\mathbb{F}} = l = |\mathscr{L}|$ la sua cardinalità. Definiamo*

$$m_{r,\mathbb{F}}(t) = m(t) = |O(t)|$$

la cardinalità dell'orbita di t, che coincide con il grado del polinomio minimo di ξ^t.

Nei prossimi paragrafi capiterà di considerare vettori, o prodotti di polinomi, o di strutture algebriche i cui elementi (ordinati) hanno una corrispondenza con le etichette contenute nell'insieme $\mathscr{L}$ (non ordinato). Per evitare confusione sull'ordine degli elementi di tali vettori considereremo quando necessario gli elementi di $\mathscr{L}$ ordinati in modo crescente.
La notazione

$$(f_v)_{v \in \mathscr{L}}$$

rappresenta quindi il vettore

$$(f_0, f_1, \dots, f_t, \dots)$$

nel quale f_t non è necessariamente al t-esimo posto per t diverso da 0 e da 1, ma corrisponde alla posizione di t in $\mathscr{L}$ ordinato in modo crescente.

Teorema 1.1.6. *Sia r intero positivo ed $\mathbb{F}$ campo perfetto.*

1. *Ad ogni orbita $O(v)$ corrisponde un polinomio irriducibile in $\mathbb{F}[x]$ definito da*

$$M^{(v)}(x) = \prod_{t \in O(v)} (x - \xi^t)$$

[9] Nel 1967 è stato pubblicato un algoritmo per fattorizzare un polinomio sui campi finti detto algoritmo di Berkelamp[3]. Questo è ancora utilizzato nell'implementazione di alcuni software, come ad esempio PARI/Gp, sebbene sia stato rimpiazzato nel 1981 dall'algoritmo di Cantor-Zassenhaus [7].

2. *La decomposizione in* $\mathbb{F}$ *di* $x^r - 1$ *in fattori irriducibili è data da*

$$x^r - 1 = \prod_{v \in \mathscr{L}} M^{(v)}(x)$$

Dimostrazione. Il secondo punto è conseguenza del primo, considerando la definizione di fattorizzazione; è allora sufficiente dimostrare il secondo punto.
Dato che le orbite sono definite dall'azione del gruppo G su $\mathbb{Z}_r$ è noto che sono una partizione di $\mathbb{Z}_r$. Quindi è possibile scrivere

$$\prod_{t=0}^{r-1}(x - \xi^t) = \prod_{v \in \mathscr{L}} \prod_{t \in O(v)} (x - \xi^t) = \prod_{v \in \mathscr{L}} M^{(v)}(x) = x^r - 1$$

Dal teorema 1.1.5 abbiamo che $M^{(v)}(x)$ è il polinomio minimo che contiene tutti i coniugati di ξ^v. Quindi è il polinomio minimo di ξ^v ed è irriducibile dal lemma 1.1.1. □

Osservazione 1.1.3. *Nel caso particolare di* $\mathbb{F} = \mathbb{Q}$*, segue che* $M^{(0)}(x) = \Phi_1(x) = x - 1$*, ed* $M^{(t)}(x) = \Phi_{r/t}(x)$ *per ogni* $t \mid r$*. Infatti* $\Phi_{r/t}(x)$ *è definito come il polinomio minimo che contiene tutte le radici* r*-esime dell'unità di periodo* t*.*

Esempio 1.1.3. *Il polinomio* $x^8 - 1$ *si scompone in* $\mathbb{Q}$ *come*

$$\begin{aligned} x^8 - 1 &= (x-1)(x+1)(x^2+1)(x^4+1) \\ \Phi_1(x) &= (x-1) = M^{(0)}(x) \\ \Phi_2(x) &= (x+1) = M^{(4)}(x) \\ \Phi_4(x) &= (x^2+1) = M^{(2)}(x) \\ \Phi_8(x) &= (x^4+1) = M^{(1)}(x) \end{aligned}$$

Noti i fattori irriducibili di $x^r - 1$ possiamo costruire i suoi divisori come prodotto dei polinomi $M^{(v)}(x)$ per qualche $v \in \mathscr{L}$. Prima di approfondire questo tema vogliamo rispondere ad una domanda:
dati l'intero positivo r ed il campo perfetto $\mathbb{F}$, quante sono le orbite $O_{r,\mathbb{F}}(t)$?

1.2 Cardinalità dell'insieme di orbite

Presentiamo alcune proprietà[10] sulle orbite dell'azione di G su $\mathbb{Z}_r$ (che valgono anche per $Gal(\mathbb{F}(\xi), \mathbb{F})$ su $E^{(r)}$ grazie agli isomorfismi introdotti).

Definizione 1.2.1. *Dato* v *intero, l'orbita* $O(-v)$ *è detta* **coniugata** *ad* $O(v)$*. Se* $O(v) = O(-v)$*, allora l'orbita è detta* **autoconiugata**.

Proprietà 1.2.1. *Valgono le seguenti proprietà:*

1. $O(1) = G$ *come insieme.*
2. $m(t) \mid m(1)$ *per ogni* t *in* $\mathbb{Z}_r$*.*

[10] [8] pag. 4 e seguenti.

3. $m(t) = m(-t)$ *per ogni* t *in* $\mathbb{Z}_r$.

4. $\sum_{t \in \mathscr{L}} m(t) = r$.

Dimostrazione. Dimostriamo i vari punti dell'asserto separatamente.

1. Segue dalla definizione:

$$O(1) = \{g \mod r \mid g \in G\} = G \subseteq \mathbb{Z}_r$$

2. Ricordando la definizione di stabilizzatore dell'elemento t

$$Stab(t) = \{g \in G \mid gt = t \mod r\}$$

ed il teorema sulla cardinalità dell'orbita di t rispetto allo stabilizzatore dello stesso elemento [11]

$$|O(t)| \cdot |Stab(t)| = |G|$$

segue che

$$|O(t)| \mid |G|$$

quindi $m(1)$ è multiplo di $m(t)$.

3. La cardinalità dello stabilizzatore di t è uguale alla cardinalità dello stabilizzatore di $-t$, dal fatto che risultano essere lo stesso insieme:

$$\begin{aligned} Stab(t) &= \{g \in G \mid gt = t \mod r\} \\ &= \{g \in G \mid -gt = -t \mod r\} \\ &= \{g \in G \mid g(-t) = -t \mod r\} \\ &= Stab(-t) \end{aligned}$$

allora dal punto precedente segue che

$$|Stab(t)| \cdot m(t) = m(1) \qquad |Stab(-t)| \cdot m(-t) = m(1)$$

Quindi la cardinalità di due orbite coniugate è la stessa:

$$m(t) = \frac{m(1)}{|Stab(t)|} = \frac{m(1)}{|Stab(-t)|} = m(-t)$$

4. Le orbite sono a due a due disgiunte e la loro unione costituisce $\mathbb{Z}_r$. Formano quindi una partizione di $\mathbb{Z}_r$.

□

La seguente proprietà caratterizza le orbite autoconiugate:

Proprietà 1.2.2. *Con le notazioni precedentemente definite, sono equivalenti le seguenti condizioni:*

(a) $-1 \in G$.

[11] [6] Corollario 5.1.17 pag. 269.

(b) $O(1)$ è autoconiugata.

(c) Ogni $O(t)$ è autoconiugata.

Dimostrazione. c $\Rightarrow$ b) Se ogni orbita è autoconiugata, allora a fortiori lo è anche $O(1)$.
b $\Rightarrow$ a) Sia $O(1)$ autoconiugata, allora

$$-1 \in O(-1) = O(1)$$

Allora dalla proprietà 1.2.1 $-1 \in O(1) = G$.
a $\Rightarrow$ c) Se -1 appartiene ad G, allora per ogni t in $\mathbb{Z}_r$

$$-1t \in O(t) = \{ht \mod r \mid h \in G\}$$

Ma questo implica che $-t \in O(t)$ e quindi $O(t) = O(-t)$. □

Osservazione 1.2.1. *Dal paragrafo 1.1.1 segue che per $\mathbb{F} = \mathbb{Q}$, allora $G = \mathbb{Z}_r^\star$, quindi $-1 \in G$ e quindi ogni orbita è autoconiugata. Inoltre l'ordine di $O(t)$ definito come $m(t)$ è dato dal grado dell'r/t-esimo polinomio ciclotomico, per t divisore di r e coincide con $\varphi(r/t)$ per φ funzione di Eulero.*
Quindi la proprietà 1.2.1 verifica la nota formula $\sum_{t|r} \varphi(t) = r$.

Un'interessante applicazione del lemma di Burnside [12] si esprime nel seguente teorema sulla cardinalità dell'insieme delle orbite [13]:

Teorema 1.2.1. *Sia $\mathbb{F}$ campo perfetto ed r intero positivo coprimo con la caratteristica di $\mathbb{F}$,*

$$l_{r,\mathbb{F}} = \frac{1}{|G|} \sum_{g \in G} (g-1, r)$$

Dimostrazione. Il lemma di Burnside applicato a questo caso particolare afferma che il numero delle orbite dell'azione del gruppo G sull'insieme $\mathbb{Z}_r$ è dato da

$$l = \frac{1}{|G|} \sum_{g \in G} |X_g|$$

dove con X_g si indica l'insieme

$$X_g = \{t \in \mathbb{Z}_r \mid gt = t \mod r\}$$

La cardinalità di tale insieme equivale al numero di soluzioni della congruenza lineare $gt \equiv t \mod r$, che è proprio il massimo comune divisore fra $g-1$ e r. □

Corollario 1.2.1. *Sia $\mathbb{F} = \mathbb{Q}$ ed r intero positivo. Se $d(r)$ rappresenta il numero dei divisori di r, allora*

$$\frac{1}{\varphi(r)} \sum_{(g,r)=1} (g-1, r) = d(r)$$

[12] [1] pag. 196, [6] pag. 271.
[13] [8] pag. 4.

Dimostrazione. Per $\mathbb{F} = \mathbb{Q}$ abbiamo che $G = \mathbb{Z}_r^\star$, quindi si ha un'orbita per ogni divisore di r: $|G| = \varphi(r)$ e $g \in G$ se e solo se $(g, r) = 1$. Inoltre

$$\mathscr{L} = \{g \in \mathbb{Z}_r^\star : g \mid r\} \qquad |\mathscr{L}| = d(r)$$

da cui segue l'implicazione

$$l = \frac{1}{|G|} \sum_{g \in G} (g - 1, r) \quad \Longrightarrow \quad d(r) = \frac{1}{\varphi(r)} \sum_{(g,r)=1} (g - 1, r)$$

□

Terminiamo il paragrafo, con la dimostrazione di un lemma che sarà utile nel prossimo capitolo.

Lemma 1.2.1. *Sia $\mathbb{F}$ campo perfetto, r intero positivo e sia $O_{r,\mathbb{F}}(1)$ orbita autoconiugata, allora*

1. *Se r è dispari*

   $$m(v) = 1 \iff v = 0$$

   e per $v \neq 0$ allora $m(v)$ è pari.

2. *Se r è pari*

   $$m(v) = 1 \iff v = 0 \vee v = r/2$$

   e per $v \neq 0, r/2$ allora $m(v)$ è pari.

Dimostrazione. Dall'ipotesi $O(1)$ autoconiugato segue che ogni orbita è autoconiugata (proprietà 1.2.1), quindi per ogni t appartenente a $O(v)$ anche $-t$ appartiene ad $O(v)$.
Sia r dispari:

⇒) Per ipotesi $m(v) = 1$. Sia per assurdo $v \neq 0$ e $t \in O(v)$ per $t \in \{1, 2, \ldots, r-1\}$, allora

$$t \equiv -t \mod r \qquad 2t \equiv 0 \mod r$$

in contraddizione con r dispari.
Inoltre $m(v)$ è pari, dato che per ogni $t \neq 0$, $O(v)$ se contiene t contiene anche $-t$ che non è congruo a t modulo r.

⇐) Viceversa se $v = 0$ allora $O(v) = \{0\}$ per definizione.

Sia r pari:

⇒) Per ipotesi $m(v) = 1$. Sia per assurdo $v \neq 0$ e $t \in O(v)$ per $t \in \{1, 2, \ldots, r-1\}$, allora come prima $r \mid 2t$, che accade solo per $t = 0$ o per $t = t/2$. Analogamente a prima, per ogni altro valore di t, $O(t)$ contiene un numero pari di elementi.

⇐) Se $v = 0$ allora $m(v) = 1$ come già dimostrato, mentre per $v = r/2$ $O(v)$ contiene solo $r/2$.

□

Conoscere la fattorizzazione di $x^r - 1$ può fornire delle informazioni sull'algebra $\mathcal{R}_{r,\mathbb{F}} = \mathbb{F}[x] \Big/ (x^r - 1)$?

1.3 Fattorizzazione di $\mathcal{R}$ come prodotto di campi

Il paragrafo precedente è stato dedicato alla fattorizzazione di $x^r - 1$. Ora analizziamo una delle conseguenze di tale fattorizzazione: la possibilità di esprimere l'algebra $\mathcal{R}_{r,\mathbb{F}}$ come prodotto di campi.

Teorema 1.3.1. *Sia $\mathbb{F}$ campo perfetto ed r intero positivo. Sia*

$$x^r - 1 = \prod_{v \in \mathscr{L}} M^{(v)}(x)$$

per $M^{(v)}(x)$ polinomi irriducibili in $\mathbb{F}$. Allora vale l'isomorfismo di algebre

$$\mathbb{F}[x] \Big/ (x^r - 1) \cong \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x)$$

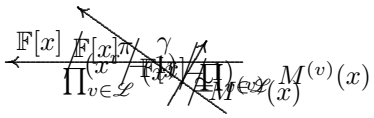
Dimostrazione. La chiave della dimostrazione consiste nel considerare la funzione γ, che definiamo come **trasformata di Winograd** e che sarà approfondita nei capitoli successivi:

$$\gamma : \mathbb{F}[x] \Big/ x^r - 1 \longrightarrow \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x)$$
$$a(x) \longmapsto (a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}}$$

È una funzione suriettiva: per il teorema cinese dei resti il sistema $a_v(x) \mod M^{(v)}(x)$ per $v \in \mathscr{L}$ ammette un'unica soluzione modulo $x^r - 1$ che è la controimmagine cercata.
Si verifica facilmente che è un isomorfismo di spazi vettoriali considerando il codominio come spazio l-dimensionale di vettori di polinomi, con il prodotto scalare e la somma usuali. Rimane da verificare che è un isomorfismo di anelli. L'ideale costituito dal nucleo di tale funzione è costituito da tutte le combinazioni di polinomi i cui addendi appartengono a $\{M^{(v)}(x)\}_{v \in \mathscr{L}}$. Quindi è definito da

$$ker(\gamma) = \{a(x) \mid a(x) \equiv 0 \mod M^{(v)}(x) \quad \forall v \in \mathscr{L}\} = (\prod_{v \in \mathscr{L}} M^{(v)}(x))$$

Osservato questo, la tesi è conseguenza del teorema fondamentale degli isomorfismi di anelli:



Da otteniamo la tesi osservando che

$$\mathbb{F}[x] \Big/ (x^r - 1) \Big/ \textstyle\prod_{v \in \mathscr{L}} M^{(v)}(x) = \mathbb{F}[x] \Big/ (x^r - 1)$$

e che π è la funzione identità. □

Osservazione 1.3.1. *Grazie all'aiuto fornito dal teorema cinese dei resti è possibile determinare un'inversa della trasformata di Winograd utilizzando la dimostrazione della suriettività nel teorema precedente. Questo tema sarà approfondita nel capitolo 3.*

Esempio 1.3.1. *Consideriamo il polinomio* $x^7 + x^2 + 1$ *in* $\mathcal{R}_{r,\mathbb{F}_q}$*, allora*

$$\gamma(x^7 + x^2 + 1) = (1, x^4 + x^2 + x + 1, 0)$$

dal fatto che

$$\begin{aligned} x^9 - 1 &= (x-1)(x^6 + x^3 + 1)(x^2 + x + 1) \\ M^{(0)}(x) &= x + 1 \\ M^{(1)}(x) &= x^6 + x^3 + 1 \\ M^{(3)}(x) &= x^2 + x + 1 \end{aligned}$$

e da

$$\begin{aligned} x^7 + x^2 + 1 &\equiv 1 \mod x + 1 \\ x^7 + x^2 + 1 &\equiv x^4 + x^2 + x + 1 \mod x^6 + x^3 + 1 \\ x^7 + x^2 + 1 &\equiv 0 \mod x^2 + x + 1 \end{aligned}$$

La controimmagine del vettore di polinomi $(1, x^4 + x^2 + x + 1, 0)$ *si calcola applicando il teorema cinese dei resti al sistema*

$$\begin{aligned} a(x) &\equiv 1 \mod x + 1 \\ a(x) &\equiv x^4 + x^2 + x + 1 \mod x^6 + x^3 + 1 \\ a(x) &\equiv 0 \mod x^2 + x + 1 \end{aligned}$$

da cui si ottiene $a(x) = x^7 + x^2 + 1$.

Il prodotto di campi definito nel teorema precedente sarà indicato, per non appesantire le notazioni con $\mathcal{Q}_{r,\mathbb{F}}$, quindi

$$\mathcal{Q}_{r,\mathbb{F}} := \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x)$$

o semplicemente con $\mathcal{Q}$ quando non ci sia bisogno di specificare r ed $\mathbb{F}$. I singoli campi che definiscono il prodotto saranno indicati con $\mathcal{Q}_{r,\mathbb{F}}^{(v)}$ per $v \in \mathscr{L}$:

$$\mathcal{Q}_{r,\mathbb{F}} = \prod_{v \in \mathscr{L}} \mathcal{Q}_{r,\mathbb{F}}^{(v)}$$

Lemma 1.3.1. *Sia* $\mathbb{F}$ *campo perfetto,* r *intero positivo e* ξ *radice primitiva* r*-esima dell'unità. Allora per* $v \in \mathscr{L}$*, l'estensione semplice* $\mathbb{F}(\xi^v)$ *è un campo che soddisfa l'isomorfismo*

$$\mathbb{F}(\xi^v) \cong \mathbb{F}[x] \Big/ M^{(v)}(x)$$

Dimostrazione. Come già osservato, il quoziente di $\mathbb{F}[x]$ su $M^{(v)}(x)$ è un campo dal fatto che $M^{(v)}(x)$ è un polinomio irriducibile. L'isomorfismo è conseguenza immediata del lemma 1.1.1 e del fatto che $M^{(v)}(x)$ è il polinomio minimo di ξ^v (e dei suoi coniugati). □

Il lemma appena concluso può rendere interessante l'oggetto definito dal prodotto dei campi $\mathbb{F}(\xi^v)$ per $v \in \mathscr{L}$. Sarà indicato, per non appesantire la notazione, con $\mathcal{P}_{r,\mathbb{F}}$ o semplicemente con $\mathcal{P}$ se non ci sono ambiguità su r e sul campo $\mathbb{F}$ sui quali è definito:

$$\mathcal{P}_{r,\mathbb{F}} := \prod_{v \in \mathscr{L}} \mathbb{F}(\xi^v)$$

I singoli campi che definiscono il prodotto saranno indicati con $\mathcal{P}_{r,\mathbb{F}}^{(v)}$ per $v \in \mathscr{L}$:

$$\mathcal{P}_{r,\mathbb{F}} = \prod_{v \in \mathscr{L}} \mathcal{P}_{r,\mathbb{F}}^{(v)}$$

Proseguiamo il capitolo con un corollario del teorema 1.3.1 e del lemma 1.3.1 che specifica la funzione μ indicata nel diagramma presentato all'inizio del capitolo.

Corollario 1.3.1. *Con le notazioni precedenti, vale la seguente catena di isomorfismi:*

$$\begin{aligned}
\mathcal{R}_{r,\mathbb{F}} &:= \mathbb{F}[x] \Big/ (x^r - 1) \\
&= \mathbb{F}[x] \Big/ \textstyle\prod_{v \in \mathscr{L}} M^{(v)}(x) \\
&\cong \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x) \\
&\cong \prod_{v \in \mathscr{L}} \mathbb{F}(\xi^v) =: \mathcal{P}_{r,\mathbb{F}}
\end{aligned}$$

Dimostrazione. Il primo isomorfismo è quello indotto da γ nel diagramma costruito dal teorema fondamentale degli isomorfismi nella dimostrazione del teorema 1.3.1. Il secondo isomorfismo è conseguenza del lemma 1.3.1: la mappa

$$\begin{aligned}
\mu : \prod_{v \in \mathscr{L}} \mathbb{F}(\xi^v) &\longrightarrow \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x) \\
(a(\xi^v))_{v \in \mathscr{L}} &\longmapsto (a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}}
\end{aligned}$$

è un isomorfismo di spazi vettoriali ed è un isomorfismo di algebre dato che ogni sua componente

$$\begin{aligned}
\mu^{(v)} : \mathbb{F}(\xi^v) &\longrightarrow \mathbb{F}[x] \Big/ M^{(v)}(x) \\
a(\xi^v) &\longmapsto a(x) \mod M^{(v)}(x)
\end{aligned}$$

è indipendente dalle altre ed è un isomorfismo di campi. □

Tornando al diagramma presentato all'inizio del capitolo, anche η è un isomorfismo di algebre perché composizione di isomorfismi:

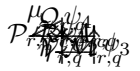
$$\begin{aligned}
\eta : \mathbb{F}[x] \Big/ (x^r - 1) &\longrightarrow \prod_{v \in \mathscr{L}} \mathbb{F}(\xi^v) \\
a(x) &\longmapsto (a(\xi^v))_{v \in \mathscr{L}}
\end{aligned}$$

Esiste una struttura algebrica soggiacente a $\mathcal{P}_{r,\mathbb{F}}$ e $\mathcal{Q}_{r,\mathbb{F}}$, sempre isomorfa ad $\mathcal{R}_{r,\mathbb{F}}$ che chiamiamo algebra dei vettori circolanti concatenati. Sarà presentata nel capitolo 3 con lo scopo di analizzare γ come trasformazione lineare nella rappresentazione vettoriale e quindi come matrice di trasformazione.

Capitolo 2

Operatori, ideali e idempotenti minimali

Fino a qui abbiamo portato avanti in parallelo il discorso della fattorizzazione di $x^r - 1$ sia sul campo dei razionali che sui campi finiti per l'analogia significativa che si svela nel calcolo dei polinomi minimi. Dato che lo scopo della tesi riguarda le applicazioni alla teoria dei codici correttori, procederemo solo sui campi di caratteristica p, tenendo presente che molti dei risultati possono essere generalizzati anche ai campi perfetti di caratteristica zero.
Ogni campo finito con $q = p^n$ elementi è definito in modo unico a meno di isomorfismi, quindi nei pedici delle algebre introdotte, non indichiamo più $\mathbb{F}_q$, ma semplicemente q. Matrici e vettori circolanti saranno indicati rispettivamente con $\mathcal{V}^c_{r,q}$ ed $\mathcal{M}^c_{r,q}$; l'algebra $\mathbb{F}_q C_r$ e l'algebra dei polinomi modulo $x^r - 1$ saranno indicati con $\mathcal{A}_{r,q}$ ed $\mathcal{R}_{r,q}$ rispettivamente e le fattorizzazioni in prodotto di campi delle algebre saranno indicati con $\mathcal{P}_{r,q}$ e $\mathcal{Q}_{r,q}$.



Il capitolo che state per leggere ha tre scopi principali: definire cinque operatori che ci permetteranno di maneggiare più agevolmente i polinomi modulo $x^r - 1$ ed analizzare alcune delle loro principali conseguenze; studiare la forma degli ideali di $\mathcal{R}$, che avranno un ruolo particolare nello studio di $\mathcal{Q}$; infine usare γ per trovare gli idempotenti di $\mathcal{R}$ ed utilizzarli come generatori dei suoi ideali.

2.1 Operatori su $\mathcal{R}$

Sia r intero positivo, $\mathbb{F}_q$ campo finito di ordine $q = p^n$ con $(r, p) = 1$, e sia $G \trianglelefteq \mathbb{Z}_r^\star$ isomorfo a $Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$ per ξ radice primitiva r-esima dell'unità.

Definizione 2.1.1. *Dato $a(x) \in \mathcal{R}_{r,q}$ si definisce* **k-shift** *di $a(x)$ il polinomio $x^k a(x)$ modulo $x^r - 1$, per $k \in \mathbb{Z}$. I k-shift sono indicati con σ_k:*

$$\begin{aligned} \sigma_k : \mathcal{R} &\longrightarrow \mathcal{R} \\ a(x) &\longmapsto x^k a(x) \end{aligned}$$

Si definisce **g-coniugato** *di $a(x)$ il polinomio $a(x^g)$ modulo $x^r - 1$ per $g \in G$, e si indica con τ_g:*

$$\begin{aligned} \tau_g : \mathcal{R} &\longrightarrow \mathcal{R} \\ a(x) &\longmapsto a(x^g) \end{aligned}$$

Come primo risultato abbiamo la

Proprietà 2.1.1. *Con le notazioni precedenti, per $a(x) = \sum_{j\in\mathbb{Z}_r} a_j x^j$*

$$\sigma_k(a(x)) = \sum_{j\in\mathbb{Z}_r} a_{j-k} x^j$$
$$\tau_g(a(x)) = \sum_{j\in\mathbb{Z}_r} a_j x^{gj}$$

inoltre per ogni g nel gruppo G, τ_g è un isomorfismo di algebre.

Dimostrazione. Le prime due equazioni sono conseguenza diretta della definizione. L'operatore τ_g è lineare ed è definito sull'automorfismo del gruppo ciclico generato da x in $\mathcal{R}$. Il fatto di elevare x per un elemento del gruppo G mantiene le eventuali radici che $a(x)$ ha in comune con $x^r - 1$ nella stessa orbita. □

Definizione 2.1.2. *Sia dato $a(x) \in \mathcal{R}$ il cui vettore circolante associato è dato da*

$$\psi_2(a(x)) = (a_0, a_1, \ldots, a_{r-1})$$

allora si definisce **riflesso** *di $a(x)$ il polinomio $a(x)^R$ il cui vettore circolante associato è definito come*

$$\psi_2(a(x)^R) = (a_{r-1}, a_{r-2}, \ldots, a_0)$$

Vale la seguente

Proprietà 2.1.2. *Con le notazioni precedenti*

$$a(x)^R = \sum_{j\in\mathbb{Z}_r} a_{r-1-j} x^j$$

Con la prossima definizione concludiamo l'elenco degli operatori su $\mathcal{R}$ presentati in questa ricerca.

Definizione 2.1.3. *Dato $a(x) \in \mathcal{R}$ si definisce* **trasposto** *di $a(x)$ il polinomio $a(x)^T$ il cui vettore circolante associato è dato da*

$$\psi_2(a(x)^T) = (a_0, a_{r-1}, a_{r-2} \ldots, a_2, a_1)$$

Si definisce **reciproco** *del polinomio $a(x)$ il polinomio*

$$a(x)^\perp = x^d a(x^{-1})$$

dove d è il grado di $a(x)$.

Il trasposto di $a(x)$ è il polinomio determinato dalla prima riga della trasposta della matrice determinata da $a(x)$. Volendo rendere le cose comode da implementare (ma meno leggibili), possiamo definire il trasposto di $a(x)$ come

$$a(x)^T = \psi_2^{-1}(\psi_1^{-1}\Big(\psi_1(\psi_2(a(x)))\Big)^t)$$

mentre il reciproco di $a(x)$ può essere definito come il polinomio riflesso, considerato però su uno spazio vettoriale di dimensione ridotta, pari al grado di $a(x)$ e poi riportato nello spazio originale di dimensione r. Per gli operatori trasposto e reciproco, abbiamo la

Proprietà 2.1.3. *Sia $a(x) \in \mathcal{R}$, allora*

$$a(x)^T = a_0 + \sum_{j\in\mathbb{Z}_r\setminus\{0\}} a_{r-j}x^j = \sum_{j\in\mathbb{Z}_r} a_{r-j}x^j$$
$$a(x)^\perp = \sum_{j\in\mathbb{Z}_r} a_{d-j}x^j$$

Dimostrazione. La prima è conseguenza immediata della definizione. Si dimostra la seconda:

$$\begin{aligned} a(x)^\perp &= x^d a(x^{-1}) = x^d \sum_{i\in\mathbb{Z}_r} a_i (x^{-1})^i \\ &= x^d \sum_{i\in\mathbb{Z}_r} a_i x^{-i} = \sum_{i\in\mathbb{Z}_r} a_i x^{d-i} \\ &= \sum_{j\in\mathbb{Z}_r} a_{d-j}x^j \end{aligned}$$

Avendo posto nell'ultimo passaggio $i = d - j$. □

Possiamo esplorare alcune relazioni fra gli operatori di $\mathcal{R}$ appena definiti:

Proprietà 2.1.4. *Sia $a(x) \in \mathcal{R}$ di grado d, allora*

$$a(x)^R = x^{-(d+1)}a(x)^\perp$$
$$a(x)^T = x^{-d}a(x)^\perp$$

Dimostrazione. La prima equazione segue da:

$$x^{-(d+1)}a(x)^\perp = x^{-1}\sum_{i\in\mathbb{Z}_r} a_i x^{r-i} = \sum_{i\in\mathbb{Z}_r} a_i x^{r-i-1} = \sum_{i\in\mathbb{Z}_r} a_{r-j-1}x^j = a(x)^R$$

Mentre la seconda segue applicando la definizione di $a(x)^\perp = x^d a(x^{-1})$:

$$x^{-d}a(x)^\perp = x^{d-d}a(x^{-1}) = a(x^{-1}) = a(x)^T$$

□

Proprietà 2.1.5. *Se $deg(a(x)) = r-1$ allora il reciproco coincide con il riflesso*

$$a(x)^\perp = a(x)^R$$

Dimostrazione.

$$a(x)^{\perp} = \sum_{j\in\mathbb{Z}_r} a_j x^{d-j} = \sum_{j\in\mathbb{Z}_r} a_j x^{r-1-j} = \\ = \sum_{j\in\mathbb{Z}_r} a_{r-j-1} x^j = a(x)^R$$

□

Se $a(x)$ è un divisore di x^r-1 (quindi un prodotto dei suoi fattori irriducibili $M^{(v)}(x)$) allora abbiamo alcune proprietà:

Proprietà 2.1.6. *Se $a(x) \in \mathcal{R}$ è un divisore di x^r-1 allora anche il suo reciproco $a(x)^{\perp}$ è un divisore di x^r-1 in $\mathcal{R}$.*

Dimostrazione. Se $a(x) = a_0 + a_1 x + \cdots + a_{d-1}x^{d-1} + a_d x^d$ allora abbiamo che, per definizione $a(x)^{\perp} = a_d + a_{d-1}x + \cdots + a_1 x^{d-1} + a_0 x^d = x^d a(1/d)$.
Indichiamo con $\hat{a}(x)$ il polinomio di grado $r-d$ dato da

$$\hat{a}(x) = (x^r-1)/a(x)$$

ben definito dal fatto che $a(x)$ è divisore di x^r-1. Allora

$$\begin{aligned} \hat{a}(x)a(x) &= x^r-1 \\ \hat{a}(1/x)a(1/x) &= (1/x)^r-1 \\ x^r\hat{a}(1/x)a(1/x) &= 1-x^r \\ x^{r-d}\hat{a}(1/x)x^d a(1/x) &= 1-x^r \\ -x^{r-d}\hat{a}(1/x)a(x)^{\perp} &= x^r-1 \end{aligned}$$

ed essendo $-x^{r-d}\hat{a}(1/x)$ polinomio di grado positivo o uguale a zero, segue che $a(x)^{\perp}$ è un divisore di x^r-1. □

La prossima proprietà[1] come la precedente caratterizza gli operatori introdotti quando agiscono sui divisori di x^r-1

Proprietà 2.1.7. *Con le notazioni introdotte si verificano i seguenti risultati:*

1. *Per ogni $g \in G$, $k \in \mathbb{Z}$ le radici che $a(x)$, $\sigma_k(a(x))$, $\tau_g(a(x))$ hanno in comune con x^r-1 coincidono fra loro.*
2. *Le radici che $a(x)^{\perp}$ ha in comune con x^r-1 coincidono con i reciproci delle radici che $a(x)$ ha in comune con x^r-1.*
3. *Le radici che $a(x)^R$, $a(x)^T$ hanno in comune con x^r-1 coincidono fra loro.*
4. *Per ogni $v \in \mathscr{L}$, $M^{(v)}(x)^{\perp} = \lambda M^{(-v)}(x)$, dove λ è il termine noto di $M^{(v)}(x)$, indicato con $M_0^{(v)}$.*
5. *Se le orbite sono autoconiugate e $v \neq 0$ allora $\lambda = 1$ ed $M^{(v)}(x)^{\perp} = M^{(-v)}(x)$.*

[1] [8] proprietà 1.11 pag 11.

Dimostrazione. Dimostriamo i diversi punti separatamente:

1. Sia $\{\xi^t \mid t \in O(v)\}$ insieme delle radici del polinomio minimo $M^{(v)}(x)$ cioè l'insieme delle radici coniugate ad ξ^v, allora

$$a(\xi^v) = 0 \iff a(\xi^{vg}) = 0 \qquad \forall g \in G$$

quindi $a(x)$ e $\tau_g(a(x))$ hanno le stesse radici in comune con $x^r - 1$.
Inoltre $\sigma_k(a(x)) = x^k a(x)$ modulo $x^r - 1$, cioè in $\mathbb{F}_q[x]$:

$$\sigma_k(a(x)) = x^k a(x) + b(x)(x^r - 1) \qquad b(x) \in \mathbb{F}_q[x]$$

Ogni ξ^t che annulla $a(x)$ ed $x^r - 1$ annulla anche $\sigma_k(a(x))$.

2. Se ξ^v è radice di $a(x)$ allora ξ^{-v} è radice di $a(x^{-1})$ ed in particolare di $x^d a(x^{-1}) = a(x)^\perp$. Viceversa se ξ^v è radice di $a(x)^\perp$ allora ξ^{-v} è radice di $a(x)$.

3. Dal fatto che

$$a(x)^R = x^{-(d+1)} a(x)^\perp = x^{-1} a(x^{-1})$$
$$a(x)^T = x^{-d} a(x)^\perp = a(x^{-1})$$

e dal punto precedente segue che almeno le radici che $a(x)^R$ ed $a(x)^T$ hanno in comune con $x^r - 1$ coincidono fra loro.

4. Le radici di $M^{(v)}(x)$ sono i reciproci delle radici di $M^{(v)}(x)^\perp$ e da quanto visto $M^{(v)}(x)^\perp$ è divisore di $x^r - 1$. Questi due polinomi hanno lo stesso grado e l'insieme delle radici di $M^{(v)}(x)^\perp$ è dato da

$$\{\xi^{-t} \mid t \in O(v)\}$$

quindi

$$M^{(v)}(x)^\perp = \lambda(\prod_{t \in O(v)} (x - \xi^{-t})) = \lambda M^{(-v)}(x)$$

Dato che $M_0^{(v)}$ è il coefficiente direttivo di $M^{(v)}(x)^\perp$, allora $\lambda = M_0^{(v)}$.

5. Se le orbite sono autoconiugate allora $M^{(v)}(x)^\perp = M_0^{(v)} M^{(v)}(x)$. Osservando che

$$M_0^{(v)} = (-1)^{m(v)} \prod_{t \in O(v)} \xi^t$$

ed applicando il lemma 1.2.1 per r dispari e $v \neq 0$, $m(v)$ è pari ed

$$M_0^{(v)} = (-1)^{m(v)} \prod \xi^t \xi^{-t} = 1$$

Mentre per r pari e $v \neq 0$ allora si considera solo il caso $v = r/2$:

$$M_0^{(r/2)} = (-1)^{m(v)} \xi^{r/2} = (-1)(-1) = 1$$

□

In considerazione della proprietà precedente risulta ragionevole distinguere le radici del generico polinomio $a(x)$ in $\mathcal{R}$ dalle radici che tale polinomio condivide con $x^r - 1$.

Definizione 2.1.4. *Sia $a(x) \in \mathcal{R}$, allora le radici che $a(x)$ possiede in comune con $x^r - 1$ sono dette* **radici principali** *(o zeri [2]).*

La distinzione fra radici e radici principali è anche conseguenza del fatto che le radici di $a(x)$ non sono in generale radici di $a(x)b(x) \mod x^r - 1$ comunque scelto $b(x)$ in $\mathcal{R}$.

2.2 Ideali di $\mathcal{R}$ e sottospazi

Gli ideali dell'algebra $\mathcal{R}$ sono generati da polinomi monici particolari ed hanno un ruolo interessante nel determinare il prodotto di campi $\mathcal{Q}$.

Teorema 2.2.1. *Sia $\mathfrak{a}$ ideale di $\mathcal{R}$, allora $\mathfrak{a}$ è principale ed è generato da un unico polinomio monico che divide $x^r - 1$.*

Dimostrazione. In generale gli anelli di polinomi a coefficienti in un campo sono ad ideali principali[3]. Riportiamo qui la dimostrazione per il caso specifico.
Sia $\mathfrak{a} \trianglelefteq \mathcal{R}$ ed $a(x)$ polinomio monico, di grado minimo, appartenente ad $\mathfrak{a}$. Dunque per definizione segue che

$$\mathfrak{a} \supseteq (a(x))$$

Verifichiamo che vale anche l'inclusione inversa (tesi intermedia: $\mathfrak{a} \subseteq (a(x))$): sia $f(x) \in \mathfrak{a}$, allora

$$f(x) = a(x)q(x) + r(x)$$

con $r(x)$ nullo o di grado strettamente inferiore al grado di $a(x)$. Inoltre $r(x) \in \mathfrak{a}$ dal fatto che

$$r(x) = f(x) - a(x)q(x)$$

Allora se $r(x)$ non fosse nullo, avremmo trovato un polinomio in $\mathfrak{a}$ di grado superiore ad $a(x)$ in contraddizione con l'ipotesi di minimalità del suo grado.
Quindi ora ogni ideale è generato da un polinomio monico. Dimostriamo che è un divisore di $x^r - 1$: con una strategia simile alla precedente consideriamo

$$x^r - 1 = a(x)q(x) + r(x)$$

dove $q(x)$ ed $r(x)$ sono diversi da prima e con $r(x)$ nullo o di grado strettamente inferiore al grado di $a(x)$. Se quozientiamo l'equazione precedente per $x^r - 1$ otteniamo

$$0 \equiv a(x)q(x) + r(x) \qquad \mod x^r - 1$$

[2] Così definite in [8] pag. 8
[3] Ad esempio [2], pag. 64.

e quindi $r(x) \in \mathfrak{a}$; ma dato che il suo grado non può essere inferiore a quello di $a(x)$, analogamente a prima abbiamo che $r(x) \equiv 0 \mod x^r - 1$ ed avendo grado minore del grado di $a(x)$ possiamo dire che è identicamente nullo. Dunque

$$x^r - 1 = a(x)q(x) + r(x)$$

cioè

$$a(x) \mid x^r - 1$$

Ora ogni ideale è generato da un polinomio monico che divide $x^r - 1$. Rimane da dimostrare che è unico. Sia $\mathfrak{a} = (a_1(x)) = (a_2(x))$ allora possiamo calcolare due polinomi $q(x)$ ed $r(x)$ tali che

$$a_1(x) = a_2(x)q(x) + r(x)(x^r - 1)$$

ma dato che per, quanto dimostrato prima, anche $a_2(x)$ è un divisore di $x^r - 1$ esiste $k(x)$ tale che

$$a_1(x) = a_2(x)q(x) + r(x)a_2(x)k(x)$$

quindi $a_1(x) \in (a_2(x))$. In modo analogo si dimostra che $a_2(x) \in (a_1(x))$ e quindi $a_1(x) = a_2(x)$. □

Riassumendo

$$\mathfrak{a} \trianglelefteq \mathcal{R} \Leftrightarrow \exists!\ a(x) \in \mathcal{R}\ \ monico, a(x) \mid x^r - 1, \mathfrak{a} = (a(x))$$

A questo punto ci poniamo due problemi: dato il polinomio $f(x)$ come possiamo stabilire qual è il più piccolo ideale $\mathfrak{a}$ in $\mathcal{R}$ che lo contiene? Dato $\mathcal{R}$, quanti sono i suoi ideali?

Corollario 2.2.1. *Sia $f(x) \in \mathfrak{a} \trianglelefteq \mathcal{R}$ allora $\mathfrak{a}$ è generato da $a(x)$ ricavato come il minimo comune multiplo fra $f(x)$ ed $x^r - 1$.*

Dimostrazione. Se $f(x) \in \mathfrak{a}$, allora per definizione di ideale $a(x) \mid f(x)$, cioè $f(x) = a(x)k(x)$ per qualche $k(x)$ in $\mathbb{F}[x]$. Inoltre, dal teorema precedente $a(x)$ è divisore di $x^r - 1$, quindi $a(x) \mid x^r - 1$.
$\Longrightarrow$ $a(x)$ è divisore comune di $f(x)$ e di $x^r - 1$, e se per assurdo non fosse il massimo dei divisori possibili avremmo una contraddizione con la sua minimalità postulata nel teorema precedente. □

Corollario 2.2.2. *L'algebra $\mathcal{R}_{r,q}$ possiede sono esattamente 2^l ideali, per l cardinalità dell'insieme delle etichette $\mathscr{L}$.*

Dimostrazione. Dal capitolo dedicato alla fattorizzazione di $x^r - 1$ sappiamo che ci sono l polinomi monici irriducibili che dividono $x^r - 1$ e dal teorema precedente sappiamo che ogni ideale è generato da un divisore monico di $x^r - 1$. Inoltre per l'ipotesi $(r, q) = 1$ i fattori di $x^r - 1$ sono tutti distinti. Dunque il generatore di un ideale è un prodotto di polinomi monici irriducibili che dividono $x^r - 1$ ciascuno dei quali può essere scelto da un insieme di l elementi. □

In un'algebra la dimensione di un ideale $\mathfrak{a}$ è la dimensione di $\mathfrak{a}$ come sottospazio vettoriale. Il prossimo corollario[4] lega la dimensione di $\mathfrak{a}$ con il grado del suo polinomio generatore.

[4]Proprietà 3.2.5 [21].

Corollario 2.2.3. *Sia* $(a(x)) = \mathfrak{a} \trianglelefteq \mathcal{R}$ *e sia* $deg(a(x)) = d = r - k$. *Allora* $dim(\mathfrak{a}) = r - d = k$.

Dimostrazione. Partiamo dall'insieme $\{\sigma_j(a(x)) = a(x)x^j\}_{j=0}^{k-1}$. Questo costituisce una base di $\mathfrak{a}$; che sia un insieme di generatori per l'ideale generato da $a(x)$ è evidente. Dimostriamo che è un insieme di elementi linearmente indipendenti:
la combinazione lineare

$$\lambda_0 a(x)x^0 + \lambda_1 a(x)x^1 + \cdots \lambda_{k-1} a(x)x^{k-1}$$

si annulla se e solo se

$$\lambda(x)a(x) = 0$$

dove $\lambda(x) = \sum_{j=0}^{k-1} \lambda_j x^j$. Il polinomio $\lambda(x)a(x)$ ha al più grado $r-1$ ed appartiene all'ideale $\mathfrak{a}$. Per questo motivo si può annullare solo se $\lambda(x)$ è identicamente nullo, cioè se e solo se $\lambda_j = 0$ per ogni j. □

Aumentando quindi il grado del generatore diminuisce la dimensione dello spazio vettoriale e viceversa.

$$deg(a(x)) + dim(\mathfrak{a}) = r$$

Vogliamo ora sfruttare il fatto che ogni ideale, in quanto sottospazio vettoriale, può essere determinato da una matrice. Cominciamo col ricordare che in generale un sottospazio U di uno spazio vettoriale E è generato dalla matrice G_U se le sue righe sono costituite dai vettori della base di E. Essa non è unica così come non è unica la scelta dei vettori che definiscono la base di U. Tuttavia, fissata una base per E ogni elemento della base di U si scrive in modo unico come combinazione lineare degli elementi della base di E.
La matrice generatrice di U, per $\{\mathbf{e}_j\}_{j=0}^{r-1}$ base di E e per $\{\mathbf{u}_j\}_{j=0}^{r-1}$ base di U, è data da

$$G_U = \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{k-1} \end{pmatrix} = \begin{pmatrix} u_{0,0} & u_{0,1} & u_{0,r-1} \\ u_{1,0} & u_{1,1} & u_{1,r-1} \\ \vdots & & \vdots \\ u_{k-1,0} & u_{k-1,1} & u_{k-1,r-1} \end{pmatrix}$$

Un generico vettore $\mathbf{a}$ di E appartiene al sottospazio U se può essere scritto come

$$\mathbf{a} = (\lambda_0, \lambda_1, \ldots, \lambda_{k-1}) \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{k-1} \end{pmatrix}$$

infatti ad ogni k-upla $(\lambda_0, \lambda_1, \ldots, \lambda_{k-1})$ corrisponde biunivocamente un vettore del sottospazio U.
Torniamo ora al nostro ideale $\mathfrak{a} \trianglelefteq \mathcal{R}$ generato dal polinomio $a(x)$; come appena visto ammette $\{a(x)x^j\}_{j=0}^{k-1}$ come sua base e dall'interludio sulle matrici generatrici dei sottospazi vettoriali possiamo dedurre immediatamente la dimostrazione del seguente

Corollario 2.2.4. *Una matrice generatrice dell'ideale* $\mathfrak{a} = (a(x)) \trianglelefteq \mathcal{R}$ *può essere scritta come*

$$G_{\mathfrak{a}} = \begin{pmatrix} a(x) \\ a(x)x^1 \\ \vdots \\ a(x)x^{k-1} \end{pmatrix} = \begin{pmatrix} a_0 & & \cdots & & a_{r-k} & 0 & \cdots & 0 \\ 0 & a_0 & & \cdots & a_{r-k-1} & a_{r-k} & 0 \cdots & 0 \\ & \vdots & & & & & \vdots & \\ 0 & \cdots & 0 & a_0 & \cdots & & \cdots & a_{r-k} \end{pmatrix}$$

che è una matrice circolante non quadrata.

Da quanto visto i mattoni fondamentali per costruire gli ideali di $\mathcal{R}$ sono i fattori irriducibili di $x^r - 1$ il cui prodotto ne determina i divisori; nella prossima definizione daremo loro un nome. Tuttavia anche il prodotto dei mattoni che vengono scartati nella costruzione di $\mathfrak{a}$ assumeranno un significato (e quindi un nome) particolare, la cui scelta sarà chiarita nel corso del capitolo sui codici lineari.

Definizione 2.2.1. *I divisori monici di* $x^r - 1$ *in* $\mathcal{R}$ *sono detti* **divisori di** $\mathcal{R}$. *Sia* $\mathfrak{a}$ *ideale generato da* $a(x)$ *determinato da un prodotto di divisori di* $\mathcal{R}$, *allora il prodotto dei divisori monici di* $x^r - 1$ *che non dividono* $a(x)$ *è un polinomio indicato come*

$$\hat{a}(x) = (x^r - 1)/a(x)$$

ed è detto **polinomio di controllo**. *Convenzionalmente se* $\mathfrak{a}$ *è generato dal polinomio* $a(x)$ *allora l'ideale generato dal polinomio* $\hat{a}(x)$ *è indicato con* $\hat{\mathfrak{a}}$.

Ogni divisore di $\mathcal{R}$ è determinato univocamente dall'insieme delle sue radici r-esime dell'unità (sottoinsieme di tutte le radici di $x^r - 1$), al quale corrisponde univocamente l'insieme dei loro esponenti.

Definizione 2.2.2. *Si definiscono* **radici principali** *dell'ideale* $\mathfrak{a}$ *le radici principali del divisore che lo genera. L'insieme degli esponenti delle radici principali di* $\mathfrak{a} = (a(x))$ *si denotano con* $Esp(\mathfrak{a})$*:*

$$Esp(\mathfrak{a}) := \{t \mid 0 \leq t \leq r-1, a(\xi^t) = 0\} \subseteq \mathbb{Z}_r$$

Chiaramente se $v \in Esp(\mathfrak{a})$ allora ogni altro intero dell'orbita $O(v)$ deve appartenere a $Esp(\mathfrak{a})$.
Presentiamo due proprietà sull'insieme $Esp(\mathfrak{a})$ appena definito.

Proprietà 2.2.1. *Sia* $a(x)$ *divisore in* $\mathcal{R}_{r,q}$, *generatore dell'ideale* $\mathfrak{a}$ *allora*

$$Esp(\mathfrak{a}) = \mathbb{Z}_r \setminus Esp(\hat{\mathfrak{a}})$$

Dimostrazione. Per definizione abbiamo che

$$a(x) = \prod_{t \in Esp(\mathfrak{a})} (x - \xi^t) \qquad \hat{a}(x) = \prod_{t \in Esp(\hat{\mathfrak{a}})} (x - \xi^t)$$

inoltre $a(x)\hat{a}(x) = x^r - 1$, da cui $Esp(\mathfrak{a})$ e $Esp(\hat{\mathfrak{a}})$ formano una partizione di $\mathbb{Z}_r$. □

Proprietà 2.2.2. *Siano* $\mathfrak{a}, \mathfrak{b}$ *ideali di* $\mathcal{R}_{r,q}$, $\mathfrak{a} \subset \mathfrak{b}$, *allora* $Esp(\mathfrak{a}) \supset Esp(\mathfrak{b})$.

Dimostrazione. Se $\mathfrak{a} = (a(x))$ e $\mathfrak{b} = (b(x))$ allora per definizione

$$(a(x)) = \{h(x) \prod_{t \in Esp(\mathfrak{a})} (x - \xi^t) \mid h(x) \in \mathcal{R}\}$$
$$(b(x)) = \{h(x) \prod_{t \in Esp(\mathfrak{b})} (x - \xi^t) \mid h(x) \in \mathcal{R}\}$$

Dato che $\mathfrak{a} \subset \mathfrak{b}$ allora esiste almeno un indice t_0 tale che $(x - \xi^{t_0})$ è un divisore di $a(x)$ ma non divide $b(x)$, quindi

$$Esp(\mathfrak{b}) \subset Esp(\mathfrak{a})$$

dato che t_0 appartiene ad $Esp(\mathfrak{a})$ ma non ad $Esp(\mathfrak{b})$. □

Osservazione 2.2.1. *Esp può dare lo spunto per definire, in parallelo alla definizione delle mappe I e V della geometria algebrica che agiscono fra insiemi algebrici affini dello spazio affine n-dimensionale e ideali dell'anello dei polinomi ad n indeterminate*[5] *altre due corrispondenze nel contesto che stiamo esaminando:*

$$I : \mathcal{P}(\mathscr{L}) \longrightarrow \{\mathfrak{a} \mid \mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}\}$$
$$A \longmapsto I(A) = (\prod_{v \in A} M^{(v)}(x)) = \mathfrak{a}_A$$
$$V : \{\mathfrak{a} \mid \mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}\} \longrightarrow \mathcal{P}(\mathscr{L})$$
$$(\prod_{v \in A} M^{(v)}(x)) = \mathfrak{a}_A \longmapsto V(\mathfrak{a}_A) = A$$

dove con $\mathcal{P}(\mathscr{L})$ intendiamo l'insieme delle parti dell'insieme delle etichette. Con queste notazioni

$$Esp(\mathfrak{a}) = \bigcup_{v \in \mathfrak{a}_A} O(v)$$

2.2.1 Ideali massimali e minimali

Ora lo scopo è quello di definire formalmente la corrispondenza biunivoca fra gli ideali ed i sottoinsiemi dell'insieme delle etichette $\mathscr{L}_{r,q}$.
Verificheremo quindi che le corrispondenze del seguente diagramma sono biiezioni e determineremo gli ideali massimali e gli ideali minimali scegliendo sottoinsiemi di $\mathscr{L}$ opportuni.

$$\{\bigcup_{t \in A} O(t) \mid A \subseteq \mathscr{L}\} \rightleftarrows \{\mathfrak{a} \mid \mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}\}$$

La prima parte della corrispondenza che vogliamo determinare è conseguenza del seguente

Teorema 2.2.2. *Per ogni sottoinsieme A dell'insieme delle etichette $\mathscr{L}$ esiste un ideale $\mathfrak{a}$ di $\mathcal{R}_{r,q}$ generato dal prodotto dei fattori irriducibili $M^{(v)}(x)$ per $v \in A$.*

[5] Presentate ad esempio in [23].

Dimostrazione. Segue dal che ogni elemento del sottoinsieme A di $\mathscr{L}$ è il rappresentante di un'orbita i cui elementi determinano le radici di un divisore irriducibile di $x^r - 1$ come esponenti della radice primitiva r-esima dell'unità.
In altre e più comprensibili parole, per ciascun elemento v di A possiamo determinare in modo unico un insieme e un polinomio fra loro correlati:

$$\{\xi^t\}_{t\in O(v)} \qquad\qquad \prod_{t\in O(v)} (x - \xi^t) = M^{(v)}(x)$$

Quindi A determina un polinomio come prodotto dei $M^{(v)}(x)$ per v in A

$$a(x) = \prod_{v\in A} M^{(v)}(x) = \prod_{v\in A} (\prod_{t\in O(v)} (x - \xi^t))$$

che è un divisore in $\mathcal{R}_{r,q}$ e, per il teorema 2.2.1, determina univocamente un ideale di $\mathfrak{a}$ di $\mathcal{R}_{r,q}$. □

La seconda parte della corrispondenza che vogliamo ottenere è data dal seguente teorema:

Teorema 2.2.3. *Per ogni ideale $\mathfrak{a} = (a(x))$ di $\mathcal{R}_{r,q}$ esiste un unico sottoinsieme dell'insieme delle etichette $\mathscr{L}$ determinato dagli indici v dei fattori irriducibili $M^{(v)}(x)$ che costituiscono il generatore $a(x)$.*

Dimostrazione. Da quanto visto nel teorema sui generatori degli ideali, $a(x)$ è il prodotto di un insieme di divisori monici irriducibili di $x^r - 1$:

$$a(x) = M^{(v_1)}(x)M^{(v_2)}(x)\cdots M^{(v_m)}(x) = \prod_{j=1}^{m} (\prod_{t\in O(v_j)} (x - \xi^t))$$

Allora l'insieme $\{v_1, \ldots, v_m\}$ costituisce un sottoinsieme di $\mathscr{L}$ unicamente determinato da $a(x)$, che è proprio l'insieme cercato. □

Per entrambe le dimostrazioni siamo passati per l'unione delle orbite determinate dalle etichette $\{\xi^t\}_{t\in O(v)}$, come indicato in modo schematico nel diagramma iniziale. Ogni ideale è unione di polinomi ciascuno dei quali corrisponde ad un'orbita rappresentata da un elemento dell'insieme delle etichette. E viceversa.

Esempio 2.2.1. *Sia $r = 9$ e $q = 2$, allora come visto in 1.0.2 sappiamo che $G = \mathbb{Z}_9^\star$ e che $\mathscr{L} = \{0, 1, 3\}$. I possibili ideali di $\mathcal{R}_{9,2}$ sono generati dai possibili prodotti dei divisori irriducibili di $x^9 - 1$, cioè da*

$$M^{(0)}(x) = x + 1$$
$$M^{(1)}(x) = x^6 + x^3 + 1$$
$$M^{(3)}(x) = x^2 + x + 1$$

Volendo elencare gli ideali definiti dal sottoinsieme A di $\mathscr{L}$ con la notazione $\mathfrak{a}_A$, si ottiene

$$\mathfrak{a}_{\{0\}} = (M^{(0)}(x)) = (x-1)$$
$$\mathfrak{a}_{\{1\}} = (M^{(1)}(x)) = (x^6+x^3+1)$$
$$\mathfrak{a}_{\{3\}} = (M^{(3)}(x)) = (x^2+x+1)$$
$$\mathfrak{a}_{\{0,1\}} = (M^{(0)}(x)M^{(1)}(x)) = (x^3-1)$$
$$\mathfrak{a}_{\{0,3\}} = (M^{(0)}(x)M^{(3)}(x)) = (x^7+x^6+x^4+x^3+x+1)$$
$$\mathfrak{a}_{\{1,3\}} = (M^{(1)}(x)M^{(3)}(x)) = (x^8+x^7+x^6+x^5+x^4+x^3+x^2+x+1)$$
$$\mathfrak{a}_{\{0,1,3\}} = (M^{(0)}(x)M^{(1)}(x)M^{(3)}(x)) = (x^9-1) = (0)$$
$$\mathfrak{a}_{\emptyset} = (1)$$

che sono, come ci aspettavamo 8 ideali, contando l'ideale improprio e l'ideale banale.

Ricordando che un ideale $\mathfrak{a}$ è massimale se fra $\mathfrak{a}$ ed (1) non ci sono ideali intermedi ed è minimale se non ci sono ideali intermedi fra $\mathfrak{a}$ ed (0), vogliamo esaminare quali sono gli ideali massimali e quali gli ideali minimali in $\mathcal{R}$. Come già osservato nel precedente esempio, scegliendo $A = \mathscr{L}$, allora l'ideale definito da A è l'ideale triviale. Se invece scegliamo $A = \emptyset$ l'ideale definito da A è l'ideale improprio.
Scegliendo invece $A = \{v\}$ allora l'ideale corrispondente è $(M^{(v)}(x))$. Inoltre $deg(M^{(v)}(x)) = |O(v)|$ ed $dim(\mathfrak{a}_{\{v\}}) = r - |O(v)|$. È facile verificare che gli ideali determinati da sottoinsiemi di $\mathscr{L}$ con un solo elemento sono ideali **massimali**, e che sono gli unici ideali massimali in $\mathcal{R}$.
Se invece $A = \mathscr{L} \setminus \{v\}$, cioè il complementare del sottoinsieme esaminato nel caso precedente, allora $deg(\hat{M}^{(v)}(x)) = r - |O(v)|$ ed $dim(\mathfrak{a}_{\mathscr{L}\setminus\{v\}}) = |O(v)|$. Si può verificare in analogia con il caso precedente, che gli ideali determinati da sottoinsiemi di $\mathscr{L}$ di questo tipo sono **minimali**, e che sono gli unici minimali in $\mathcal{R}$.
Possiamo allora aggiungere un pezzo al corollario 2.2.2 che oltre alla quantità fornisce informazioni sulla qualità degli ideali.

Corollario 2.2.5. *In $\mathcal{R}_{r,q}$ ci sono esattamente 2^l ideali, l dei quali massimali ed l dei quali minimali.*

Dimostrazione. Il sottoinsieme A di $\mathscr{L}$ può essere scelto come singleton esattamente in l modi, al quale corrispondono l possibili ideali massimali; può essere scelto come complementare di un singleton in l modi diversi al quale corrispondono l ideali minimali. □

2.2.2 Ideali ortogonali

Definiamo l'ortogonalità geometrica di due polinomi in $\mathcal{R}_{r,q}$ a partire dal prodotto scalare usuale nell'algebra dei vettori $\mathcal{V}^c_{r,q}$ che viene ereditata da $\mathcal{R}$ grazie all'isomorfismo ψ_2.

Definizione 2.2.3. *Dati $a(x)$ e $b(x)$ in $\mathcal{R}$ si definisce loro* **prodotto scalare** *l'elemento del campo $\mathbb{F}_q$ determinato da*

$$\langle a(x), b(x)\rangle = \sum_{k\in\mathbb{Z}_r} a_k b_k$$

che corrisponde all'usuale prodotto scalare dei vettori circolanti $\psi_2(a(x)), \psi_2(b(x))$ corrispondenti.

Definizione 2.2.4. *Si dice che i polinomi $a(x)$ e $b(x)$ di $\mathcal{R}$ sono geometricamente ortogonali, o* **g-ortogonali**, *se il loro prodotto scalare è nullo. Mentre si dice che sono algebricamente ortogonali, o* **a-ortogonali**, *se il loro prodotto in $\mathcal{R}$ è nullo.*

Osservazione 2.2.2. *Possiamo osservare che $a(x)$ e $b(x)$ sono g-ortogonali se*

$$\sum_{k\in\mathbb{Z}_r} a_k b_k = 0$$

Mentre sono a-ortogonali se per ogni j in $\mathbb{Z}_r$

$$\sum_{k\in\mathbb{Z}_r} a_k b_{j-k} = 0$$

Con la definizione precedente possiamo considerare i sottospazi ortogonali e gli ideali ortogonali di $\mathcal{R}$.

Definizione 2.2.5. *sia $S \subseteq \mathcal{R}$ sottoinsieme, allora si definisce* **sottospazio a-ortogonale** *generato da S l'insieme*

$$S_a^{\perp} := \{f(x) \in \mathcal{R} \mid f(x)g(x) = 0 \quad \forall g(x) \in S\}$$

e si definisce **sottospazio g-ortogonale** *generato da S l'insieme*

$$S_g^{\perp} := \{f(x) \in \mathcal{R} \mid \langle f(x), g(x)\rangle = 0 \quad \forall g(x) \in S\}$$

Se $S = \mathfrak{a}$ ideale di $\mathcal{R}$ allora $\mathfrak{a}_a^{\perp}$ è ancora un ideale, detto **ideale a-ortogonale**.

Esempio 2.2.2. *In $\mathcal{R}_{5,2}$ i polinomi x^3+1 ed x^2+x sono g-ortogonali ma non sono a-ortogonali.*

Si può trovare una correlazione fra la a-ortogonalità e la g-ortogonalità[6]:

Osservazione 2.2.3. *Dati i polinomi $a(x)$ e $b(x)$, segue, dall'osservazione 2.2.2 e dalle definizioni, che*

$$b(x^{-1}) = \sum_{k\in\mathbb{Z}_r} b_{-k}x^k = 0$$

$$a(x)b(x) = 0 \iff \forall j \in \mathbb{Z}_r \sum_{k\in\mathbb{Z}_r} a_k b_{j-k} = 0$$

$$\langle a(x), x^j b(x^{-1})\rangle = \sum_{k\in\mathbb{Z}_r} a_k b_{j-k}$$

Da cui risulta immediato verificare che $a(x)$ e $b(x)$ sono a-ortogonali se e solo se $a(x)$ è g-ortogonale a $b(x^{-1})$ e ad ogni suo shift.

[6] [8] proprietà 1.14 pag. 17.

Sapendo che tutti gli ideali in $\mathcal{R}$ sono generati da un divisore, dato $\mathfrak{a} = (a(x))$, qual è il divisore che genera $\mathfrak{a}_a^\perp$?

Lemma 2.2.1. *Sia $\mathfrak{a}$ ideale di $\mathcal{R}$ generato da $a(x)$, allora*

1. $\mathfrak{a}_a^\perp = (\hat{a}(x))$
2. $\mathfrak{a}_g^\perp$ *non è in generale un ideale.*

Dimostrazione. 1. Tutti i polinomi a-ortogonali ad $a(x)$ sono quelli il cui prodotto con $a(x)$ risulta essere 0 modulo $x^r - 1$. Il più piccolo polinomio con tale proprietà è il polinomio di controllo

$$\hat{a}(x) = (x^r - 1)/a(x)$$

e quindi $\mathfrak{a}_a^\perp = (\hat{a}(x))$.

2. $\mathfrak{a}_g^\perp$ è chiuso per la somma grazie alla linearità del prodotto scalare, ma all'esempio 2.2.2 in $\mathcal{R}_{5,2}$ per $a(x) = x^2 + x$ e $b(x) = x^3 + 1$ si ha che $b(x) \in \mathfrak{a}_g^\perp$, ma $x^2 b(x) \notin \mathfrak{a}_g^\perp$:

$$\langle x^3 + 1, x^2 + x \rangle = 0$$
$$\langle x^2 + 1, x^2 + x \rangle \neq 0$$

□

L'ideale a-ortogonale può essere semplicemente detto **ideale ortogonale**, senza che sorgano ambiguità.
La prossima definizione permette di riformulare il lemma in una proprietà nella quale si esamina una nuova relazione fra ideali ortogonali e spazi g-ortogonali.

Definizione 2.2.6. *sia $\mathfrak{a} \trianglelefteq \mathcal{R}$ allora l'ideale $\bar{\mathfrak{a}}$ le cui radici principali sono le inverse delle radici principali di $\mathfrak{a}$ è detto* **ideale coniugato**.

Osservazione 2.2.4. *Si verifica che*

$$Esp(\bar{\mathfrak{a}}) = \cup_{v \in Esp(\mathfrak{a})} O(-v)$$

Inoltre se $a(x)$ è il divisore che genera l'ideale $\mathfrak{a}$, allora $\bar{\mathfrak{a}} = (a(x^{-1}))$.

Proprietà 2.2.3. *Con le notazioni che abbiamo introdotto valgono le seguenti proprietà:*

1. *Sia $\mathfrak{a} = (a(x))$ ideale di $\mathcal{R}$, allora*

$$\mathfrak{a}_a^\perp = (\hat{a}(x)) \qquad \mathfrak{a}_a^\perp = \bar{\mathfrak{a}}_g^\perp$$

2. *Per $v \in \mathscr{L}$ segue che*

$$(M^{(v)}(x))_a^\perp = (\hat{M}^{(v)}(x)) = (M^{(-v)}(x))_g^\perp$$

Dimostrazione. Il secondo punto è conseguenza del primo, mentre per il primo punto la prima equazione è conseguenza del lemma 2.2.1. Rimane da dimostrare per doppia inclusione che

$$\mathfrak{a}_a^{\perp} = \bar{\mathfrak{a}}_g^{\perp}$$

Se $f(x) \in \mathfrak{a}_a^{\perp}$ allora $f(x)a(x) = 0$ modulo $x^r - 1$, e dall'osservazione 2.2.3 segue che per ogni j in $\mathbb{Z}_r$

$$\langle f(x), x^j a(x^{-1}) \rangle = 0$$

Ma dato che dal corollario 2.2.3 $\{a(x^{-1})x^j\}_{j=0}^{k-1}$ è un sistema di generatori di $\bar{\mathfrak{a}}$, segue che $f(x) \in \bar{\mathfrak{a}}_g^{\perp}$.
Viceversa se $f(x) \in \bar{\mathfrak{a}}_g^{\perp}$, allora segue che $f(x)$ è ortogonale ad ogni elemento della base di $\bar{\mathfrak{a}}$, quindi come prima, per ogni j in $\mathbb{Z}_r$

$$\langle f(x), x^j a(x^{-1}) \rangle = 0$$

da cui dall'osservazione 2.2.3 $f(x) \in \mathfrak{a}_a^{\perp}$. □

In questo paragrafo abbiamo potuto osservare come gli isomorfismi presentati nel capitolo 0 non siano solo un esercizio di stile. Consentono ad $\mathcal{R}_{r,q}$ di ereditare il prodotto scalare da cui l'ortogonalità geometrica[7]. Nel prossimo paragrafo continuiamo ad analizzare i polinomi nella loro rappresentazione vettoriale, passando però da $\mathcal{Q}_{r,q}$.

2.3 Elementi idempotenti

Gli elementi idempotenti giocano un ruolo particolare sia nella teoria dei codici correttori che nello studio degli ideali dell'algebra $\mathcal{R}_{r,q}$.
Ricordiamo che la fattorizzazione di $\mathcal{R}$ come prodotto di campi è indicata con

$$\mathcal{Q}_{r,q} := \prod_{v \in \mathscr{L}} {}^{\mathbb{F}_q[x]}\Big/_{(M^{(v)}(x))} = \prod_{v \in \mathscr{L}} \mathcal{Q}_{r,q}^{(v)}$$

dove il singolo campo appartenente al prodotto del secondo membro è indicato con

$$\mathcal{Q}_{r,q}^{(v)} := {}^{\mathbb{F}_q[x]}\Big/_{(M^{(v)}(x))}$$

per $v \in \mathscr{L}$.
La funzione γ, isomorfismo fra le due strutture $\mathcal{R}$ e $\mathcal{Q}$ è stata definita nel teorema 1.3.1 come

$$\begin{aligned} \gamma : \mathcal{R}_{r,q} &\longrightarrow \prod_{v \in \mathscr{L}} \mathcal{Q}_{r,q}^{(v)} \\ a(x) &\longmapsto (a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}} \end{aligned}$$

che composta con μ_1^{-1} consente di passare da $\mathcal{R}$ al prodotto di campi $\mathcal{P} = \prod_{v \in \mathscr{L}} \mathcal{P}^{(v)}$, dove $\mathcal{P}^{(v)} := \mathbb{F}_q(\xi^v)$.

[7] L'isomorfismo di $\mathcal{R}_{r,q}$ con l'algebra delle matrici circolanti consente di definire sull'algebra dei polinomi una forma r-lineare alternante che ad ogni matrice corrispondente associa il suo determinante. Una ricerca in questa direzione è stata seguita in [21] e da [29]. Altre strade possono dirigersi verso lo studio degli ideali e degli idempotenti nella rappresentazione di $\mathcal{R}_{r,q}$ come algebra di matrici circolanti e di vettori circolanti.

Definizione 2.3.1. *Un polinomio $a(x)$ appartenente all'algebra $\mathcal{R}_{r,q}$ è detto* **idempotente** *se $a(x)^2 = a(x)$.*

In generale non è un problema semplice trovare gli idempotenti in $\mathcal{R}$, senza sfruttare la sua fattorizzazione in campi. Infatti in $\mathcal{Q}$, i cui elementi sono vettori di lunghezza $m(v)$ costituiti da elementi appartenenti ai campi $\mathcal{Q}^{(v)}$, gli idempotenti sono solo i vettori che contengono 1 o 0 in ogni posizione.

Proprietà 2.3.1. *Gli idempotenti in $\mathcal{Q}$ sono tutti e soli i vettori costituiti da* 1 *e da* 0.

Dimostrazione. È conseguenza della definizione di campo: gli unici idempotenti nel campo $\mathcal{Q}^{(v)}$ sono l'unità e lo zero, quindi il prodotto di due vettori uguali di $\mathcal{Q}$ dà se stesso se e solo se è costituito da zeri e da unità □

Osservazione 2.3.1. *Si possono considerare in modo equivalente gli idempotenti in $\mathcal{P}$. Ricordando che*

$$\eta : \mathcal{R} \longrightarrow \mathcal{P} = \prod_{v \in \mathscr{L}} \mathbb{F}(\xi^v)$$
$$a(x) \longmapsto (a(\xi^v))_{v \in \mathscr{L}}$$

analogamente a prima gli unici idempotenti del campo $\mathbb{F}(\xi^v)$ sono l'unità e lo zero.

Possiamo notare una stretta correlazione fra idempotenti ed ideali, dato che ogni ideale è generato da un idempotente di $\mathcal{Q}$: in un campo gli unici idempotenti sono 0 ed 1 e gli unici ideali sono (0) ed (1), inoltre ci sono 2^l idempotenti, esattamente quanti sono gli ideali. Quali idempotenti generano ideali massimali e quali generano ideali minimali?

Definizione 2.3.2. *Il vettore l-dimensionale* $\mathbf{a}$ *di $\mathcal{Q}$ è detto* **idempotente minimale** *(o primitivo) se è costituito da un vettore di $l-1$ zeri ed ha un solo* 1.

Gli idempotenti minimali sono evidentemente idempotenti e la loro somma genera tutti gli altri idempotenti. Il "complementare" di un idempotente minimale, cioé il vettore di $\mathcal{Q}$ che contiene 1 in ogni posizione tranne in una che contiene 0 è detto **idempotente massimale**.
Indichiamo con con $\mathbf{e}_v$ il v-esimo idempotente minimale, dove v non corrisponde propriamente alla posizione dell'unità nel vettore ma è un pedice dell'insieme $\mathscr{L}$. Si indica invece con $\mathbf{e}_v(x)$ il polinomio in $\mathcal{R}$ la cui immagine tramite γ coincide con $\mathbf{e}_v$.
Dimostriamo quanto detto fino ad ora:

Teorema 2.3.1. *Sia $\mathbf{e}_i$ l'i-esimo idempotente minimale, $\mathbb{1}$ il vettore nel quale ogni elemento è un* 1 *ed $\mathbb{0}$ il vettore nel quale ogni elemento è uno zero in $\mathcal{Q}_{r,q}$. Allora valgono le seguenti proprietà*

1. *Il numero dei fattori di $x^r - 1$ in $\mathbb{F}_q$ coincide con il numero degli idempotenti primitivi e con il numero degli idempotenti massimali.*
2. *Gli idempotenti sono ortogonali: $\mathbf{e}_i \mathbf{e}_j = \mathbb{0}$ per $i \neq j$.*

3. *Gli idempotenti decompongono l'unità:* $\sum_{i\in\mathscr{L}} \mathbf{e}_i = \mathtt{1}$.

4. *Le combinazioni lineari di idempotenti generano tutti gli ideali.*

5. *Ogni elemento di* $\mathcal{Q}$ *si decompone come combinazione lineare a coefficienti in* $\mathbb{F}_q$ *degli idempotenti primitivi.*

Dimostrazione. 1. Infatti la lunghezza dei vettori di $\mathcal{Q}$ coincide con il numero dei fattori di $x^r - 1$.

2. L'ortogonalità degli idempotenti è una conseguenza della definizione di prodotto nell'algebra $\mathcal{Q}$.

3. Per definizione di somma in $\mathcal{Q}$ si ha che: $\sum_{j\in\mathscr{L}} \mathbf{e}_j = \mathtt{1}$.

4. Si verifica facilmente che $\mathbf{e}_i$ è un ideale in $\mathcal{Q}$, così come lo è la somma $\mathbf{e}_i + \mathbf{e}_j$. Per induzione si ha la tesi.

5. Sia $(q_v(x))_{v\in\mathscr{L}} \in \mathcal{Q}$, allora per la linearità dell'algebra $\mathcal{Q}$ posso scrivere

$$(q_v(x))_{v\in\mathscr{L}} = \sum_{j\in\mathscr{L}} q_v(x)\mathbf{e}_j$$

□

Corollario 2.3.1. *Ogni idempotente minimale genera un ideale minimale, ogni idempotente massimale genera un ideale massimale.*

Dimostrazione. È immediato verificare che non ci sono ideali fra $\mathbf{e}_j$ e (0) e fra il suo complementare e (1). □

Capitolo 3

Trasformata di Winograd

Nei capitoli precedenti abbiamo introdotto l'algebra dei polinomi $\mathcal{R}_{r,q}$, che dal teorema 1.3.1 risulta essere isomorfa al prodotto dei campi $\mathcal{Q}_{r,q}^{(v)}$ e dei campi $\mathcal{R}_{r,q}^{(v)}$ per $v \in \mathscr{L}$.

$$\mathcal{R}_{r,q} \cong \prod_{v \in \mathscr{L}} \mathcal{Q}_{r,q}^{(v)} \cong \prod_{v \in \mathscr{L}} \mathcal{P}_{r,q}^{(v)}$$

ricordando che

$$\begin{aligned}
\mathcal{R}_{r,q} &:= {}^{\mathbb{F}_q[x]}\Big/_{(x^r-1)} \\
\mathcal{Q}_{r,q} = \prod_{v \in \mathscr{L}} \mathcal{Q}_{r,q}^{(v)} &:= \prod_{v \in \mathscr{L}} {}^{\mathbb{F}_q[x]}\Big/_{M^{(v)}(x)} \\
\mathcal{P}_{r,q} = \prod_{v \in \mathscr{L}} \mathcal{P}_{r,q}^{(v)} &:= \prod_{v \in \mathscr{L}} \mathbb{F}_q(\xi^v)
\end{aligned}$$

e che $\mathcal{V}_{r,q}^c$ è lo spazio dei vettori circolanti.
In questo capitolo presentiamo l'isomorfismo γ studiandolo come trasformazione lineare. Dato che per questo scopo è più efficace rappresentare gli elementi di uno spazio vettoriale come r-uple e non come vettori di polinomi ciascuno dei quali considerato modulo $M^{(v)}(x)$ (come fatto fino ad ora per rappresentare $\mathcal{Q}$) costruiremo una nuova rappresentazione, nella quale gli elementi sono vettori di l vettori circolanti: ad ogni polinomio modulo $M^{(v)}(x)$ di posto corrispondente alla posizione di v in $\mathscr{L}$ faremo corrispondere un vettore di dimensione $m(v)$.
Quindi accanto al diagramma presentato nell'introduzione del secondo capitolo, considereremo i diagrammi

Con la notazione

$$\mathcal{V}_{r,q}^{\mathscr{L}} := \coprod_{v \in \mathscr{L}} \mathcal{V}_{m(v),q}^c$$

indicheremo lo spazio dei vettori circolanti concatenati, che sarà definito in questo capitolo. Ricaveremo poi le matrici di trasformazione corrispondenti a γ ed η, ed esaminando i loro blocchi, che determinano una scomposizione naturale di γ ed η in componenti.

3.1 Algebra dei vettori circolanti concatenati

Fissati r e q le immagini degli elementi della base di $\mathcal{R}_{r,q}$ tramite la trasformata di Winograd

$$\gamma : \mathbb{F}[x] \Big/ x^r - 1 \longrightarrow \prod_{v \in \mathscr{L}} \mathbb{F}[x] \Big/ M^{(v)}(x)$$
$$a(x) \longmapsto (a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}}$$

dipendono fortemente dalla fattorizzazione di $x^r - 1$. L'immagine di $1 \in \mathcal{R}$ è un vettore lungo l costituito dalle unità di ogni campo: $\gamma(1) = (1, 1, \ldots, 1)$ dato che $1 \equiv 1 \mod M^{(v)}(x)$ comunque scelto v nell'insieme delle etichette $\mathscr{L}$. Stessa cosa non si può dire per l'immagine di x, che risulta essere $x \equiv 1 \mod M^{(0)}(x)$.

Esempio 3.1.1. *Per $\mathcal{R}_{2,7}$, con $m = 3$ e con campo di spezzamento $\mathbb{F}_{2^3}$ abbiamo:*

$$\gamma(x^0) = (1, 1, 1)$$
$$\gamma(x^1) = (1, x, x)$$
$$\gamma(x^2) = (1, x^2, x^2)$$
$$\gamma(x^3) = (1, 1 + x^2, 1 + x + x^2)$$
$$\gamma(x^4) = (1, 1 + x, x + x^2)$$
$$\gamma(x^5) = (1, 1 + x + x^2, 1 + x^2)$$
$$\gamma(x^6) = (1, x + x^2, 1 + x)$$

Le immagini degli elementi della base sono quindi vettori di polinomi definiti dai loro resti per i divisori di $x^r - 1$.

Ricordando che la matrice di trasformazione risulta essere determinata dalle immagini degli elementi della base come vettori colonna, risulta più efficace considerare gli elementi di $\mathcal{Q}$ non come vettori di polinomi, ma come vettori di vettori i cui coefficienti sono quelli dei polinomi corrispondenti.
Per ottenere questa rappresentazione, senza che sia persa la struttura algebrica originale abbiamo bisogno di definire una nuova struttura.
Da due vettori circolanti $\mathbf{u} \in \mathcal{V}^c_{d_1,q}$, $\mathbf{v} \in \mathcal{V}^c_{d_2,q}$ di dimensioni differenti sullo stesso campo $\mathbb{F}_q$, è possibile definire un terzo vettore concatenandoli:

$$concat(\mathbf{u}, \mathbf{v}) = (u_0, \ldots, u_{d_1-1}, v_0, \ldots, v_{d_2-1})$$

L'insieme degli elementi di questo tipo potrebbe coincidere con $\mathcal{V}^c_{d_1+d_2,q}$ se ci limitassimo a considerarlo con le operazioni di somma e di prodotto per scalari. Volendo considerare invece anche il prodotto di convoluzione per creare una nuova struttura di algebra bisogna mantenere separati i prodotti.

Definizione 3.1.1. *Date due algebre di vettori circolanti $\mathcal{V}^c_{d_1,q}, \mathcal{V}^c_{d_2,q}$ di dimensioni diverse tali che $d_1 + d_2 = r$, si definisce* **algebra dei vettori circolanti concatenati** *la struttura*

$$\mathcal{V}^{(d_1,d_2)}_{r,q} = \mathcal{V}^c_{d_1,q} \coprod \mathcal{V}^c_{d_2,q} = \{concat(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \in \mathcal{V}^c_{d_1,q}, \mathbf{v} \in \mathcal{V}^c_{d_2,q}\}$$

considerata con la somma ed il prodotto per scalari usuali e con il prodotto di convoluzione sui singoli vettori. Indicando tale prodotto con $\star$ segue che

$$concat(\mathbf{u}_1, \mathbf{v}_1) \star concat(\mathbf{u}_2, \mathbf{v}_2) = concat(\mathbf{u}_1 \star \mathbf{u}_2, \mathbf{v}_1 \star \mathbf{v}_2)$$

Risulta immediato verificare che $\mathcal{V}_{r,q}^{(d_1,d_2)}$ è un'algebra.
L'operazione di concatenazione fra due vettori non è commutativa e può possedere un elemento neutro dato dal vettore triviale zerodimensionale $\epsilon \in \mathcal{V}_{0,q}^{c}$. Più interessante notare che è associativa: siano $\mathbf{u} \in \mathcal{V}_{d_1,q}^{c}$ $\mathbf{v} \in \mathcal{V}_{d_2,q}^{c}$ $\mathbf{w} \in \mathcal{V}_{d_3,q}^{c}$, allora

$$concat(concat(\mathbf{u}, \mathbf{v}), \mathbf{w}) = concat(\mathbf{u}, concat(\mathbf{v}, \mathbf{w})) =: concat(\mathbf{u}, \mathbf{v}, \mathbf{w})$$

Questo ci permette di definire l'operazione di concatenazione di più vettori circolanti e di costruire un'algebra concatenando più di due strutture:

Definizione 3.1.2. *Date m algebre di vettori circolanti $\mathcal{V}_{d_0,q}^{c}, \ldots, \mathcal{V}_{d_{m-1},q}^{c}$ di dimensioni diverse la cui somma sia r, si definisce* **algebra dei vettori circolanti concatenati** *la struttura*

$$\mathcal{V}_{r,q}^{(0,\ldots,m-1)} = \coprod_{j=0}^{m-1} \mathcal{V}_{d_j,q}^{c} = \{\mathbf{u} = concat(\mathbf{u}_0, \ldots, \mathbf{u}_{m-1}) \mid \mathbf{u}_j \in \mathcal{V}_{d_j,q}^{c}\}$$

*considerata con la somma ed il prodotto scalare elemento per elemento e con il prodotto di convoluzione sui singoli vettori che ne costituiscono gli elementi per concatenazione. Il vettore $\mathbf{u}_j \in \mathcal{V}_{d_j,q}$ che appartiene ad $\mathbf{u} = concat(\mathbf{u}_0, \ldots, \mathbf{u}_{m-1})$ è detto j-**esimo sottovettore** di $\mathbf{u}$.*

Per induzione si verifica facilmente che anche $\mathcal{V}_{r,q}^{(0,\ldots,m-1)}$ è un'algebra[1].
È di particolare interesse l'algebra $\mathcal{V}_{r,q}^{\mathscr{L}}$ costituita dalla concatenazione delle sottoalgebre $\mathcal{V}_{m(v),q}^{c}$. Nel prossimo teorema verificheremo che questa algebra è proprio quello che ci serve per poter ridefinire γ ed η come trasformazioni lineari fra vettori e quindi ricavarne le matrici corrispondenti.

Teorema 3.1.1. *Siano r e q fissati, $\mathscr{L}_{r,q}$ insieme delle etichette determinato dalla fattorizzazione di $x^r - 1$, allora*

1. *$\mathcal{V}_{r,q}^{\mathscr{L}}$ è un'algebra isomorfa a $\mathcal{Q}_{r,q}$.*
2. *$\mathcal{V}_{r,q}^{\mathscr{L}}$ è un'algebra isomorfa a $\mathcal{P}_{r,q}$.*

Dimostrazione. Per il primo punto la chiave della dimostrazione è la definizione di ψ_5:

$$\psi_5 : \prod_{v \in \mathscr{L}} {\mathbb{F}_q[x]}\Big/{M^{(v)}(x)} \longrightarrow \coprod_{v \in \mathscr{L}} \mathcal{V}_{m(v),q}^{c}$$
$$(a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}} \longmapsto ((a_v)_0, (a_v)_1, \ldots (a)_{m(v)-1})_{v \in \mathscr{L}}$$

[1]Sempre per induzione è possibile costruire $\mathcal{V}_{r,q}^{A}$ per A (insieme di indici delle dimensioni delle algebre di vettori circolanti concatenati) di cardinalità numerabile. Ad esempio si può considerare una variante dell'operatore della teoria dei linguaggi formali chiamato stella di Kleene applicata a questo contesto:

$$\mathcal{V}_q^{\star} = \mathcal{V}_{0,q}^{c} \coprod \mathcal{V}_{1,q}^{c} \coprod \mathcal{V}_{2,q}^{c} \coprod \cdots = \coprod_{j=0}^{\infty} \mathcal{V}_{j,q}^{c}$$

insieme dei vettori infinito-dimensionali ottenuti concatenando in successione i vettori circolanti di tutte le dimensioni possibili. A meno di un riordinamento di indici ogni algebra del tipo $\mathcal{V}_{r,q}^{A}$ è un sottoinsieme della stella di Kleene [12].

dove $(a_v)_j$ è il coefficiente di x^j del polinomio $a(x) \mod M^{(v)}(x)$.
Per la restrizione ai sottocampi ψ_5 è un isomorfismo di algebre

$$\psi_5^{(v)} : \mathcal{Q}_{r,q}^{(v)} \longrightarrow \mathcal{V}_{m(v),q}^{c}$$

e dato che il prodotto fra vettori circolanti concatenati è definito sui singoli elementi delle sottoalgebre, abbiamo che l'isomorfismo si mantiene anche per il prodotto:

$$\prod_{v \in \mathscr{L}} \mathcal{Q}_{r,q}^{(v)} \cong \coprod_{v \in \mathscr{L}} \mathcal{V}_{m(v),q}^{c}$$

Per il secondo punto potremmo ricorrere all'isomorfismo μ fra $\mathcal{Q}_{r,q}$ e $\mathcal{P}_{r,q}$, ma dato che vogliamo esplicitare l'isomorfismo fra $\mathcal{V}_{r,q}^{\mathscr{L}}$ è $\mathcal{P}_{r,q}$ evitiamo questa scorciatoia.
Definiamo ψ_6 come

$$\psi_6 : \prod_{v \in \mathscr{L}} \mathbb{F}_q(\xi^v) \longrightarrow \coprod_{v \in \mathscr{L}} \mathcal{V}_{m(v),q}^{c}$$
$$(a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}} \longmapsto ((a_v)_0, (a_v)_1, \dots (a)_{m(v)-1})_{v \in \mathscr{L}}$$

dove $(a_v)_j$ è il coefficiente di ξ^{jv} dell'elemento $a(\xi^v)$.
Come prima ψ_6 è un isomorfismo di algebre, dato che è un isomorfismo per ciascuna delle sue componenti:

$$\mathcal{P}_{r,q}^{(v)} \cong \mathcal{V}_{m(v),q}^{c}$$

e quindi abbiamo anche la tesi del secondo punto

$$\prod_{v \in \mathscr{L}} \mathcal{P}_{r,q}^{(v)} \cong \prod_{v \in \mathscr{L}} \mathcal{V}_{m(v),q}^{c}$$

□

Corollario 3.1.1. *Per ogni v etichetta di $\mathscr{L}_{r,q}$ $\mathcal{V}_{m(v),q}^{c}$ è un campo.*

Dimostrazione. Dal teorema precedente

$$\mathcal{V}_{m(v),q}^{c} \cong \mathcal{Q}_{r,q}^{(v)}$$

e $\mathcal{Q}_{r,q}^{(v)}$ è un campo per l'irriducibilità di $M^{(v)}(x)$. □

Nel teorema precedente abbiamo determinato gli isomorfismi ψ_5 e ψ_6 che agiscono entrambi fra prodotti di campi. Costituiscono semplicemente un cambiamento di notazione dove gli stessi coefficienti vengono scritti non come elementi dell'estensione di un campo ma come elementi di un vettore di vettori. Lo stesso accadeva per gli isomorfismi $\psi_j, j = 1, \dots, 4$ che riposizionavano i coefficienti in strutture diverse, senza cambiare il loro valore.
Gli isomorfismi γ ed η sono di natura diversa. Con loro passiamo da un'algebra di polinomi ad un prodotto di campi. Li riformuliamo come trasformazioni dall'algebra dei vettori circolanti $\mathcal{V}_{r,q}^{c}$ all'algebra dei vettori circolanti concatenati $\mathcal{V}_{r,q}^{\mathscr{L}}$.

Corollario 3.1.2. *Il prodotto di campi $\mathcal{V}_{r,q}^{\mathscr{L}}$ è un'algebra isomorfa a $\mathcal{V}_{r,q}^{c}$.*

Dimostrazione. Affrontiamo la dimostrazione in due modi diversi con lo scopo di definire isomorfismi δ e ρ analoghi di γ ed η.

1. Definiamo δ come composizione di 3 isomorfismi di algebre:

$$\begin{array}{ccc} \mathcal{R}_{r,q} & \xrightarrow{\quad\gamma\quad} & \mathcal{Q}_{r,q} \\ \Big\downarrow{\scriptstyle \psi_2} & & \Big\downarrow{\scriptstyle \psi_5} \\ \mathcal{V}_{r,q}^{c} & \xrightarrow{\quad\delta\quad} & \mathcal{V}_{r,q}^{\mathscr{L}} \end{array}$$

   Quindi $\delta = \psi_5 \circ \gamma \circ \psi_2^{-1}$ è un isomorfismo perché composizione di isomorfismi.

2. Definiamo ρ come composizione di 3 isomorfismi di algebre:

$$\begin{array}{ccc} \mathcal{R}_{r,q} & \xrightarrow{\quad\eta\quad} & \mathcal{P}_{r,q} \\ \Big\downarrow{\scriptstyle \psi_2} & & \Big\downarrow{\scriptstyle \psi_6} \\ \mathcal{V}_{r,q}^{c} & \xrightarrow{\quad\rho\quad} & \mathcal{V}_{r,q}^{\mathscr{L}} \end{array}$$

   Quindi $\rho = \psi_6 \circ \eta \circ \psi_2^{-1}$ è un isomorfismo perché composizione di isomorfismi.

□

Abbiamo due classi di algebre: la classe delle algebre scomposte in campi[2] il cui rappresentante privilegiato per questo capitolo è $\mathcal{V}_{r,q}^{\mathscr{L}}$ e la classe delle algebre non scomposte in campi[3] il cui rappresentante privilegiato è $\mathcal{V}_{r,q}^{c}$.
Per concentrare l'attenzione solo sugli isomorfismi γ ed η e per rendere la notazione meno pesante ometteremo gli altri ψ_j per $j = 1, \ldots, 6$.
Ad esempio per $a(x) = 1 + 2x^2 \in \mathcal{R}_{3,5}$ [4] scriveremo

$$(1, 0, 2) = circ((1, 0, 2)) = 1 + 2g^2 = 1 + 2x^2$$

e per $\gamma(a(x)) = (3, 4 + 2x)$ elemento di $\mathcal{Q}_{3,5}$ scriveremo

$$(3, 4, 2) = (3, 4 + 2\xi) = (3, 4 + 2x)$$

Abbiamo tutti gli ingredienti per definire formalmente la trasformata di Winograd:

[2] Alla quale appartengono $\mathcal{V}_{r,q}^{\mathscr{L}}, \mathcal{P}_{r,q}, \mathcal{Q}_{r,q}$.
[3] Alla quale appartengono $\mathcal{V}_{r,q}^{c}, \mathcal{M}_{r,q}^{c}, \mathbb{F}_q C_r, \mathcal{R}_{r,q}$.
[4] $x^3 - 1 = (x + 1)(x^2 + 4x + 1)$ in $\mathbb{F}_5$.

Definizione 3.1.3. *La matrice della trasformazione δ fra le algebre $\mathcal{R}_{r,q}$ e $\mathcal{Q}_{r,q}$ nelle rispettive rappresentazioni vettoriali $\mathcal{V}^{c}_{r,q}$ e $\mathcal{V}^{\mathscr{L}}_{r,q}$ è detta* **matrice di Winograd** *ed è indicata con Γ.*
Mentre la matrice della trasformazione ρ fra le algebre $\mathcal{R}_{r,q}$ e $\mathcal{P}_{r,q}$ nelle rispettive rappresentazioni vettoriali $\mathcal{V}^{c}_{r,q}$ e $\mathcal{V}^{\mathscr{L}}_{r,q}$ è indicata con H.

Esempio 3.1.2. *Tornando all'esempio 3.1.1, con la nuova struttura dei vettori circolanti concatenati possiamo costruire la matrice di trasformazione. Abbiamo:*

$$\begin{aligned}
\gamma(x^0) &= (1,1,1) = (1,1,0,0,1,0,0)\\
\gamma(x^1) &= (1,x,x) = (1,0,1,0,0,1,0)\\
\gamma(x^2) &= (1,x^2,x^2) = (1,0,0,1,0,0,1)\\
\gamma(x^3) &= (1,1+x^2,1+x+x^2) = (1,1,0,1,1,1,1)\\
\gamma(x^4) &= (1,1+x,x+x^2) = (1,1,1,0,0,1,1)\\
\gamma(x^5) &= (1,1+x+x^2,1+x^2) = (1,1,1,1,1,0,1)\\
\gamma(x^6) &= (1,x+x^2,1+x) = (1,0,1,1,1,1,0)
\end{aligned}$$

da cui segue che la matrice Γ è definita come:

$$\Gamma = \begin{pmatrix} \Gamma^{(0)} \\ \\ \Gamma^{(1)} \\ \\ \Gamma^{(3)} \end{pmatrix} = \left(\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{array}\right)$$

nella quale abbiamo mantenuto la suddivisione in blocchi ereditata dalla struttura $\mathcal{Q}$. Possiamo inoltre osservare che è una matrice circolante a blocchi.
La sua inversa, i cui dettagli saranno visti nei prossimi paragrafi, è data da

$$\Delta = \left(\begin{array}{c|ccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Osserviamo che la dimensione dei blocchi $\Gamma^{(i)}$ è data dai gradi dei divisori $M^{(v)}(x)$ cosa che si accorda con il fatto che che la somma dei gradi è uguale ad r.

Esempio 3.1.3. *Sia $a(x) = x^5 + x^3 + 1 = (1,0,0,1,0,1,0)$ elemento di $\mathcal{R}_{2,7}$. La sua immagine mediante la trasformata di Winograd è data da $\gamma(a(x)) =$*

$\Gamma(a(x))^t$ *per* $(a(x))^t$ *trasposto del vettore* $a(x)$*:*

$$\gamma(1,0,0,1,0,1,0) = \left(\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{array}\right) \left(\begin{array}{c} 1 \\ \hline 0 \\ 0 \\ 1 \\ \hline 0 \\ 1 \\ 0 \end{array}\right) = (1,1,1,0,1,1,0)$$

Quindi $\gamma(a(x)) = (1,1,1,0,1,1,0)$.

3.2 Scomposizione di γ e di η

La suddivisione in blocchi della matrice γ suggerisce una naturale suddivisione dell'isomorfismo γ in l epimorfismi, che può fornire un modo per calcolare i singoli blocchi separatamente:

$$\begin{aligned} \gamma^{(v)} : \mathcal{R}_{r,q} &\longrightarrow \mathcal{Q}_{r,q}^{(v)} \\ a(x) &\longmapsto a(x) \mod M^{(v)}(x) \end{aligned}$$

o considerando la rappresentazione sulle algebre dei vettori circolanti e circolanti concatenati

$$\begin{aligned} \delta^{(v)} : \mathcal{V}_{r,q}^{c} &\longrightarrow \mathcal{V}_{m(v),q}^{c} \\ (a_0,\ldots,a_{r-1}) &\longmapsto \psi_5(\gamma(\psi_2^{-1}(a_0,\ldots,a_{r-1}))) \end{aligned}$$

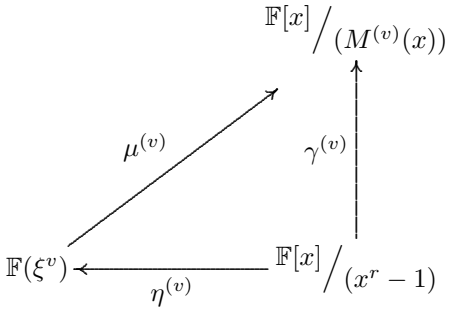
La matrice $\Gamma^{(v)}$ è la matrice $(m(v)-1)\times r$ dell'epimorfismo $\gamma^{(v)}$ come matrice di trasformazione fra $\mathcal{R}_{r,q}$ con base canonica $\{1,x,x^2,\ldots,x^{r-1}\}$ ed $\mathcal{Q}_{r,q}^{(v)}$ con base canonica $\{1,x,x^2,\ldots,x^{m(v)-1}\}$. Le colonne di $\Gamma^{(v)}$ sono le immagini degli elementi della base di $\mathcal{R}_{r,q}$ tramite γ.
Allo stesso modo suddividiamo in l epimorfismi η:

$$\begin{aligned} \eta^{(v)} : \mathcal{R}_{r,q} &\longrightarrow \mathcal{Q}_{r,q}^{(v)} = \mathbb{F}(\xi^v) \\ a(x) &\longmapsto a(\xi^v) \\ x^j &\longmapsto \xi^{jv} \end{aligned}$$

e quindi abbiamo la rappresentazione sulle algebre dei vettori circolanti e circolanti concatenati

$$\begin{aligned} \rho^{(v)} : \mathcal{V}_{r,q}^{c} &\longrightarrow \mathcal{V}_{m(v),q}^{c} \\ (a_0,\ldots,a_{r-1}) &\longmapsto \psi_6(\eta(\psi_2^{-1}(a_0,\ldots,a_{r-1}))) \end{aligned}$$

La matrice $H^{(v)}$ è la matrice $(m(v)-1)\times r$ dell'epimorfismo $\eta^{(v)}$ come matrice di trasformazione fra $\mathcal{R}_{r,q}$ con base canonica $\{1,x,x^2,\ldots,x^{r-1}\}$ ed $\mathcal{Q}_{r,q}^{(v)}$ con base canonica $\{1,x,x^2,\ldots,x^{m(v)-1}\}$. Le colonne di $H^{(v)}$ sono le immagini degli elementi della base di $\mathcal{R}_{r,q}$ tramite η.
Gli epimorfismi commutano il diagramma seguente dove μ a dominio e codominio opportunamente ristretti è un isomorfismo:



Conseguenza diretta della definizione di Γ ed H sono le seguenti proprietà[5]:

Proprietà 3.2.1. *Sia $0 \leq k \leq r-1$ allora il polinomio*

$$\sum_{j=0}^{m(v)-1} \Gamma_{j,k}^{(v)} x^j$$

è il resto della divisione di x^k per $M^{(v)}(x)$ in $\mathbb{F}_q$.

Dimostrazione. I coefficienti del polinomio in questione si trovano sulla colonna k-esima di $\Gamma^{(v)}$, blocco di Γ, le cui colonne sono le immagini degli elementi della base di $\mathcal{R}_{r,q}$. Alla k-esima colonna del blocco $\Gamma^{(v)}$ troviamo $x^k \mod M^{(v)}(x)$. □

Proprietà 3.2.2. *Sia γ trasformata di Winograd fra $\mathcal{R}_{r,q}$ ed $\mathcal{Q}_{r,q}$ e sia $\mathbf{e}_v$ il v-esimo idempotente minimale di $\mathcal{Q}_{r,q}$, allora*

1. $ker(\gamma^{(v)}) = (M^{(v)}(x))$
2. $\gamma^{-1}((\mathbf{e}_v)) = (\hat{M}^{(v)}(x))$

Dimostrazione. Il primo punto è una conseguenza immediata della definizione:

$$ker(\gamma^{(v)}) = \{a(x) \mid a(x) \equiv 0 \mod M^{(v)}(x)\} = (M^{(v)}(x))$$

Dimostriamo il secondo punto:
$\gamma^{-1}((\mathbf{e}_v))$ è ancora un ideale in $\mathcal{R}_{r,q}$ dato che è controimmagine di un ideale tramite isomorfismo. La controimmagine del generatore, che genera l'ideale controimmagine, è

$$\gamma^{-1}(0,\ldots,0|1,0,\ldots,0|0,\ldots,0) = ?(x)$$

dove $?(x)$ è quel polinomio di $\mathcal{R}_{r,q}$ il cui resto modulo $M^{(v)}(x)$ dà come risultato 1 e modulo $M^{(u)}(x)$ dà come risultato 0 per ogni u diverso da v. Pertanto $?(x) = \hat{M}^{(v)}(x)$. □

Dimostriamo finalmente che H coincide con Γ:

Teorema 3.2.1. *Sia v elemento dell'insieme delle etichette $\mathscr{L}_{r,q}$, Γ ed H definite come sopra, allora per ogni $i = 0,\ldots,m(v)-1$ e $j = 0,\ldots,r$, segue che*

$$\Gamma_{i,j}^{(v)} = H_{i,j}^{(v)}$$

[5] Proprietà 2.3 pag. 23 e variante della 1.13 pag. 15 [8].

Dimostrazione. Ricordiamo inizialmente che per definizione abbiamo

$$\gamma^{(v)}(a(x)) = a(x) \mod M^{(v)}(x)$$
$$\eta^{(v)}(a(x)) = a(\xi^v)$$

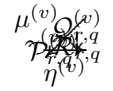
quindi le immagini degli elementi della base, che nella corrispondente rappresentazione dei vettori circolanti determinano le colonne delle matrici di trasformazione Γ e H sono date da

$$\gamma^{(v)}(x^k) = x^k \mod (M^{(v)}(x))$$
$$\eta^{(v)}(x^k) = \xi^{vk}$$

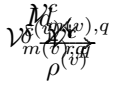
Dall'isomorfismo fra le strutture $\mathcal{P}_{r,q}^{(v)}$ e $\mathcal{Q}_{r,q}^{(v)}$ presentato nel lemma 1.1.1 abbiamo

$$\forall a(x) \in \mathcal{R}_{r,q} \quad \psi_5(\gamma^{(v)}(a(x))) = \psi_6(\eta^{(v)}(a(x)))$$

dal fatto che i coefficienti di $\gamma^{(v)}(a(x))$ coincidono con i coefficienti di $\eta^{(v)}(a(x))$ sulle basi delle rispettive strutture.
In altre parole, dal diagramma seguente

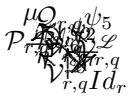


siamo passati alle algebre di vettori circolanti concatenati isomorfe:

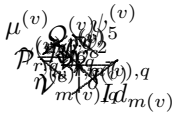


Dato che $\mu^{(v)}$ nel passaggio alla nuova rappresentazione risulta essere l'identità, le trasformazioni γ^v ed η^v coincidono e quindi anche le rispettive trasformazioni lineari $H^{(v)}$ e $\Gamma^{(v)}$ coincidono. □

Possiamo riassumere il teorema precedente con un diagramma che contiene tutti gli isomorfismi che abbiamo utilizzato:



Il diagramma degli epimorfismi corrispondenti per $v \in \mathscr{L}$ risulta essere:



Una conseguenza del teorema precedente è il fatto che

$$\gamma^{(v)}(x^k) = \sum_{j=0}^{r} \Gamma_{k,j}^{(v)} x^j \qquad \eta^{(v)}(x^k) = \sum_{j=0}^{r} H_{k,j}^{(v)} \xi^{vj} \qquad k = 0, \ldots, m(v) - 1$$

e quindi (a meno di isomorfismi), dato che H e Γ coincidono

$$\gamma^{(v)}(x^k) = \sum_{j=0}^{r} \Gamma_{k,j}^{(v)} x^j = \sum_{j=0}^{r} \Gamma_{k,j}^{(v)} \xi^{vj}$$

3.3 Proprietà strutturali di Γ

In questo paragrafo esaminiamo alcune proprietà di Γ.

Lemma 3.3.1. *Sia*

$$M^{(v)}(x) = x^{m(v)} - \sum_{j=1}^{m(v)} a_j^{(v)} x^{m(v)-j}$$

il v-esimo polinomio minimo nella fattorizzazione di $x^r - 1$ per v appartenente all'insieme delle etichette. Allora i suoi coefficienti soddisfano l'equazione

$$a_j^{(v)} = \Gamma^{(v)}_{m(v)-j,m(v)}$$

Dimostrazione. Dalla proprietà 3.2.1 abbiamo che

$$x^{m(v)} = \sum_{j=0}^{m(v)-1} \Gamma^{(v)}_{j,m(v)} x^j$$

quindi

$$x^{m(v)} \equiv \sum_{j=0}^{m(v)-1} \Gamma^{(v)}_{j,m(v)} x^j \mod M^{(v)}(x)$$

che con una sostituzione opportuna degli indici ci dà

$$x^{m(v)} \equiv \sum_{j=0}^{m(v)-1} \Gamma^{(v)}_{m(v)-j,m(v)} x^{m(v)-j} \mod M^{(v)}(x)$$

e quindi

$$M^{(v)}(x) = x^{m(v)} - \sum_{j=0}^{m(v)-1} \Gamma^{(v)}_{m(v)-j,m(v)} x^{m(v)-j}$$

□

Teorema 3.3.1. *Sia v appartenente all'insieme delle etichette $\mathscr{L}_{r,q}$, allora per ogni $i \in \{0, 1, \ldots, m(v) - 1\}$ segue che*

$$\Gamma^{(v)}_{i,j} = \delta_{i,j} \qquad j \in \{0, 1, \ldots, m(v) - 1\} \tag{3.1}$$

$$\Gamma^{(v)}_{i,n} = \sum_{k=1}^{m(v)} a_k^{(v)} \Gamma^{(v)}_{i,n-k} \qquad \forall n \in \mathbb{Z} \tag{3.2}$$

Dimostrazione. Sia ξ radice primitiva r-esima dell'unità, allora ξ^v è radice di $M^{(v)}(x)$ e soddisfa quindi la relazione di ricorrenza

$$\xi^{v(n+1)} = a_1^{(v)} \xi^{vn} + a_2^{(v)} \xi^{v(n-1)} + \cdots + a_{m(v)}^{(v)} \xi^{v(n-m(v)+1)} \qquad \forall n \in \mathbb{Z}$$

Considerando l'isomorfismo η ed il conseguente epimorfismo $\eta^{(v)}$ analogo di $\gamma^{(v)}$, si ottiene per $j \in \{0, 1, \ldots, r\}$

$$\eta^{(v)} : \mathbb{F}[x] \Big/ x^r - 1 \longrightarrow \mathbb{F}(\xi^v)$$
$$x^j \longmapsto \xi^{jv} = \sum_{i=1}^{m(v)} \Gamma_{i,j}^{(v)} \xi^{iv}$$

Quindi appoggiandoci a questa rappresentazione

$$\eta^{(v)}(x^n) = \xi^{vn} = \sum_{k=1}^{m(v)} a_k^{(v)} \xi^{v(j-k)}$$

e dato che $(\gamma^{(v)}(x^n))_i = \Gamma_{i,n}^{(v)}$ e tramite η al coefficiente dell'elemento $(\xi^{v(j-k)})_i$ della base corrisponde $\Gamma_{i,j-k}^{(v)}$ segue la tesi. □

Corollario 3.3.1. *Con le condizioni precedenti*

$$a_h^{(v)} = \sum_{k=0}^{m(v)-1} a_k^{(v)} \Gamma_{m(v)-h,m(v)-k}^{(v)} \tag{3.3}$$

Dimostrazione. Dal lemma segue che

$$a_h^{(v)} = \Gamma_{m(v)-h,m(v)}^{(v)}$$

che sostituita nella seconda equazione della tesi del teorema 3.3.1 determina la tesi cercata. □

Il teorema precedente[6] mette a disposizione un sistema per definire la matrice Γ per ricorrenza. Possiamo quindi calcolare la matrice senza effettuare divisioni ma semplicemente applicando le equazioni 3.1 e possiamo osservare che il primo blocco di dimensione $m(v) \times m(v)$ (con righe e colonne sempre numerate partendo da zero) di Γ è sempre la matrice identità da $\Gamma_{i,j}^{(v)} = \delta_{i,j}$ e che il blocco successivo al primo è un blocco circolante da $\Gamma_{i,n}^{(v)} = \sum_{k=1}^{m(v)} a_k^{(v)} \Gamma_{i,n-k}^{(v)}$.
Terminiamo il paragrafo con un lemma che, oltre a servire nelle dimostrazioni del prossimo paragrafo, avrà una interessante conseguenza[7] presentata nel capitolo sulle applicazioni:

Lemma 3.3.2. *Sia $m(x)$ polinomio di $\mathcal{R}_{r,q}$ e sia $v \in \mathscr{L}$, allora $m(x) \in (M^{(v)}(x))$ se e solo se $\Gamma^{(v)}\mathbf{m}^t = 0$.*

Dimostrazione. Ricordando che

1. $\eta^{(v)}(x^j) = \xi^{vj} = \sum_{i=0}^{m(v)-1} \Gamma_{i,j}^{(v)} \xi^{iv}$ per $j \in \mathbb{Z}_r$.
2. Per $H^{(v)}$ matrice dell'epimorfismo $\eta^{(v)}$ si ha $\eta^{(v)} m(x) = H^{(v)}\mathbf{m}^t$.
3. $H = \Gamma$ per H e Γ matrici di trasformazione di η e γ.

allora vale la catena di biimplicazioni:
$m(x) \in (M^{(v)}(x))$ se e solo se $(M^{(v)}(x))$ divide $m(x)$ se e solo se $m(\xi^v) = 0$ se e solo se $\eta^{(v)} m(x) = H^{(v)}\mathbf{m}^t = 0$ se e solo se $\Gamma^{(v)}\mathbf{m}^t = 0$. □

[6] [8] pag. 25.
[7] [8] pag. 33 e successive.

3.4 Matrice inversa Δ

Indichiamo con Δ la matrice inversa della matrice di Winograd Γ. Possiamo suddividere Δ in componenti verticali $\Delta^{(v)}$ in modo analogo a quanto fatto per Γ:

$$\begin{aligned}(\gamma^{(v)})^{-1} : \mathcal{Q}_{r,q}^{(v)} &\longrightarrow \mathcal{R}_{r,q} \\ a(x) \mod M^{(v)}(x) &\longmapsto b(x)\end{aligned}$$

dove $b(x)$ è il risultato del sistema di congruenze $b(x) \equiv a(x) \mod M^{(v)}(x)$ ricavato con il teorema cinese dei resti polinomiale.
La matrice di trasformazione corrispondente $\Delta^{(v)}$ è una matrice $r \times m(v)$; indicheremo la sua colonna j-esima con $\Delta^{(v)}_{\sim,j}$ e la sua riga i-esima con $\Delta^{(v)}_{i,\sim}$.
Presentiamo alcune proprietà che avranno delle applicazioni nella teoria dei codici correttori:

Proprietà 3.4.1. *Sia $v \in \mathscr{L}_{r,q}$ elemento dell'insieme delle etichette. La prima colonna di ogni blocco $\Delta^{(v)}$ di Δ è l'idempotente miminale $\mathbf{e}_v$ che genera l'ideale minimale*

$$(\hat{M}^{(v)}(x)) = \Big(\frac{1-x^r}{M^{(v)}(x)}\Big)$$

nell'algebra isomorfa $\mathcal{R}_{r,q}$.

Dimostrazione. Δ è la matrice dell'isomorfismo fra le due strutture isomorfe $\mathcal{V}^{\mathscr{L}}_{r,q}$ e $\mathcal{V}^{c}_{r,q}$, quindi manda idempotenti minimali in idempotenti minimali: l'immagine tramite Δ della rappresentazione vettoriale dell'idempotente minimale $e_v(x) \in (\hat{M}^{(v)}(x))$ è un idempotente minimale.
Dobbiamo dimostrare che ad $\mathbf{e}_v$ corrisponde proprio l'idempotente dell'ideale $(\hat{M}^{(v)}(x))$ e non di qualche altro ideale.
Osserviamo che

$$\mathbf{e}_v = (0,0,\ldots,0,1,0,\ldots,0) = \mathit{concat}(\mathbf{0},\ldots,\mathbf{0},(1,0,\ldots,0),\mathbf{0},\ldots,\mathbf{0})$$

che corrisponde al vettore circolante $(1,0,\ldots,0) \in \mathcal{V}^{c}_{m(v),q}$ concatenato in un vettore di $\mathcal{V}^{\mathscr{L}}_{r,q}$ alla posizione contrassegnata dall'etichetta v, nel quale tutti gli altri vettori sono nulli. Quindi a meno di isomorfismi sulle algebre $\mathcal{V}^{\mathscr{L}}_{r,q}$ e $\mathcal{R}_{r,q}$:

$$\Delta\mathbf{e}_v^t = \Delta \begin{pmatrix} 0 \\ \hline 0 \\ \vdots \\ 0 \\ \hline \vdots \\ \hline 1 \\ 0 \\ 0 \\ \hline \vdots \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix} \cong \sum_{i=0}^{r-1} \Delta^{(v)}_{i,0} x^i = a(x)$$

per $a(x) \in \mathcal{R}_{r,q}$ polinomio i cui elementi sono definiti dalla prima colonna del blocco $\Gamma^{(v)}$. Dato che Γ è la matrice inversa di Δ, $a(x) \cong \mathbf{a}$ soddisfa la seguente condizione:

$$\Gamma \mathbf{a}^t = \mathbf{e}_v^t$$

e considerandone i blocchi, per ogni $u \in \mathscr{L}$, $u \neq v$, si ottiene

$$\Gamma^{(u)} \mathbf{a}^t = \mathbf{0}$$

Pertanto dal lemma 3.3.2 segue che $a(x) \in (\hat{M}^{(v)}(x))$ ed essendo un idempotente primitivo, è proprio quello che genera $(\hat{M}^{(v)}(x))$ □

Proprietà 3.4.2. *L'insieme $\{\Delta^{(v)}_{\sim,0}\}_{v\in\mathscr{L}}$ delle prime colonne di tutti i blocchi di Δ costituisce l'insieme di tutti gli idempotenti minimali che generano tutti gli ideali minimali dell'algebra $\mathcal{V}^{\mathscr{L}}_{r,q}$ isomorfa a $\mathcal{R}_{r,q}$.*

Dimostrazione. È una generalizzazione della proprietà precedente: ogni prima colonna dei blocchi di Δ è un idempotente minimale in $\mathcal{V}^{\mathscr{L}}_{r,q}$, e dal corollario 2.3.1 segue che ogni idempotente genera un ideale minimale. □

La seguente proprietà fornisce un metodo per costruire in modo ricorrente i blocchi di Δ.

Proprietà 3.4.3. *Sia $v \in \mathscr{L}$ e $j \in \{0, 1 \ldots, m(v) - 1\}$. La j-esima colonna del v-esimo blocco di Δ, indicata con $\Delta^{(v)}_{\sim,j}$ è uno shift circolare verso il basso di j posti di $\Delta^{(v)}_{\sim,0}$:*

$$\left(\Delta^{(v)}_{\sim,j}\right)^t \in \mathcal{V}^c_{m(v),q} \qquad \left(\Delta^{(v)}_{\sim,j}\right)^t = (0,1,0,\ldots,0)^j \star \left(\Delta^{(v)}_{\sim,0}\right)^t$$

Dimostrazione. Consideriamo la rappresentazione polinomiale di $\mathcal{V}^{\mathscr{L}}_{r,q}$ e l'idempotente minimale $\mathbf{e}_v$ in questa rappresentazione:

$$d(x) = \psi_2^{-1}\left((\Delta^{(v)}_{\sim,j})^t\right) \qquad d(x) \in \mathcal{R}_{r,q}$$
$$e_v(x) = \psi_2^{-1}\left((\Delta^{(v)}_{\sim,0})^t\right)$$

L'immagine del v-esimo epimorfismo η^v di $x^j e_v(x)$ cioè del j-esimo shift dell'idempotente minimale $e_v(x)$ risulta essere

$$\eta^{(v)}(x^j e_v(x)) = \eta^{(v)}(x^j)\eta^{(v)}(e_v(x)) = \xi^{vj} e_v(\xi^v) \in \mathcal{P}_{r,q}$$

che rappresenta $e_v(x)$ shiftato di j posti.
La sua immagine tramite $\psi_5^{(v)}$ nello spazio dei vettori circolanti $\mathcal{V}^c_{m(v),q}$ è costituita dal vettore che ha 1 al j-esimo posto e tutti gli altri nulli. Anche considerando la sua immagine al codominio esteso $\mathcal{V}^{\mathscr{L}}_{r,q}$ si ottiene un vettore che ha 1 al j-esimo posto del blocco v-esimo e zero in tutti gli altri, infatti per $u \in \mathscr{L}$ diverso da v segue che

$$\eta^{(u)}(x^j e_v(x)) = 0 \qquad \forall j = 0, \ldots, m(u) - 1$$

Quindi la proprietà vale per tutti gli idempotenti minimali, e dato che ogni idempotente è somma di idempotenti, per la linearità degli isomorfismi utilizzati, vale in generale. □

In $\mathcal{V}^c_{m(v),q}$ ogni shift dell'idempotente minimale $e_v(x) = (1, 0, \ldots, 0)$ determina l'elemento successivo della base canonica come spazio vettoriale. A meno di isomorfismi $x^j e_v(x) = (0, \ldots 0, 1, 0, \ldots, 0)$ dove l'1 si trova al j-esimo posto. Le colonne $\Delta^{(v)}_{\sim,j}$ per $j = 0, \ldots, m(v)-1$ determinano una base di $\mathcal{V}^c_{m(v),q}$. Abbiamo quindi

Proprietà 3.4.4. *L'insieme dei vettori colonna* $\{\Delta^{(v)}_{\sim,j}\}_{j=0}^{m(v)-1}$ *determina una base dell'ideale* $(M^{(v)}(x))$.

Dimostrazione. La tesi segue da quanto osservato: le colonne $\Delta^{(v)}_{\sim,j}$ per $j = 0, \ldots, m(v) - 1$ definiscono una base di $\mathcal{V}^c_{m(v),q}$ ed inoltre dalla proprietà 3.4.1 ogni colonna è uno shift dell'idempotente che genera l'ideale minimale $(M^{(v)}(x))$. □

Concludiamo il capitolo proponendo, senza dimostrazione, una formula[8] che lega la matrice Γ alla sua inversa Δ:

$$\Delta_{i,j} = \frac{1}{r} \sum_{k=1}^{m(v)-1} \Gamma^{(v)}_{k,j-i+k}$$

Dalle proprietà di questo capitolo abbiamo visto che i blocchi di Γ determinano gli ideali e che le prime colonne di Δ determinano gli idempotenti minimali di $\mathcal{R}_{r,q}$; questo fatto sarà una delle basi delle applicazioni della trasformata di Winograd alla teoria dei codici correttori proposte nell'ultimo capitolo.

[8] [8] pag. 28 e successive, oppure usando i risultati del . capitolo 6 di questa tesi ed il teorema 6.23 di pag 145 di [2].

Capitolo 4

Codici lineari

La teoria dei codici correttori si basa sulla necessità di trasmettere dei vettori di informazione attraverso un canale che per motivi tecnici può alterare parte del messaggio. In ogni sistema di comunicazione infatti la ricezione può essere disturbata da segnali di interferenza chiamati genericamente "rumore"che si sommano all'informazione originariamente trasmessa. Riconoscere che un errore è entrato nel messaggio ricevuto ed eventualmente correggerlo è possibile utilizzando determinati strumenti algebrici.
In questo capitolo riportiamo le definizioni e i risultati fondamentali della teoria dei codici lineari utilizzando le notazioni orientate allo sviluppo della terza parte [1] .

4.1 Codici rivelatori e codici correttori

I messaggi inviati attraverso un sistema di comunicazione sono generalmente vettori di lunghezza k appartenenti ad uno spazio metrico. Si vedrà nel corso del paragrafo che è conveniente aggiungere a tale spazio la struttura di spazio vettoriale.

Definizione 4.1.1. *Sia $\mathbb{F} = \mathbb{F}_q$ campo finito di ordine p e di cardinalità $q = p^n$. L'insieme delle r-uple ordinate ad elementi in $\mathbb{F}$ definito come*

$$\mathbb{F}_q^r = \{\mathbf{x} = (x_0, x_1, \ldots, x_{r-1}) \mid x_j \in \mathbb{F}\}$$

e considerato congiuntamente con l'operazione binaria

$$\begin{aligned} \rho : \mathbb{F}^r \times \mathbb{F}^r &\longrightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\longmapsto \rho(\mathbf{x}, \mathbf{y}) = |\{j \mid x_j \neq y_j\}| \end{aligned}$$

è detto **spazio delle parole** *di lunghezza r e soddisfa gli assiomi di spazio metrico. I suoi elementi sono detti* **parole** *ed $\mathbb{F}_q$ è detto* **alfabeto** *dello spazio. Un qualsiasi sottoinsieme C di $\mathbb{F}^r$ è detto* **codice**, *mentre i sottospazi vettoriali sono detti* **codici lineari**.

Definizione 4.1.2. *Si definisce* **distanza minima** *la più piccola distanza fra due parole del codice:*

$$d(C) := min\{\rho(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in C\}$$

[1] I risultati citati compaiono in [2], [18], [4] e [24].

Il vantaggio principale nel considerare i codici C come sottospazi vettoriali dello spazio delle parole, consiste nel poter definire una norma e nel poter esprimere i codici in modo compatto tramite matrici.

Definizione 4.1.3. *Si definisce* **peso di Hamming** *di una parola* $\mathbf{x}$*, la sua distanza dall'origine, cioè il numero delle sue componenti non nulle.*

$$w(\mathbf{x}) = |\{j \mid x_j \neq 0\}| = \rho(\mathbf{x}, \mathbf{0})$$

Si verifica che w *soddisfa la definizione di norma dello spazio vettoriale* $\mathbb{F}^r$ *e dei suoi sottospazi.*
Il più piccolo peso delle parole di un codice è detto **peso minimo** *ed equivale alla distanza minima.*

Si introduce la definizione di rumore, di funzione rivelatore e la pseudo-funzione correttore.

Definizione 4.1.4. *Si definisce* **rumore** *un qualsiasi vettore* $\mathbf{r}$ *di* $\mathbb{F}^r$ *che viene sommato alla parola inviata durante la trasmissione. Il livello di rumore è dato dal suo peso di Hamming e si dice* **rumore minimo** *se il suo peso di Hamming è pari ad* 1*. Dato un codice* C *con distanza minima* d *si definisce* **rumore massimo riconoscibile** *il vettore* $\mathbf{r}$ *di* $\mathbb{F}^r$ *tale che suo peso di Hamming equivale a* $\llcorner d/2 \lrcorner$*, cioè alla parte intera di* $d/2$*.*
Mentre per $\mathbf{c}$ *parola di* C *si definisce* **rumore critico** *di* $\mathbf{c}$ *il vettore* $\mathbf{r}$ *di* $\mathbb{F}^r$ *tale che* $\mathbf{c} + \mathbf{r}$ *appartenga a* $C \setminus \{\mathbf{c}\}$*.*

Definizione 4.1.5. *Dato il codice* C *con distanza minima* d*, la funzione che associa ad ogni parola dello spazio* $\mathbb{F}^r$ *il simbolo* 0 *se la parola non appartiene al codice ed il simbolo* 1 *se la parola vi appartiene è detta funzione* **rivelatore**

$$\begin{aligned} r : \mathbb{F}^r &\longrightarrow \{0,1\} \\ \mathbf{v} &\longmapsto 0 \quad if \quad \mathbf{v} \notin C \\ \mathbf{v} &\longmapsto 1 \quad if \quad \mathbf{v} \in C \end{aligned}$$

Definizione 4.1.6. *Dato il codice* C *con distanza minima* d*, la pseudofunzione che associa ad ogni parola dello spazio* $\mathbb{F}^r$ *la parola (o le parole) del codice ad essa più vicina è detta funzione* **correttore**

$$\begin{aligned} c : \mathbb{F}^r &\longrightarrow C \\ \mathbf{x} + \mathbf{r} &\longmapsto \mathbf{x} \end{aligned}$$

dove $\mathbf{x}$ *è una parola del codice ed* $\mathbf{r}$ *è il rumore.*

La funzione c è una pseudofunzione se esiste un vettore di $\mathbb{F}^r$ avente distanza pari a $\llcorner d/2 \lrcorner$ da due parole distinte del codice C. Sulle definizioni appena enunciate si basano le prossime di codice rivelatore e codice correttore:

Definizione 4.1.7. *Un codice* C *con distanza minima* d *si dice* e**-rivelatore** *se* e *è il peso massimo del rumore che può essere sommato ad una parola del codice, senza che l'immagine del correttore di tale somma risulti essere* 1 *(cioè senza che il rumore la trasformi in un'altra parola del codice). In questo caso* $e + 1$ *è il più piccolo peso dei rumori critici delle parole del codice.*

Definizione 4.1.8. *Un codice C con distanza minima d si dice* **e-correttore** *se la funzione c è ben definita per la restrizione del dominio ai vettori di $\mathbb{F}^r$ del tipo $\mathbf{v} = \mathbf{x} + \mathbf{r}$ dove x è un vettore di C ed r ha peso di Hamming minore o uguale ad e.*

Presentiamo tre proprietà che mettono in relazione le caratteristiche di un codice come sottospazio vettoriale alle sue capacità di correggere e rivelare gli errori.

Proprietà 4.1.1. *Sia C codice di lunghezza r. C è e-rivelatore se e solo se $d(C) = e + 1$.*

Dimostrazione. Dimostriamo separatamente le due inclusioni.

$\Rightarrow$) Se per assurdo $d(C) < e + 1$ allora esistono due vettori del codice $\mathbf{x}, \mathbf{y}$, la cui distanza è minore di e. Quindi C non può essere e-rivelatore.

$\Leftarrow$) Viceversa, siano $\mathbf{x} \in C$, $\mathbf{r} \in \mathbb{F}^r$ tale che $w(\mathbf{r}) \leq e$. Allora

$$\begin{aligned} &\rho(\mathbf{x}, \mathbf{x} + \mathbf{r}) \leq e < d(C) \\ &\Rightarrow \mathbf{x} + \mathbf{r} \notin C \end{aligned}$$

Se invece $w(\mathbf{r}) > e + 1$ allora

$$\rho(\mathbf{x}, \mathbf{x} + \mathbf{r}) \geq e + 1 > d(C)$$

quindi il vettore $\mathbf{x} + \mathbf{r}$ può eventualmente appartenere al codice.

□

Proprietà 4.1.2. *Sia C codice di lunghezza r. C è e-correttore se e solo se $d(C) = 2e + 1$.*

Dimostrazione. Dimostriamo separatamente le due inclusioni.

$\Rightarrow$) Se C è e-correttore, allora per ogni coppia di parole distinte del codice $\mathbf{x}, \mathbf{y}$ e comunque scelti $\mathbf{r}_1$ ed $\mathbf{r}_2$ rumori di peso massimo e, $\mathbf{x} + \mathbf{r}_1 \neq \mathbf{y} + \mathbf{r}_2$ cioè la funzione correttore è ancora ben definita, quindi $\rho(\mathbf{x}, \mathbf{y}) = 2e + 1$.

$\Leftarrow$) Viceversa se $d(C) = 2e + 1$ allora comunque scelti $\mathbf{x}, \mathbf{y} \in C$ parole distinte del codice $\rho(\mathbf{x}, \mathbf{y}) \geq 2e + 1$. Sia $\mathbf{r}$ rumore di peso $w(\mathbf{r}) \leq e$ allora $\rho(\mathbf{x} + \mathbf{r}, \mathbf{y}) \geq e + 1$, quindi la funzione correttore è ben definita per il dominio ristretto a tutti gli elementi del tipo $\mathbf{x} + \mathbf{r}$.

□

Proprietà 4.1.3. *Sia C codice di lunghezza r. C è $\llcorner(d-1)/2\lrcorner$-correttore.*

Dimostrazione. Segue dalla proprietà 4.1.2, infatti se $d = 2e + 1$ allora $e = \llcorner(d-1)/2\lrcorner$. □

Se nei casi in cui la funzione rivelatore è ben definita allora la codifica del codice è detta **completa**. Abbiamo una codifica **incompleta** quando c'è una parola che non appartiene al codice e che può essere decodificata indifferentemente con due parole distinte del codice.

4.1.1 Codici correttori perfetti e limitazione di Hamming

La metrica nello spazio delle parole $\mathbb{F}^r$ induce una topologia i cui chiusi sono definiti da sfere indicate con

$$S_t(\mathbf{x}) := \{\mathbf{y} \in \mathbb{F}^r \mid \rho(\mathbf{x}, \mathbf{y}) \leq t\}$$

la cui frontiera è una superficie sferica definita da

$$fr(S_t(\mathbf{x})) = \sigma_t(\mathbf{x}) := \{\mathbf{y} \in \mathbb{F}^r \mid \rho(\mathbf{x}, \mathbf{y}) = t\}$$

Ciascuna sfera è unione di superfici sferiche concentriche di raggio compreso fra 0 ed il raggio della sfera:

$$S_t(\mathbf{x}) = \bigcup_{s=0}^{t} \sigma_s(\mathbf{x})$$

da cui la cardinalità di una sfera è pari alla somma delle cardinalità delle superfici sferiche concentriche che lo definiscono, essendo queste fra loro disgiunte

$$|S_t(\mathbf{x})| = \bigcup_{s=0}^{t} |\sigma_s(\mathbf{x})|$$

Si verifica inoltre che

$$|\sigma_s(\mathbf{x})| = \binom{r}{s}(q-1)^s$$

quindi

$$|S_t(\mathbf{x})| = \sum_{s=0}^{t} \binom{r}{s}(q-1)^s$$

Proprietà 4.1.4. *Sia C codice di lunghezza r. C è e-correttore se e solo se per ogni coppia di parole distinte del codice $\mathbf{x}, \mathbf{y}$, l'intersezione fra le due sfere $S_e(\mathbf{x}), S_e(\mathbf{y})$ è vuota.*

Dimostrazione. Dalla proprietà 4.1.2 C è e correttore se e solo se $d(C) = 2e+1$, allora segue

$$\begin{aligned} d(C) = 2e+1 &\iff \forall \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y} \quad \rho(\mathbf{x}, \mathbf{y}) \geq 2e+1 \\ &\iff S_e(\mathbf{x}) \cap S_e(\mathbf{x}) = \emptyset \end{aligned}$$

□

Sono stati appena visti alcuni vincoli per una scelta della distanza minima d ottimale, che possono essere formalizzati nella seguente

Proprietà 4.1.5. *Sia C codice con distanza minima d pari, allora C è un codice $(d/2-1)$-correttore ed è un codice $d/2$-rivelatore*

Dimostrazione. Sia $\mathbf{r} \in \mathbb{F}^r$ tale che $w(\mathbf{r}) = d/2$, allora per ogni $\mathbf{x}$ parola del codice si ha che $\mathbf{x}+\mathbf{r} \in S_{d/2}(\mathbf{x})$, ma contemporaneamente può esistere $\mathbf{y} \in C$ tale che $\mathbf{x}+\mathbf{r} \in S_{d/2}(\mathbf{y})$ senza contraddizioni dato che $\rho(\mathbf{x}, \mathbf{y}) \geq d/2 + d/2 = d$. □

Considerando lo spazio delle parole, se esiste un raggio ed un insieme di centri tali che le sfere siano disgiunte e che non ci siano parole che non appartengano ad alcuna di queste sfere, allora i centri costituiscono un codice nello spazio delle parole di tipo particolare.

Definizione 4.1.9. *Dato un codice C e-correttore, si definisce* **insieme dei punti di difetto**

$$set(\delta_C) := \mathbb{F}^r \setminus \bigcup_{\mathbf{x}\in C} S_e(\mathbf{x})$$

e si definisce **difetto** *il numero*

$$\begin{aligned}\delta_C &:= |\mathbb{F}^r| - |\bigcup_{\mathbf{x}\in C} S_e(\mathbf{x})| \\ &= |\{\mathbf{y} \in \mathbb{F}^r \mid \mathbf{y} \notin \cup_{\mathbf{x}\in C} S_e(\mathbf{x})\}|\end{aligned}$$

Definizione 4.1.10. *Un codice C e-correttore si dice* **perfetto** *se la famiglia di sfere $\{S_e(\mathbf{x})\}_{\mathbf{x}\in C}$ costituisce una partizione di $\mathbb{F}^r$, cioè se il suo difetto è pari a zero.*

Dalle considerazioni sulla cardinalità delle sfere segue la dimostrazione del prossimo teorema, detto **teorema di limitazione di Hamming**

Teorema 4.1.1. *Sia C codice e-correttore di lunghezza r definito sull'alfabeto $\mathbb{F} = \mathbb{F}_q$, allora*

$$|C||S_e(\mathbf{x})| \leq |\mathbb{F}^r|$$

per $\mathbf{x}$ generico elemento di C.
Se il codice C è perfetto vale allora

$$|C||S_e(\mathbf{x})| = |\mathbb{F}^r|$$

La tesi può essere riformulata come:

$$|C| \sum_{s=0}^{t} \binom{r}{s} (q-1)^s \leq q^r$$

Tre conseguenze del teorema di limitazione di Hamming sono

Corollario 4.1.1. *Il codice e-correttore C di lunghezza r è perfetto se*

$$|S_e(\mathbf{x})| \mid |\mathbb{F}^r|$$

o equivalentemente

$$\sum_{s=0}^{t} \binom{r}{s} (q-1)^s \mid q^r$$

Corollario 4.1.2. *Il codice 1-correttore C di lunghezza r è perfetto se*

$$(1 + r(q-1)) \mid q^r$$

Corollario 4.1.3. *Non esistono codici 1-correttori perfetti di lunghezza pari.*

4.1.2 (r,k)-codici e limitazione di Singleton

Presentiamo un ulteriore limite alla progettazione di codici correttori lineari, detto limitazione di Singleton. Il seguente teorema è un primo passo che conduce alla definizione di *posti di informazione*, che sarà cruciale nei prossimi capitoli.

Teorema 4.1.2. *Sia C codice di lunghezza r e distanza minima d, allora*

$$|C| \leq q^{r-d+1}$$

Dimostrazione. Segue da

$$|C|q^{d-1} \leq |\mathbb{F}^r|$$

□

Dal teorema precedente è possibile assegnare $k \leq r - d + 1$ posti di ogni parola del codice che la determinano univocamente.

Definizione 4.1.11. *Sia C codice di lunghezza r e distanza minima d. Fissati $k \leq r - d + 1$ interi positivi compresi fra 1 ed r ed indicati con $j_1, \ldots j_k$, si dicono* **posti di informazione** *se comunque scelta una k-upla $(y_1, \ldots, y_k)$ ad elementi in $\mathbb{F}$, esiste una ed una sola parola di C tale che nel posto j_i compaia y_i per ogni i compreso fra 1 e k. I posti che non sono di informazione sono detti di* **ridondanza**.

Nel processo di codifica quindi ogni messaggio, inteso come sequenza di lettere, viene suddiviso in vettori di lunghezza k e ad ogni vettore vengono aggiunti $r-k$ simboli di ridondanza. Il rapporto r/k viene chiamato tasso di informazione.

Definizione 4.1.12. *Un codice C di lunghezza r avente k posti di informazione si dice (r,k)-***codice**. *Se è necessario specificare anche la distanza minima d, il codice C sarà detto (r,k,d)-codice.*

La tesi del seguente teorema è nota come **limitazione di Singleton**:

Teorema 4.1.3. *Sia C un (r,k,d)-codice, allora $d \leq r - k + 1$*

Dimostrazione. Per la corrispondenza fra le parole del codice e i vettori di lunghezza k definita per gli (r,k)-codici segue che $|C| = q^k$, mentre dal teorema segue che $|C| = q^{r-d+1}$. □

Definizione 4.1.13. *Un (r,k)-codice C è detto a* **massima distanza separabile** *(o MDS) se comunque scelti k posti di un qualsiasi vettore del codice, questi sono di informazione.*

4.2 Matrice generatrice e matrice di controllo

Utilizzando il fatto che i codici lineari sono in particolare dei sottospazi vettoriali, possiamo scrivere gli elementi di un codice dalla matrice di passaggio dallo spazio $\mathbb{F}^r$ al sottospazio C.

Sia $\{\mathbf{b}_i\}_{i=1}^{r}$ base di $\mathbb{F}^r$ e sia $\{\mathbf{e}_i\}_{i=1}^{k}$ base di C, allora è noto che ogni vettore $\mathbf{b}_i$ può essere scritto come combinazione lineare dei $\mathbf{e}_i$:

$$\mathbf{b}_i = \sum_{j=1}^{r} a_{ij}\mathbf{e}_j$$

A tale sistema corrisponde la matrice di caratteristica k, indicata con G e detta **matrice generatrice** del codice C

$$G = \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1r} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2r} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & \cdots & a_{kr} \end{pmatrix}$$

In questo caso il codice C è detto generato da G. Utilizzando i risultati dell'algebra lineare [2] è possibile scrivere la matrice G in modo tale che le prime k colonne coincidano con le prime k colonne della matrice identità $k \times k$:

$$G = (I_k \mid E)$$

In questa forma G è detta **in forma standard**. Risulta immediato il rapporto fra gli (r, k)-codici sopra definiti ed i codici lineari:

Teorema 4.2.1. *Ogni matrice $k \times r$ di caratteristica k definisce un (r, k)-codice.*

Dimostrazione. Ad ogni matrice corrisponde infatti un sottospazio di dimensione k dello spazio delle parole rispetto a delle basi fissate. Tutte le combinazioni lineari degli elementi della base del sottospazio sono vettori di un sottospazio k-dimensionale che è quindi un (r, k)-codice. □

Dato che esistono matrici generatrici diverse che definiscono lo stesso sottospazio vettoriale, non c'è una corrispondenza biunivoca fra le matrici $k \times r$ di caratteristica k ed i codici lineari. I codici lineari equivalenti saranno approfonditi nei prossimi paragrafi.
Fra i vantaggi della rappresentazione matriciale si annovera un modo rapido per capire (e per costruire) codici a massima distanza separabile.

Teorema 4.2.2. [3] *Un codice lineare C k-dimensionale è un MDS se e solo se ogni minore di ordine k di una matrice generatrice del codice è non nullo.*

Dimostrazione. Dimostriamo il risultato per doppia inclusione:

⇒) C è MDS se k posti qualsiasi sono di informazione. Se per assurdo la matrice generatrice di C avesse un minore di ordine k nullo, i vettori di tale matrice non sarebbero linearmente indipendenti e due parole diverse di lunghezza k private dei simboli di ridondanza coinciderebbero.

⇐) Per contrapposizione: sia C codice avente G matrice generatrice che possiede un minore nullo di ordine k. Allora esistono almeno due parole diverse che private dei simboli di ridondanza risultano essere uguali. Quindi C non è un codice MDS.

[2] Ad esempio con una generalizzazione di pag. 99 [28].
[3] [2] pag. 131, teorema 6.9.

□

Si pone ora il problema di cercare un modo computazionalmente efficiente per determinare se una parola dello spazio $\mathbb{F}^r$ appartiene ad un dato codice lineare C. Per tale scopo è necessario introdurre i codici duali definiti sull'usuale prodotto scalare fra vettori:

Definizione 4.2.1. *Sia un codice lineare C di $\mathbb{F}^r$ si definisce* **spazio ortogonale** *l'insieme*

$$C^\perp := \{\mathbf{x} \mid \mathbf{x} \cdot \mathbf{c} = 0 \quad \forall \mathbf{c} \in C\}$$

che è a sua volta un sottospazio vettoriale di $\mathbb{F}^r$ detto **codice ortogonale** *(o codice duale).*

Dato che sui campi finiti possono esserci vettori isotropi, C e $C^\perp$ possono avere in comune vettori diversi dal vettore nullo. Nel caso in cui $C = C^\perp$ allora C è detto codice **autoduale**; gli unici codici autoduali possibili sono quelli in cui $k = n/2$. Il fatto che il codice ortogonale sia un sottospazio vettoriale di $\mathbb{F}^r$, porta alla seguente

Definizione 4.2.2. *Sia C (n,k)-codice lineare, allora la matrice generatrice del codice ortogonale è detta* **matrice di controllo** *ed è indicata con H.*

I prossimi teoremi stabiliscono un modo per determinare l'appartenenza di una parola ad un codice tramite la matrice di controllo.

Teorema 4.2.3. *Sia C (n,k)-codice lineare generato dalla matrice G. Allora*

$$\mathbf{x} \in C^\perp \iff G\mathbf{x}^t = \mathbf{0}^t$$

Dimostrazione. Dimosrtiamo per doppia inclusione:

⇒) Sia $\mathbf{x} \in C^\perp$ allora $\mathbf{x}$ è ortogonale ad ogni parola di C, quinde a maggior ragione è ortogonale alle parole di ogni base di C, anche della base i cui vettori costituiscono la matrice G: $G\mathbf{x}^t = \mathbf{0}^t$.

⇐) Viceversa se $G\mathbf{x}^t = \mathbf{0}^t$, allora $\mathbf{x}$ è ortogonale a una scelta di vettori $\{\mathbf{e}_0, \ldots, \mathbf{e}_{r-1}\}$ base di C. Dato che ogni parola del codice è definita come combinazione lineare $\mathbf{c} = \lambda_0\mathbf{e}_0 + \cdots + \lambda_{r-1}\mathbf{e}_{r-1}$ allora

$$\begin{aligned}\mathbf{x} \cdot \mathbf{c} &= \mathbf{x} \cdot (\lambda_0\mathbf{e}_0 + \cdots + \lambda_{r-1}\mathbf{e}_{r-1}) \\ &= \lambda_0(\mathbf{x} \cdot \mathbf{e}_0) + \cdots + \lambda_{r-1}(\mathbf{x} \cdot \mathbf{e}_{r-1}) \\ &= 0\end{aligned}$$

Da cui segue che $\mathbf{x}$ è ortogonale ad ogni parola di C ed appartiene quindi al codice duale.

□

Corollario 4.2.1. *Se C (n,k)-codice lineare allora $C^\perp$ è un $(r, r-k)$-codice lineare e la sua matrice generatrice è di dimensione $(r-k) \times r$.*

Dimostrazione. Dal teorema precedente i vettori $\mathbf{x} = (x_1, \ldots, x_r)$ di $C^\perp$ sono vettori di $\mathbb{F}_q^r$ ortogonali ai vettori di una base di C. Sia $\{\mathbf{e}_0 \ldots \mathbf{e}_k\}$ tale base, allora $\mathbf{x}$ appartiene a $C^\perp$ se le sue componenti soddisfano il sistema

$$(\mathbf{e}_i)_j x_j = 0 \qquad \forall i = 1, \ldots, k \quad \forall j = 1, \ldots, r$$

Tale sistema possiede esattamente $n-k$ soluzioni linearmente indipendenti, che costituiscono una base per $C^\perp$. □

Teorema 4.2.4. *Sia C (n,k)-codice lineare generato dalla matrice G ed avente H come matrice di controllo.*
Allora

$$\mathbf{x} \in C \iff H\mathbf{x}^t = \mathbf{0}^t$$

Dimostrazione. $\mathbf{x} \in C$ se e solo se $\mathbf{x} \cdot \mathbf{c}^\perp = \mathbf{0}$ $\forall \mathbf{c}^\perp \in C^\perp$. Ma questo accade in particolare per tutti i vettori della base che costituiscono le righe di H, e viceversa se accade per tutti i vettori della base accade per ogni altro vettore. □

Risulta quindi essere di importanza cruciale il valore di $H\mathbf{x}^t$ che prende il nome di **sindrome** del vettore $\mathbf{x} \in \mathbb{F}^r$. È possibile ricavare la matrice di controllo H costruendola a partire dalla matrice generatrice G, utilizzando il seguente

Teorema 4.2.5 (di correlazione fra G ed H)**.** *Sia C (r,k)-codice lineare di matrice generatrice G scritta in forma standard come $G = (I_k \mid E)$, allora la matrice di controllo risulta essere $H = (-E^t \mid I_{n-k})$.*

Dimostrazione. Se G, rispetto alla base standard di $\mathbb{F}_q^r$ e alla base $\{\mathbf{e}_0 \ldots \mathbf{e}_k$ di C, è data da

$$G = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & e_{1,k+1} & \cdots & e_{1,k+1} \\ 0 & 1 & \cdots & 0 & e_{1,k+1} & \cdots & e_{1,k+1} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & e_{1,k+1} & \cdots & e_{1,k+1} \end{array}\right)$$

allora un vettore $\mathbf{x}$ è una parola del codice C se e solo se

$$\left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & e_{1,k+1} & \cdots & e_{1,k+1} \\ 0 & 1 & \cdots & 0 & e_{1,k+1} & \cdots & e_{1,k+1} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & e_{1,k+1} & \cdots & e_{1,k+1} \end{array}\right) \left(\begin{array}{c} x_1 \\ \vdots \\ x_k \\ \hline x_{k+1} \\ \vdots \\ x_r \end{array}\right) = \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \end{array}\right)$$

che ha esattamente $n-k$ soluzioni linearmente indipendenti. Queste possono essere costuite assegnando ai vetttori $(r-k)$-dimensionali delle incognite i vettori della base standard di $\mathbb{F}_q^{r-k}$ per ottenere una matrice generatrice dello spazio duale della forma $H = (? \mid I_{n-k})$.
Tali soluzioni sono date da

$$\begin{array}{l} (-e_{1,k+1}, -e_{2,k+1}, \ldots, -e_{k,k+1}, 1, 0, \ldots, 0) \\ (-e_{1,k+2}, -e_{2,k+2}, \ldots, -e_{k,k+2}, 0, 1, \ldots, 0) \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ (-e_{1,r}, -e_{2,r}, \ldots, -e_{k,r}, 0, 0, \ldots, 1) \end{array}$$

e consentono di ricavare i vettori che occupano lo spazio che avevamo indicato con ?:

$$H = \left(\begin{array}{cccc|cccc} -e_{1,k+1} & -e_{2,k+1} & \dots & -e_{k,k+1} & 1 & 0 & \dots & 0 \\ -e_{1,k+2} & -e_{2,k+2} & \dots & -e_{k,k+2} & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots & \vdots & & \vdots \\ -e_{1,k+r} & -e_{2,k+r} & \dots & -e_{k,k+r} & 0 & 0 & \dots & 1 \end{array}\right)$$

da cui risulta la tesi. □

4.2.1 Codici lineari equivalenti

Come già accennato due sottospazi vettoriali isomorfi possono essere generati da basi diverse, quindi serve un criterio per distinguere i codici isomorfi da quelli non isomorfi.

Definizione 4.2.3. *Due codici lineari si dicono* **equivalenti** *se è possibile ottenere tutte le parole di uno a partire dall'altro, applicando una successione delle seguenti operazioni*

1. *Permutare la posizione di due elementi delle parole.*

2. *Moltiplicare le lettere di una posizione all'interno di ogni parola per una lettera non nulla.*

Essendo possibile applicare alla matrice G delle trasformazioni elementari che modificano gli elementi delle righe e delle colonne della matrice senza cambiare il sottospazio da essa generato, a meno di permutazioni sugli elementi di ogni vettore che vi appartiene, segue la dimostrazione della

Proprietà 4.2.1. *Due codici lineari C e C' sono detti equivalenti, se è possibile ottenere la matrice generatrice G di C dalla matrice generatrice G' di C tramite una sequenza di trasformazioni elementari dei seguenti tipi:*

1. *Scambiare due righe.*

2. *Moltiplicare gli elementi di una riga per un elemento non nullo in $\mathbb{F}$.*

3. *Scambiare due colonne.*

4. *Moltiplicare gli elementi di una colonna per un elemento non nullo in $\mathbb{F}$.*

4.3 Codifica e decodifica nei codici lineari

Esistono dei procedimenti standard per la codifica e la decodifica dei codici lineari che possono essere "raffinati"per codici particolari. In questa sezione sono presentati un metodo di codifica ed un metodo di decodifica validi per qualsiasi codice lineare.

4.3.1 Codifica tramite matrice generatrice

Codificare un messaggio di un (r, k)-codice significa associare tramite un algoritmo ad ogni parola di $\mathbb{F}^k$ una parola di $\mathbb{F}^r$. La matrice generatrice G del codice determina un modo computazionalmente efficiente per effettuare la codifica: sia $\mathbf{m} \in \mathbb{F}^k$ messaggio da codificare, allora $\mathbf{m}$ viene codificato nella parola del codice $\mathbf{c}$ mediante il prodotto:

$$\mathbf{c} = \mathbf{m}G = (m_1, m_2, \dots m_k) \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1r} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2r} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & \cdots & a_{kr} \end{pmatrix}$$

E quindi la parola del codice $\mathbf{c}$ generata dal messaggio $\mathbf{m}$ risulta avere componenti:

$$\begin{cases} c_1 = m_1 e_{11} + m_2 e_{21} + \cdots + m_k e_{k1} \\ c_2 = m_2 e_{12} + m_2 e_{22} + \cdots + m_k e_{k2} \\ \quad \vdots \qquad\qquad\qquad\qquad \vdots \\ c_k = m_1 e_{1r} + m_2 e_{2r} + \cdots + m_k e_{kr} \end{cases}$$

Il procedimento descritto associa ad ogni messaggio una ed una sola parola, infatti per definizione di G messaggi diversi diventano parole diverse tramite il prodotto per G.

Se la matrice generatrice è in forma standard allora il messaggio $\mathbf{m} = (m_1, m_2, \dots m_k)$ viene trasformato tramite prodotto con $G = (I_k \mid E)$ nel vettore $\mathbf{m} = (m_1, \dots m_k, c_{k+1}, \dots, c_r)$ dove le lettere $c_{k+1}, \dots, c_r$ sono i simboli di controllo.

4.3.2 Decodifica con la sindrome

Per la decodifica è utile considerare un (r, k)-codice lineare C come un sottogruppo abeliano di ordine q^k di $\mathbb{F}^r$ oltre che come sottospazio vettoriale; in questo modo il sottogruppo C determina la partizione del gruppo $\mathbb{F}^r$ in laterali C_j. Se viene trasmessa la parola codificata $\mathbf{x}$ e viene ricevuta la parola $\mathbf{y} = \mathbf{x} + \mathbf{e}$, allora l'errore è un elemento dello stesso laterale a cui appartiene il vettore ricevuto:

Teorema 4.3.1. *Sia C un (r, k)-codice, sia $\mathbf{x} \in C$ parola codificata ed inviata ed $\mathbf{y} = \mathbf{x} + \mathbf{e}$ parola ricevuta con l'errore $\mathbf{e}$, allora $\mathbf{y}$ appartiene al laterale C_j che è lo stesso laterale al quale appartiene anche $\mathbf{e}$.*

Dimostrazione. Dato che i laterali formano una partizione, esiste sempre un laterale C_j al quale la parola trasmessa $\mathbf{y}$ appartiene. Cioè $\mathbf{y} = \mathbf{a_j} + \mathbf{c}$ dove $\mathbf{a_j}$ è il **leader** di C_j, cioè l'elemento del laterale di peso minimo e $\mathbf{c}$ è una parola del codice C, allora

$$\mathbf{e} = \mathbf{y} - \mathbf{x} = \mathbf{a_j} + \mathbf{c} - \mathbf{x} = \mathbf{a_j} + \mathbf{c}'$$

dove $\mathbf{c}' \in C$, da cui segue che $\mathbf{e} \in C_j$. □

Basandosi su questa idea ed assumendo che l'errore sia di peso minore del rumore di peso massimo riconoscibile il decodificatore che riceve $\mathbf{y}$ individua come errore il leader del laterale di $\mathbf{y}$: $\mathbf{e} = \mathbf{a_j}$. Determina quindi la parola corretta come $\mathbf{y} - \mathbf{a_j}$, che è la parola del codice più vicina alla parola trasmessa:

Teorema 4.3.2. *Sia C un (r,k)-codice, sia $\mathbf{y} \in C_j$ laterale di C avente leader $\mathbf{a_j}$. Allora*

$$\rho(\mathbf{y}, \mathbf{y} - \mathbf{a_j}) \leq \rho(\mathbf{y}, \mathbf{c}) \qquad \forall \mathbf{c} \in C$$

Dimostrazione. Utilizzando la notazione $\mathbf{y} = \mathbf{a_j} + \mathbf{c}'$ dove $\mathbf{a_j}$ è il leader di C_j e $\mathbf{c}'$ è una parola del codice C segue che

$$\rho(\mathbf{y}, \mathbf{c}') = \rho(\mathbf{a_j} + \mathbf{c}', \mathbf{c}') = w(\mathbf{a_j} + \mathbf{c}' - \mathbf{c}') = w(\mathbf{a_j})$$

Cioé $\rho(\mathbf{y}, \mathbf{c}') = w(\mathbf{a_j})$.
Inoltre per ogni parola $\mathbf{c}$ appartenente al codice C

$$\rho(\mathbf{y}, \mathbf{c}) = \rho(\mathbf{a_j} + \mathbf{c}', \mathbf{c}) = w(\mathbf{a_j} + \mathbf{c}' - \mathbf{c})$$

Cioé $w(\mathbf{a_j} + \mathbf{c}' - \mathbf{c}) = \rho(\mathbf{y}, \mathbf{c})$.
La tesi segue allora dal fatto che

$$w(\mathbf{a_j}) \leq w(\mathbf{a_j} + \mathbf{c}' - \mathbf{c})$$

che è una conseguenza della definizione di leader e dal fatto che $\mathbf{a_j} + \mathbf{c}' - \mathbf{c} \in C_j$. □

Il metodo che abbiamo appena esposto per la decodifica di un codice lineare non sfrutta la struttura matriciale del codice. Esiste infatti una procedura più efficiente, conseguenza del fatto che l'appartenenza di un vettore $\mathbf{x} \in \mathbb{F}^r$ ad un laterale di C è determinato univocamente dalla sua sindrome $H\mathbf{x}^t$ (dove H è la matrice di controllo).

Teorema 4.3.3. *Sia C codice lineare di lunghezza r generato dalla matrice G e di matrice di controllo H. Due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{F}^r$ appartengono allo stesso laterale di C se e solo se hanno la stessa sindrome.*

Dimostrazione. I vettori $\mathbf{x}$ ed $\mathbf{y}$ hanno la stessa sindrome se e solo se $H\mathbf{x}^t = H\mathbf{y}^t$, cioè se e solo se $H(\mathbf{x}-\mathbf{y})^t = \mathbf{0}$. Dal teorema 4.2.4 $H(\mathbf{x}-\mathbf{y})^t = \mathbf{0}$ se e solo se $\mathbf{x} - \mathbf{y} \in C$ quindi se e solo se $\mathbf{x}$ ed $\mathbf{y}$ appartengono ad uno stesso laterale di C. □

Dal momento che la sindrome di un vettore appartenente al laterale C_j equivale alla sindrome del leader a_j di tale laterale, si ha una corrispondenza biunivoca fra i leader dei laterali e la loro sindrome. Si delinea quindi un algoritmo di decodifica suddiviso in 3 passi da applicare al vettore $\mathbf{y}$ ricevuto nota H:

1. Si calcola la sindrome $\mathbf{s} = H\mathbf{y}^t$, che è uguale a quella del leader del laterale al quale $\mathbf{y}$ appartiene.

2. Si determina il leader $\mathbf{a_j}$ avente sindrome $\mathbf{s}$.

3. Si decodifica $\mathbf{y}$ con la parola $\mathbf{x} = \mathbf{y} - \mathbf{a_j}$.

4.4 Limitazione di Gilbert-Varshamov per i codici lineari binari

Dai limiti di Hamming e di Singleton sono determinate delle caratteristiche sui parametri del codice, ma non è garantita l'esistenza di un codice avente tali parametri. Il problema generale dell'esistenza non è affrontato in questa tesi e per una soluzione generale si rimanda alla bibliografia. Però se il codice è lineare e binario (cioè definito sull'alfabeto $GF(2)$) allora un modo per risolvere la questione è quella di fornire una condizione di esistenza della matrice di controllo, la quale determina (non univocamente) il codice.
Il teorema che presentiamo, la cui tesi è nota come limitazione di Gilbert-Varshamov, procede in tale direzione.

Teorema 4.4.1. *Un (r,k)-codice lineare binario C con distanza minima d esiste se è verificata la disuguaglianza*

$$\sum_{j=0}^{d-2} \binom{r-1}{j} < 2^{r-k}$$

Dimostrazione. Dimostriamo una tesi equivalente: esiste una matrice di controllo H di dimensione $(r-k)\times r$ tale che $d-1$ sue colonne siano linearmente indipendenti.
Come prima colonna possiamo considerare una qualsiasi non nulla, come seconda possiamo considerare un'altra colonna non nulla ed indipendente dalla prima, fino ad arrivare ad avere $j \leq r-1$ colonne tali che siano linearmente indipendenti da qualsiasi $d-2$ colonne scelte precedentemente. Possiamo aggiungere una $(j+1)$-esima colonna linearmente indipendente con $d-2$ qualsiasi delle precedenti se

$$\binom{j}{1} + \binom{j}{2} + \cdots + \binom{j}{d-2} < 2^{r-k} - 1$$

infatti la possibilità di scegliere un vettore linearmente indipendente da un insiseme di $d-2$ può essere sempre fatta se la somma delle scelte possibili dei vettori precedente è inferiore al numero dei possibili vettori non nulli $(r-k)$-dimensionali. Procedendo su j fino alla $(r-1)$-esima si ha che $d-1$ delle colonne di H sono linearmente indipendenti se vale la tesi. □

Esempio 4.4.1. *Esiste un $(9,2)$-codice avente distanza minima 5, ma non esiste a distanza minima 6. Infatti*

$$\sum_{j=0}^{3} \binom{8}{j} = 93 < 2^7$$

mentre

$$\sum_{j=0}^{4} \binom{8}{j} = 163 > 2^7$$

Il teorema precedente può essere generalizzato al caso in cui il codice non sia binario, ma sia definito sull'alfabeto $\mathbb{F}_q$.

Corollario 4.4.1. *Un (r, k)-codice lineare C definito su $\mathbb{F}_q$ con distanza minima d esiste se è verificata la disuguaglianza*

$$\sum_{j=0}^{d-2}(q-1)^j\binom{r-1}{j} < q^{r-k}$$

Dimostrazione. Segue dal teorema precedente e dal fatto che è possibile scegliere un insieme di $d-2$ colonne linearmente indipendenti da un insieme di j colonne ad un alfabeto di ordine q se

$$(q-1)\binom{j}{1} + (q-1)^2\binom{j}{2} + \cdots + (q-1)^{d-2}\binom{j}{d-2} < q^{r-k} - 1$$

□

Capitolo 5

Codici ciclici

I codici ciclici costituiscono una importante sottoclasse dei codici lineari, con una codifica ed una decodifica particolarmente efficienti. L'idea alla base è quella di poter utilizzare la teoria dei campi finiti sullo spazio delle parole del codice, che in tal modo diventa oltre che spazio vettoriale, anche campo dotato di rappresentazione polinomiale, ed in particolare anello i cui ideali ed idempotenti che abbiamo già introdotto giocano un ruolo essenziale.

5.1 Introduzione

Sia $\mathbb{F}^r$ spazio vettoriale r-dimensionale sul campo finito $\mathbb{F}$ di ordine $q = p^n$ per p primo ed $(r, q) = 1$, affinché nella decomposizione di $x^r - 1$ tutti i polinomi irriducibili abbiamo molteplicità 1.

Definizione 5.1.1. *Un codice lineare C di lunghezza r (cioè sottospazio vettoriale di $\mathbb{F}^r$) si dice* **ciclico** *se è chiuso rispetto alla permutazione ciclica dei suoi elementi verso destra:*

$$\mathbf{c} = (c_0, c_1, \ldots, c_{r-1}) \in C \Longrightarrow (c_{r-1}, c_0, \ldots, c_{r-2}) \in C$$

I codici ciclici possono essere rappresentati utilizzando le algebre viste nel primo capitolo; potremmo considerare C come sottospazio vettoriale di $\mathbb{F}^r$ i cui vettori sono chiusi rispetto allo shifter nel prodotto di convoluzione ($\mathbf{c}$ è una parola del codice allora anche $(0, 1, 0, \ldots, 0) \star \mathbf{c}$ è una parola del codice). Ma possiamo anche considerarli come sottospazi vettoriali di $\mathcal{R}_{r,q}$ chiusi rispetto al prodotto per x (se $f(x)$ è una parola del codice allora anche $xf(x)$ è una parola del codice).
Una caratterizzazione algebrica che permette di sfruttare quanto visto sulla struttura $\mathcal{R}_{r,q}$ è data dal seguente teorema:

Teorema 5.1.1. *Un codice lineare C di lunghezza r sull'alfabeto $\mathbb{F}_q$ è ciclico se e solo se è un ideale di $\mathcal{R}_{r,q}$.*

Dimostrazione. $\Rightarrow$) Sia C codice ciclico di $\mathcal{R}_{r,q}$: se $c(x)$ è una parola di C allora anche $x^k c(x)$ è una parola di C comunque scelto k. Per linearità

tutte le combinazioni lineari di $x^k c(x)$ sono elementi del codice:

$$\lambda_0 c(x) + \lambda_1 x^1 c(x) + \cdots + \lambda_{r-1} x^{r-1} c(x) = \\ = (\lambda_0 + \lambda_1 x^1 + \cdots + \lambda_{r-1} x^{r-1}) c(x) \in C$$

Quindi ogni polinomio della forma $f(x)c(x)$ è un elemento di C.

$\Leftarrow$) Sia C ideale di $\mathcal{R}_{r,q}$. Se $c(x)$ è una parola di C segue che $f(x)c(x) \in C$ comunque scelto $f(x) \in \mathcal{R}_{r,q}$. Quindi a fortiori

$$xc(x) \in C$$

Pertanto C è un codice lineare.

□

Abbiamo stabilito che i codici ciclici sono ideali di $\mathcal{R}_{r,q}$; li indicheremo con $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \ldots$

5.2 Polinomi generatori

Correndo il rischio di essere ripetitivi riformuliamo alcuni teoremi del capitolo 1 nella teoria dei codici correttori.

Teorema 5.2.1. *Sia $\mathfrak{a}$ codice ciclico di lunghezza r, allora ogni elemento $f(x)$ di tale codice può essere scritto come $f(x) = g(x)a(x)$ per qualche $g(x)$ e per $a(x)$ polinomio monico di grado minimo fissato in $\mathfrak{a}$.*

Dimostrazione. Sia per assurdo $f(x) \in \mathfrak{a}$ ed $f(x)$ non multiplo di $a(x)$. Segue che

$$f(x) = q(x)a(x) + r(x)$$

dove il grado di $r(x)$ è positivo e minore del grado di a ed $r(x) = f(x) - q(x)a(x) \in \mathfrak{a}$. Ma questo è in contraddizione con il fatto che $a(x)$ ha grado minimo, pertanto $\mathfrak{a} = (a(x))$. □

Abbiamo quindi ritrovato una conferma del fatto che $\mathcal{R}_{r,q}$ è un anello a ideali principali, cioè ogni ideale e quindi ogni codice lineare può essere identificato semplicemente dal polinomio che lo genera. Per descrivere un codice ciclico quindi non è necessario fornire la matrice generatrice come nel caso dei codici lineari, ma è sufficiente indicare un determinato polinomio.

Definizione 5.2.1. *Dato $\mathfrak{a}$ codice ciclico, un polinomio monico e di grado minimo in $\mathfrak{a}$ è detto* **polinomio generatore**. *Esso genera $\mathfrak{a}$ come ideale nell'anello $\mathcal{R}_{r,q}$.*

Ripetiamo tre proprietà che definiscono i polinomi generatori di un codice ciclico, formulate e dimostrate nei teoremi 2.2.1, 2.2.2 e 2.2.3.

Teorema 5.2.2. *Sia $\mathfrak{a}$ ideale di $\mathcal{R}_{r,q} = \mathbb{F}[x] \Big/ x^r - 1$, cioè codice ciclico di dimensione r. Allora*

1. *Se $a(x)$ è polinomio generatore di $\mathfrak{a}$ allora è un divisore di $x^r - 1$.*

2. *Se $a(x)$ è un divisore monico di $x^r - 1$ allora genera il codice ciclico $\mathfrak{a}$, cioè*

$$C = \{f(x)c(x) \mid f(x) \in \mathcal{R}_{r,q}\}$$

3. *Esiste quindi una corrispondenza biunivoca fra i codici ciclici ed i divisori di $x^r - 1$.*

Possiamo individuare tutti i codici ciclici presenti nella struttura $\mathcal{R}_{r,q}$ valutando la fattorizzazione di $x^r - 1$. Ogni fattore irriducibile di $x^r - 1$ genera quindi un codice ciclico, ma anche ogni prodotto di una scelta di tali fattori genera un codice ciclico essendo ancora un fattore di $x^r - 1$. La seguente definizione ha lo scopo di sottolineare tale distinzione:

Definizione 5.2.2. *I codici ciclici di $\mathcal{R}_{r,q}$ generati dai fattori irriducibili di $x^r - 1$ sono detti codici* **primitivi**, *mentre quelli generati dal prodotto di almeno due dei fattori irriducibili non triviali sono detti codici* **composti**.

Dal corollario 2.2.2 segue che

Corollario 5.2.1. *Sia h il numero dei fattori irriducibili di $x^r - 1$, allora i codici ciclici in $\mathcal{R}_{r,q}$ sono 2^h.*

Se si vogliono escludere i codici generati dai divisori banali 1 ed $x^r - 1$, che sono inutili nella pratica essendo rispettivamente i sottospazi impropri $\mathcal{R}_{r,q}$ e 0, allora $\mathcal{R}_{r,q}$ possiede $2^h - 2$ codici ciclici non banali. Fin ora abbiamo dimostrato che ogni codice ciclico di lunghezza r è un ideale di $\mathcal{R}_{r,q}$ univocamente determinato da un polinomio particolare, detto polinomio generatore, costruito come prodotto dei divisori irriducibili di $x^r - 1$. Rimane da dimostrare che anche ogni parola di un codice ciclico è univocamente determinata come prodotto di un polinomio per il polinomio generatore.

Teorema 5.2.3. *Sia $\mathfrak{a}$ codice ciclico di lunghezza r generato dal polinomio monico $a(x)$ di grado $r - k$, allora ogni parola di $\mathfrak{a}$ può essere rappresentata in modo unico come il prodotto $f(x)a(x)$, per $f(x) \in \mathcal{R}_{r,q}$ di grado minore o uguale a $k - 1$.*

Dimostrazione. Per definizione di codice ciclico ogni parola è un elemento dell'ideale $\mathfrak{a}$ quindi è in particolare della forma $f(x)a(x)$. La dimostrazione è suddivisa in due parti:

- Se $deg(f(x)) \leq k - 1$ e la parola del codice $t(x)$ può essere rappresentata con due scritture equivalenti

$$t(x) = f(x)a(x) = g(x)a(x)$$

  allora $f(x)a(x)$ ha grado minore ad r ed $(f(x) - g(x))a(x) = 0$ modulo $x^r - 1$. Quindi $f(x) = g(x)$.

- Se $deg(f(x)) \geq k$ allora il caso è meno immediato del precedente considerando la struttura di $\mathcal{R}$. Sia $t(x) = f(x)a(x)$ polinomio che quozientato per $x^r - 1$ rappresenta una parola del codice $\mathfrak{a}$. Dato che $deg(f(x)a(x)) \geq r$ si considera divisione di $t(x)$ per $x^r - 1$

$$t(x) = f(x)a(x) = q(x)(x^r - 1) + r(x)$$

dove $r(x)$ è nullo oppure ha grado minore o uguale ad $r-1$. Ma $a(x)$ è un divisore di x^r-1 quindi dall'equazione precedente $a(x)$ divide $r(x)$:

$$r(x) = s(x)a(x)$$

dove il grado di $s(x)$ è minore o uguale a $k-1$. Si può dunque scrivere

$$t(x) = f(x)a(x) = q(x)(x^r-1) + s(x)a(x)$$

che in $\mathcal{R}$ diventa $t(x) = s(x)a(x)$ e dal caso precedente $s(x)$ è unicamente determinato, avendo grado minore o uguale a $k-1$.

□

5.3 Matrice generatrice di un codice ciclico

I codici ciclici sono in particolare codici lineari, quindi si possono applicare i metodi di codifica e di decodifica presentati nel capitolo precedente. Per poter applicare tali metodi è però necessario risalire alla matrice generatrice dal polinomio generatore del codice. In questo paragrafo, continuando a seguire [2], si presentano due metodi per determinare la matrice cercata.
Il **primo metodo** è conseguenza del corollario 2.2.4:

Teorema 5.3.1. *Sia* $\mathfrak{a}$ *codice ciclico di lunghezza* r*, generato dal polinomio monico* $a(x)$ *di grado* $r-k$

$$a(x) = a_0 + a_1x^1 + \cdots + a_{r-k-1}x^{r-k-1} + x^{r-k}$$

Allora una possibile matrice generatrice del codice $\mathfrak{a}$ *è data da*

$$G = \begin{pmatrix} a_0 & a_1 & \dots & a_{r-k-1} & 1 & 0 & \dots & 0 \\ 0 & a_0 & a_1 & \dots & a_{r-k-1} & 1 & \dots & 0 \\ & \ddots & & \ddots & & \ddots & & \\ 0 & \dots & 0 & a_0 & a_1 & \dots & a_{r-k-1} & 1 \end{pmatrix}$$

Oltre a fornire la matrice generatrice e quindi a consentire l'applicazione dei metodi di codifica e di decodifica, il teorema precedente afferma che i codici ciclici generati da un polinomio di grado $r-k$ sono (r,k)-codici.
Il **secondo metodo** per determinare la matrice generatrice di un codice ciclico a partire dal polinomio generatore $a(x)$ utilizza i resti delle divisioni dei polinomi x^j per $a(x)$. Se il grado del polinomio generatore è uguale ad $r-k$ allora dalle divisioni

$$x^j = q_j(x)a(x) + r_j(x) \qquad j = r-k, \dots, r-1$$

si ottengono k resti nulli o di grado minore di $r-k$

$$r_j(x) = r_{0,j} + r_{1,j}x^1 + r_{2,j}x^2 + \cdots + r_{r-k,j}x^{r-k} \qquad j = r-k, \dots, r-1$$

Considerando i polinomi $\{x^j - r_j(x)\}_{j=r-k}^{r-1}$ come le righe dei una matrice G, si ottiene

$$G = \begin{pmatrix} -r_{0,r-k} & -r_{1,r-k} & \dots & -r_{r-k-1,r-k} & 1 & 0 & \dots & 0 \\ -r_{0,r-k+1} & -r_{1,r-k+1} & \dots & -r_{r-k-1,r-k+1} & 0 & 1 & \dots & 0 \\ \vdots & & & \vdots & \vdots & & \ddots & \vdots \\ -r_{0,r-1} & -r_{1,r-1} & \dots & -r_{r-k-1,r-1} & 0 & \dots & 0 & 1 \end{pmatrix}$$

Rimane da dimostrare che G è effettivamente una matrice generatrice del codice $\mathfrak{a}$:

Teorema 5.3.2. *La matrice G precedentemente definita è una matrice generatrice del codice $\mathfrak{a}$.*

Dimostrazione. Le righe sono parole del codice e la matrice G ha caratteristica k, dato che contiene la matrice identità, quindi i vettori corrispondenti ai polinomi $\{x^j - r_j(x)\}_{j=r-k}^{r-1}$ sono linearmente indipendenti e sono generatori del codice $\mathfrak{a}$. □

5.4 Polinomi e matrici di controllo

Oltre al polinomio ed alla matrice generatrice, si può associare ad un codice ciclico un polinomio ed una matrice di controllo. Per questo scopo si utilizzerà la proprietà che il polinomio generatore $a(x)$ è un divisore di $x^r - 1$.

Definizione 5.4.1. *Sia $\mathfrak{a}$ codice ciclico di dimensione r generato dal polinomio $a(x)$ di grado $r-k$. Il polinomio monico $h(x)$ di grado k tale che $a(x)h(x) = x^r - 1$ è detto* **polinomio di controllo** *del codice $\mathfrak{a}$.*

Il polinomio di controllo, già introdotto precedentemente con la notazione $\hat{a}(x)$, stabilisce un criterio di appartenenza di una parola ad un codice in modo analogo alla sindrome nei codici lineari:

Teorema 5.4.1. *Sia $\mathfrak{a}$ codice ciclico di dimensione r generato dal polinomio $a(x)$ ed avente polinomio di controllo $h(x)$. Allora la parola $t(x)$ appartiene al codice $\mathfrak{a}$ se e solo se $t(x)h(x) = 0$ in $\mathcal{R}$*

Dimostrazione. $\Rightarrow$) Se $t(x)$ è una parola di $\mathfrak{a}$ allora può essere scritto come $t(x) = f(x)a(x)$ per qualche $f(x) \in \mathcal{R}$. Moltiplicando ambo i membri per il polinomio di controllo si ottiene la tesi:

$$t(x)h(x) = t(x)a(x)h(x) = 0 \in \mathcal{R}$$

$\Leftarrow$) Se $t(x)h(x) = 0$ allora in $\mathbb{F}[x]$

$$t(x)h(x) = q(x)(x^t - 1) = q(x)a(x)h(x) \in \mathbb{F}[x]$$

Dividendo primo ed ultimo membro per $h(x)$ si ottiene la tesi.

□

Il polinomio di controllo $h(x)$ di un codice $\mathfrak{a}$ non è in generale un generatore del codice duale $\mathfrak{a}^\perp$, dal fatto che se $a(x)h(x) = 0$ allora non necessariamente i vettori associati ai polinomi $a(x)$ ed $h(x)$ sono ortogonali. Però i coefficienti di $h(x)$ ed il polinomio che genera $\mathfrak{a}^\perp$ non sono del tutto scorrelati:

Teorema 5.4.2. *Sia $\mathfrak{a}$ un $(r, r-k)$-codice ciclico generato da (x), ed avente come polinomio di controllo*

$$h(x) = h_0 + h_1 x^1 + \cdots + h_{k-1}x^{k-1} + x^k$$

Allora la matrice

$$H = \begin{pmatrix} 1 & h_{k-1} & \dots & h_1 & h_0 & 0 & \dots & 0 \\ 0 & 1 & h_{k-1} & \dots & h_1 & h_0 & \dots & 0 \\ & \ddots & & \ddots & & \ddots & & \\ 0 & \dots & 0 & 1 & h_{k-1} & \dots & h_1 & h_0 \end{pmatrix}$$

è una matrice di controllo per $\mathfrak{a}$.
Inoltre il polinomio reciproco di $h(x)$*, dato da*

$$h(x)^{\perp} = 1 + h_{k-1}x^1 + \cdots + h_1 x^{k-1} + h_0 x^k$$

genera il codice duale $\mathfrak{a}^{\perp}$.

Dimostrazione. Dal teorema 5.4.1 una parola $a(x)$ è in $\mathfrak{a} = (h(x))$ se e solo se $a(x)h(x) = 0 \mod x^r - 1$ cioè se e solo se tutti i coefficienti del polinomio prodotto sono nulli. Questo accade se e solo se, per $1 = h_k$ valgono le equazioni

$$\begin{aligned} \sum_{j=0}^{k} a_j h_{k-j} &= 0 \\ \sum_{j=0}^{k} a_{j+1} h_{k-j} &= 0 \\ \sum_{j=0}^{k} a_{j+2} h_{k-j} &= 0 \\ &\vdots \\ \sum_{j=0}^{k} a_{j+r-k-1} h_{k-j} &= 0 \end{aligned}$$

che equivale al prodotto

$$H \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix}$$

Quindi le righe di H sono ortogonali a $\mathbf{a}$ e sono quindi parole del duale di $\mathfrak{a}$.
La seconda parte del teorema segue dal fatto che, per 2.1.6 $h(x)^{\perp}$ è un divisore di $x^r - 1$ e che dal teorema 5.3.1 H è una matrice generatrice del codice $(h(x)^{\perp})$, quindi è di controllo di $\mathfrak{a}$ e generatrice di $\mathfrak{a}^{\perp}$. □

5.5 Codifica e decodifica dei codici ciclici

Dal fatto che i codici ciclici sono particolari codici lineari, la codifica e la decodifica possono avvenire tramite i sistemi introdotti nel capitolo precedente. Ha senso chiedersi però se ci sono dei metodi di codifica e decodifica che permettano di sfruttare la struttura algebrica di cui i codici ciclici sono dotati.

Sia $\mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}$ un (r,k)-codice ciclico generato dal polinomio $a(x)$. Allora una codifica standard del messaggio $\mathbf{m}$ nella rappresentazione polinomiale $m(x)$, può essere data direttamente dalla moltiplicazione per $a(x)$, infatti $a(x)m(x)$ equivale al prodotto $G\mathbf{m}^t$ per G matrice generatrice del codice. La decodifica invece può variare a seconda della tipologia di codice ciclico in questione (BCH, Reed-Solomon...). Nel prossimo paragrafo presentiamo i codici BCH con la decodifica Peterson-Gorenstein-Zierler.

5.6 Codici BCH

Durante la ricerca di un codice lineare che avesse una codifica ed una decodifica ottimizzate rispetto ai codici ciclici generici, si è arrivati alla creazione di diversi tipi di codici che generalmente portano i nomi dei loro scopritori. I codici BCH sono stati scoperti da Hocquenghem nel 1959, ed indipendentemente da Bose e Chaudhuri nel 1960. La possibilità di implementare facilmente un decodificatore e il fatto che la lunghezza delle parole dei BCH possa essere scelta fra un range abbastanza ampio li ha resi ottimali per la correzione degli errori di lettura dei CD, dei DVD, delle memorie Flash[1], dei codici a barra e in passato per la comunicazione satellitare[2].
Dopo una presentazione della definizione, giustificata da alcuni importanti risultati teorici e dopo alcuni esempi, proponiamo due sistemi di decodifica. Accenniamo infine ai codici di Reed-Solomon scoperti nel 1960 ed individuati come sottoclasse dei BCH codici nel 1961 (Gorenstein e Zierler)[3]
L'idea di fondo è quella di costruire il generatore di un codice ciclico, facendo in modo che la distanza minima sia maggiore di un intero δ prefissato. Per fare ciò si sceglie una radice primitiva r-esima dell'unità ξ e si considera il più piccolo divisore di $x^r - 1$ che ha come radice δ potenze consecutive di ξ.
Questo procedimento è giustificato dal seguente teorema[4]

Teorema 5.6.1. *Sia $a(x)$ polinomio generatore di un (r,k)-codice ciclico $\mathfrak{a}$ ideale di $\mathcal{R}_{r,q}$, e siano $\{\xi_1, \ldots, \xi_{r-k}\}$ radici di $a(x)$ nel suo campo di spezzamento $\mathbb{F}_{q^m}$ per m periodo di q in $\mathbb{Z}_r^\star$. Un polinomio $c(x)$ in $\mathbb{F}_q[x]$ è una parola di $\mathfrak{a}$ se e solo se $c(\xi_j) = 0$ per ogni $j = 1, \ldots, r-k$.*

Dimostrazione. $\Rightarrow$) Sia $c(x) \in \mathfrak{a}$, allora è sufficiente osservare che $c(x) = q(x)a(x)$ per qualche $q(x)$ in $\mathbb{F}_q[x]$.

$\Leftarrow$) Siano $M_j(x)$ polinomi minimi di ξ_j per ogni $j = 1, \ldots, r-k$. Segue che ogni $M_j(x)$ divide $c(x)$, e quindi $a(x)$ è un divisore di $c(x)$:

$$c(x) = q(x) \prod_{j=1}^{r-k} M_j(x) = q(x)a(x)$$

da cui segue che $c(x) \in \mathfrak{a}$.

□

[1] ad esempio la SSD SandForce SF-2600.
[2] Nell'articolo [9] sono riportati i dati dell'utilizzo dei codici BCH in previsione di una missione su Phobos, luna di Marte, proposta nel 1988 dall'unione sovietica e mai realizzata.
[3] Per la bibliogafia dettagliata, con i riferimenti agli articoli originali si rimanda a [4].
[4] Da [2] pag. 217.

Il teorema appena dimostrato possiede una formulazione matriciale, nella quale le parole sono rappresentate da vettori invece che da polinomi:

Teorema 5.6.2. *Sia $a(x)$ polinomio generatore di un (r,k)-codice ciclico $\mathfrak{a}$ ideale di $\mathcal{R}_{r,q}$, e siano $\{\xi_1,\dots,\xi_{r-k}\}$ radici di $a(x)$ nel suo campo di spezzamento $\mathbb{F}_{q^m}$ per m periodo di q in $\mathbb{Z}_r^\star$. Sia K matrice $(r-k)\times r$ ad elementi in $\mathbb{F}_{q^m}$ definita come*

$$K = \begin{pmatrix} 1 & \xi_1 & \xi_1^2 & \dots & \xi_1^{r-1} \\ 1 & \xi_2 & \xi_2^2 & \dots & \xi_2^{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \xi_{r-k} & \xi_{r-k}^2 & \dots & \xi_{r-k}^{r-1} \end{pmatrix}$$

Un polinomio $c(x)$ avente vettore associato $\mathbf{c} = c(x)$ è una parola del codice $\mathfrak{a}$ se e solo se

$$K\mathbf{c}^t = \mathbf{0}$$

Dimostrazione. È sufficiente osservare che

$$K\mathbf{c}^t = \begin{pmatrix} c(\xi_1) \\ c(\xi_2) \\ \vdots \\ c(\xi_{r-k}) \end{pmatrix}$$

dal fatto che $c(x) = \sum_{j=0}^{r-1} c_j x^j$. □

Osserviamo che le righe della matrice K sono vettori in $\mathbb{F}_{q^m}$ ortogonali ad ogni parola del codice $\mathfrak{a}$ la cui definizione è sconseguenza diretta della rappresentazione polinomiale del codice. Inoltre ha un comportamento simile a quello della matrice di controllo H ma è fondamentalmente differente.
Nei teoremi precedenti siamo sempre partiti da $a(x)$ e poi abbiamo considerato le sue radici. Partiamo questa volta con il definire $a(x)$ direttamente dalle sue radici. Questa idea fornisce direttamente un limite alla distanza del codice e costituisce la base per la definizione di codice BCH[5].

Teorema 5.6.3. *Sia ξ radice primitiva r-esima e siano $\{\xi,\xi^2,\dots,\xi^{\delta-1}\}$ $\delta-1$ potenze consecutive di ξ distinte. Sia $a(x)$ il più piccolo polinomio in $\mathbb{F}_q[x]$ ad avere $\{\xi,\xi^2,\dots,\xi^{\delta-1}\}$ come radici. Allora il codice generato da $a(x)$ ha come distanza minima al più δ.*

Dimostrazione. Osserviamo che l'enunciato del teorema è sensato dato che $a(x)$, per avere ξ^j fra le sue radici, dovrà avere $M^{(j)}(x)$ fra i suoi fattori, quindi è costituito dal minimo comune multiplo di una scelta di polinomi minimi e quindi è un divisore di x^r-1. Essendo un divisore genera un ideale ed è pertanto un codice ciclico.
Sia allora $(a(x)) = \mathfrak{a}$. Sfruttando il teorema precedente, possiamo definire la matrice K, **detta matrice estesa di controllo** sull'insieme $\{\xi,\xi^2,\dots,\xi^{\delta-1}\}$

[5]Teorema 8, pag. 201 [24], Teorema 10.5 pag 220 [2].

che diventa

$$K = \begin{pmatrix} 1 & \xi & \xi^2 & \dots & \xi^{r-1} \\ 1 & \xi^2 & \xi^4 & \dots & \xi^{2(r-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \xi^{\delta-1} & \xi^{2(\delta-1)} & \dots & \xi^{(r-1)(\delta-1)} \end{pmatrix}$$

che in una forma più compatta può essere scritta come

$$K = (\xi^{jk})_{j=1,k=0}^{\delta-1,r-1}$$

Sia per assurdo $\mathbf{c}$ parola di $\mathfrak{a}$ con distanza minima strettamente inferiore a δ: $w(\mathbf{c}) = w < \delta$. Segue che esistono w elementi di $\mathbf{c}$ diversi da zero:

$$c_j \neq 0 \quad \forall j \in \{i_1, \dots, i_w\} \subseteq \{0, 1, \dots, r-1\}$$

Dato che $K\mathbf{c}^t = \mathbf{0}$ implica che:

$$\begin{pmatrix} \xi^{i_1} & \xi^{i_2} & \dots & \xi^{i_w} \\ \xi^{2i_1} & \xi^{2i_2} & \dots & \xi^{2i_w} \\ \vdots & \vdots & & \vdots \\ \xi^{(w-1)i_1} & \xi^{(w-1)i_2} & \dots & \xi^{(w-1)i_w} \end{pmatrix} \begin{pmatrix} c_{i_1} \\ c_{i_2} \\ \vdots \\ c_{i_w} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Quindi il determinante della matrice ricavata considerando solo le colonne di pedici $i_1, \dots, i_w$ in K risulta avere determinante uguale a zero. Quindi

$$det \begin{pmatrix} \xi^{i_1} & \xi^{i_2} & \dots & \xi^{i_w} \\ \xi^{2i_1} & \xi^{2i_2} & \dots & \xi^{2i_w} \\ \vdots & \vdots & & \vdots \\ \xi^{(w-1)i_1} & \xi^{(w-1)i_2} & \dots & \xi^{(w-1)i_w} \end{pmatrix} =$$

$$\xi^{i_1+i_2+\dots+i_w} det \begin{pmatrix} 1 & 1 & \dots & 1 \\ \xi^{2i_1} & \xi^{2i_2} & \dots & \xi^{2i_w} \\ \vdots & \vdots & & \vdots \\ \xi^{(w-1)i_1} & \xi^{(w-1)i_2} & \dots & \xi^{(w-1)i_w} \end{pmatrix} = 0$$

che è una contraddizione, dato che il determinante di una matrice di Vandermonde non può essere uguale a zero[6]. □

La dimostrazione appena presentata equivale a dimostrare che è sempre possibile scegliere un insieme di δ o meno colonne linearmente indipendenti nella matrice K.
Avremmo potuto enunciare il teorema precedente scegliendo anziché ξ, già una potenza fissata di ξ. Può quindi essere riformulato come:

Corollario 5.6.1. *Sia ξ radice primitiva r-esima e sia b intero positivo. Siano $\{\xi^b, \xi^{b+1}, \dots, \xi^{b+\delta-2}\}$ $\delta - 1$ potenze consecutive di ξ^b distinte. Sia $a(x)$ il più piccolo polinomio in $\mathbb{F}_q[x]$ ad avere $\{\xi^b, \xi^{b+1}, \dots, \xi^{b+\delta-2}\}$ come radici. Allora il codice generato da $a(x)$ ha come distanza minima al più δ.*

[6] Ad esempio lemma 17 pag 116 [24].

Dimostrazione. Segue dal teorema 5.6.3 considerando la matrice K definita sull'insieme $\{\xi^b, \xi^{b+1}, \ldots, \xi^{b+\delta-2}\}$:

$$K = \begin{pmatrix} 1 & \xi^b & \xi^{2b} & \ldots & \xi^{(r-1)b} \\ 1 & \xi^{b+1} & \xi^{2(b+1)} & \ldots & \xi^{(r-1)(b+1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \xi^{b+\delta-2} & \xi^{2(b+\delta-2)} & \ldots & \xi^{(r-1)(b+\delta-2)} \end{pmatrix}$$

La dimostrazione prosegue in modo analogo a quella del teorema 5.6.3 già valida per $b = 1$. □

5.6.1 Definizione dei codici BCH ed esempi

Con i risultati fin qui introdotti è possiamo definire i codici BCH.

Definizione 5.6.1. *Dato un intero positivo δ ed ξ radice primitiva r-esima dell'unità, un codice ciclico di lunghezza r sull'alfabeto $\mathbb{F}_q$ è detto BCH codice se il suo polinomio generatore è il minimo comune multiplo dei polinomi minimi di $\{\xi^b, \xi^{b+1}, \ldots, \xi^{b+\delta-2}\}$ per b intero positivo.*

Come predetto inizialmente, il numero δ scelto per definire il codice BCH $\mathfrak{a}$ è la più piccola distanza minima che il codice può avere, quindi dalla scelta dell'insieme $\{\xi^b, \xi^{b+1}, \ldots, \xi^{b+\delta-2}\}$ si avranno codici differenti nei quali un limite alla distanza minima può essere dato a priori.
Quindi il parametro δ può essere chiamato **distanza minima garantita** del codice BCH. Rimangono da determinare i parametri del codice

Teorema 5.6.4. *Un BCH codice $\mathfrak{a}$ in $\mathcal{R}_{r,q}$ con distanza minima garantita δ ha dimensione pari almeno a $r - m(\delta - 1)$, per m periodo di q in $\mathbb{Z}_r^\star$.*

Dimostrazione. Come sottospazio vettoriale, la dimensione di $\mathfrak{a}$ può essere data da r meno il numero di righe linearmente indipendenti della matrice di controllo H. È possibile ricavare la matrice di controllo direttamente da K, sostituendo ogni suo elemento, che appartiene al campo $\mathbb{F}_{q^m}$, con la m-upla corrispondente. In questo modo H ha dimensione $m(\delta-1) \times r$ ed ha quindi $m(\delta-1)$ righe non necessariamente linearmente indipendenti. □

Dato che il codice è ciclico si può utilizzare il metodo generico per la codifica dei codici ciclici. Per la decodifica rimane sempre valido il metodo generico ma abbiamo a disposizione alcuni strumenti che la rendono più efficiente. Prima di presentarli, proponiamo alcuni esempi.

Esempio 5.6.1. *Consideriamo l'algebra $\mathcal{R}_{15,2}$. Il periodo di 2 in $\mathbb{Z}_{15}^\star$ è 4, quindi ξ radice primitiva r-esima dell'unità è un generatore di $\mathbb{F}_{2^4}$. Il sottogruppo di $\mathbb{Z}_{15}^\star$ isomorfo al gruppo $Gal(\mathbb{F}_2(\xi), \mathbb{F}_2)$, come insieme è dato da $G = \{1, 2, 4, 8\}$ e definisce le classi ciclotomiche come orbite dell'azione di G su $\mathbb{Z}_r$ date da:*

$$\begin{aligned} C_0 &= \{0\} \\ C_1 &= \{1, 2, 4, 8\} \\ C_3 &= \{3, 6, 9, 12\} \\ C_5 &= \{5, 10\} \\ C_7 &= \{7, 11, 13, 14\} \end{aligned}$$

L'insieme delle etichette è dato quindi da $\mathscr{L} = \{0, 1, 3, 5, 7\}$ *ed* $l = 5$. *I polinomi minimi di* ξ^b *per* $b \in \mathscr{L}$ *sono dati quindi da*

$$\begin{aligned}
M^{(0)}(x) &= (x - \xi^0) = x + 1 \\
M^{(1)}(x) &= (x - \xi^1)(x - \xi^2)(x - \xi^4)(x - \xi^8) = x^4 + x + 1 \\
M^{(3)}(x) &= (x - \xi^3)(x - \xi^6)(x - \xi^9)(x - \xi^{12}) = x^4 + x^3 + x^2 + 1 \\
M^{(5)}(x) &= (x - \xi^5)(x - \xi^{10}) = x^2 + x + 1 \\
M^{(7)}(x) &= (x - \xi^7)(x - \xi^{11})(x - \xi^{13})(x - \xi^{14}) = x^4 + x^3 + 1
\end{aligned}$$

che costituiscono la fattorizzazione di $x^r + 1$ *e il prodotto degli elementi di ciascuno dei loro* 2^l *sottoinsiemi determina il generatore di un ideale.*
Scegliendo $\delta = 5$ *come distanza minima garantita ed* $\{\xi, \xi^2, \xi^3, \xi^4\}$ *insieme di radici consecutive, si ottiene* $a(x) = M^{(1)}(x)M^{(3)}(x)$ *come generatore del BCH codice. Dato che* $a(x) = x^8 + x^7 + x^6 + x^4 + 1$ *ha grado* 8 *e peso di Hamming* 5, *allora* $(a(x))$ *è un* $(15, 7)$*-codice con distanza minima pari a* 5 *coincidente con* δ.

Esempio 5.6.2. *Se nell'esempio precedente scegliamo* $\delta = 7$ *come distanza minima garantita ed* $\{\xi^9, \xi^{10}, \xi^{11}, \xi^{12}, \xi^{13}, \xi^{14}\}$ *come insieme di radici consecutive su cui costruire il BCH codice, allora*

$$a(x) = M^{(3)}(x)M^{(5)}(x)M^{(7)}(x) = x^{10} + x^8 + x^7 + x^6 + x^5 + x^2 + 1$$

quindi $(a(x))$ *è un* $(15, 5)$*-codice con distanza minima paria a* 7 *coincidente con* δ.

Esempio 5.6.3. *Consideriamo* $\mathcal{R}_{8,3}$. *Il periodo di* 3 *in* $\mathbb{Z}_8^\star$ *è* 2, *quindi* ξ *radice primitiva* r*-esima dell'unità è un generatore di* $\mathbb{F}_{3^2}$. *Il sottogruppo di* $\mathbb{Z}_8^\star$ *isomorfo al gruppo* $Gal(\mathbb{F}_3(\xi), \mathbb{F}_3)$, *come insieme è dato da* $G = \{1, 3\}$ *e definisce le classi ciclotomiche come orbite dell'azione di* G *su* $\mathbb{Z}_8$ *date da:*

$$\begin{aligned}
C_0 &= \{0\} \\
C_1 &= \{1, 3\} \\
C_2 &= \{2, 6\} \\
C_4 &= \{4\} \\
C_5 &= \{5, 7\}
\end{aligned}$$

L'insieme delle etichette è dato quindi da $\mathscr{L} = \{0, 1, 2, 4, 5\}$ *ed* $l = 5$. *I polinomi minimi di* ξ^b *per* $b \in \mathscr{L}$ *sono*

$$\begin{aligned}
M^{(0)}(x) &= (x - \xi^0) \\
M^{(1)}(x) &= (x - \xi^1)(x - \xi^3) \\
M^{(2)}(x) &= (x - \xi^2)(x - \xi^6) \\
M^{(4)}(x) &= (x - \xi^4) \\
M^{(5)}(x) &= (x - \xi^5)(x - \xi^7)
\end{aligned}$$

Scegliendo $\delta = 4$ *come distanza minima garantita ed* $\{\xi^5, \xi^6, \xi^7\}$ *come insieme di radici consecutive su cui costruire il BCH codice, allora si ottiene* $a(x) = M^{(2)}(x)M^{(5)}(x)$.

5.6.2 Decodifica Peterson-Gorenstein-Zierler

Sia $\mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}$ codice BCH con δ distanza minima garantita e $\{\xi^b, \xi^{b+1}, \dots, \xi^{b+\delta-2}\}$ insieme delle radici primitive r-esime che lo definiscono. Sia $c(x) \in \mathfrak{a}$ messaggio inviato e $v(x)$ parola ricevuta che durante la trasmissione è stata sommata ad un errore:

$$v(x) = c(x) + e(x) \qquad e(x) = \sum_{j \in \mathbb{Z}_r} e_j x^j$$

Consideriamo i tutti i coefficienti di $e(x)$ nulli tranne un numero minore o uguale a t, dove t è un intero positivo fissato, minore di r

$$e_{j_k} \neq 0 \qquad j_1 \leq j_2 \leq \cdots \leq j_w \qquad 0 \leq w \leq t$$

cioè

$$e(x) = \sum_{k=1}^{w} e_{j_k} x^{j_k}$$

Lo scopo di un decodificatore è individuare $e(x)$ e risalire alla parola inviata sottraendo $e(x)$ alla parola ricevuta.
L'informazione di partenza per raggiungere questo scopo è data dal fatto che

$$v(\xi^k) = c(\xi^k) + e(\xi^k) = e(\xi^k)$$

Da cui segue la definizione di sindrome.

Definizione 5.6.2. *L'elemento di $\mathbb{F}_{q^m}$ dato da*

$$S_{b+j} := e(\xi^{j+k}) \qquad j = b+1, b+2, \dots, b+\delta-2$$

è detta **sindrome k-esima** *del polinomio $v(x)$.*

Abbiamo quindi

$$\begin{aligned}
S_b &= e(\xi^b) = \sum_{k=1}^{w} e_{j_k} \xi^{b j_k} \\
S_{b+1} &= e(\xi^{b+1}) = \sum_{k=1}^{w} e_{j_k} \xi^{(b+1) j_k} \\
S_{b+2} &= e(\xi^{b+2}) = \sum_{k=1}^{w} e_{j_k} \xi^{(b+2) j_k} \\
&\vdots \qquad\qquad \vdots \\
S_{b+\delta-2} &= e(\xi^{b+\delta-2}) = \sum_{k=1}^{w} e_{j_k} \xi^{(b+\delta-2) j_k}
\end{aligned}$$

La scelta del nome sindrome non è casuale, infatti per $v(x) = \mathbf{v}$ e per K matrice estesa di controllo di $\mathfrak{a}$ abbiamo che

$$\begin{aligned}
K\mathbf{v}^t &= K(\mathbf{c} + \mathbf{e})^t = K\mathbf{c}^t + K\mathbf{e}^t = \\
&= K\mathbf{e}^t = (e(\xi^b), e(\xi^{b+1}), \dots, e(\xi^{b+\delta-2}))^t = \\
&= (S_b, S_{b+1}, \dots, S_{b+\delta-2})^t
\end{aligned}$$

Dato che vogliamo concentrare la nostra attenzione solo sugli elementi non nulli del polinomio $e(x)$, effettuiamo una modifica nella notazione che porterà a considerare la ricerca dell'errore come la ricerca dei valori di un insieme di coppie della forma $(X_k, Y_k) \in \mathbb{F}_{q^m} \times \mathbb{F}_q$ per $k = 1, \ldots, w$.
Indichiamo

$$X_k := \xi^{j_k} \qquad k = 1, \ldots, w$$

ed

$$Y_k := e_{j_k} \qquad k = 1, \ldots, w$$

così da avere

$$\begin{aligned}
S_b &= e(\xi^b) = \sum_{k=1}^{w} e_{j_k} \xi^{b j_k} = \sum_{k=1}^{w} Y_k X_k^b \\
S_{b+1} &= \sum_{k=1}^{w} Y_k X_k^{b+1} \\
S_{b+2} &= \sum_{k=1}^{w} Y_k X_k^{b+2} \\
&\vdots \\
S_{b+\delta-2} &= \sum_{k=1}^{w} Y_k X_k^{b+\delta-2}
\end{aligned}$$

Le X_k forniscono la posizione dell'errore all'interno del polinomio e sono dette **locatori dell'errore** mentre le Y_k forniscono il valore dell'errore nella posizione j_k-eisma e sono dette **magnitudini dell'errore**.
Possiamo osservare che l'errore è univocamente determinato dalle coppie (X_k, Y_k) ed in generale ogni metodo che permette di risolvere il sistema di equazioni non lineari

$$S_{b+k} = \sum_{j=1}^{w} Y_j X_j^{b+k} \qquad k = 0, 1, \ldots, b + \delta - 2 \tag{5.1}$$

fornisce una decodifica per i codici BCH. In questo paragrafo presentiamo la decodifica Peterson-Gorenstein-Zierler[7] che prevede di trovare le informazioni necessarie per determinare w e le X_k mediante la definizione di un polinomio intermedio.

Definizione 5.6.3. *Si definisce* **polinomio locatore** *degli errori il polinomio*

$$\lambda(x) = \prod_{k=1}^{w} (1 - X_k x) = \sum_{j=1}^{w} \lambda_j x^j$$

dove $\lambda_0 = 1$ *e le cui radici sono* $x = 1/X_k$.

[7]Per i titoli originali degli articoli nei quali è stata presentata per la prima volta si rimanda a [4] nota di pag. 206 e referenze di pag. 468.

Le radici del polinomio locatore degli errori sono i reciproci dei locatori degli errori e i cui coefficienti sono noti grazie al prossimo teorema.

Teorema 5.6.5. *Sia $\mathfrak{a} \trianglelefteq \mathcal{R}_{r,q}$ codice BCH con δ distanza minima garantita e $\{\xi^b, \xi^{b+1}, \ldots, \xi^{b+\delta-2}\}$ insieme delle radici primitive r-esime che lo definiscono. Indicando con $\{X_1, X_2, \ldots, X_w\}$ i locatori dell'errore incogniti, allora è possibile ricavare i coefficienti del polinomio locatore degli errori*

$$\lambda(x) = \prod_{k=1}^{w}(1 - X_k x) = \sum_{j=1}^{w} \lambda_j x^j$$

Prima di procedere alla dimostrazione anticipiamo un risultato[8] centrale nell'algoritmo di decodifica in quanto utilizzato per determinare il grado di $\lambda(x)$ (cioè il numero di errori che entrano nella parola $c(x)$ durante la trasmissione) che non è noto a priori.

Lemma 5.6.1. *La* **matrice delle sindromi** *definita da*

$$M = \begin{pmatrix} S_b & S_{b+1} & S_{b+2} & \ldots & S_{b+u-1} \\ S_{b+1} & S_{b+2} & S_{b+3} & \ldots & S_{b+u} \\ S_{b+2} & S_{b+3} & S_{b+4} & \ldots & S_{b+u+1} \\ \vdots & \vdots & \vdots & & \vdots \\ S_{b+u-1} & S_{b+u} & S_{b+u+1} & \ldots & S_{b+2u-1} \end{pmatrix}$$

è non-singolare se $u = w$ numero degli errori effettivamente sommati al messaggio durante la trasmissione, mentre la matrice è singolare se $u > w$.

Dimostrazione. (del lemma 5.6.1) Dato che $S_k = e(\xi^k) = \sum_{j=1}^{w} Y_j X_j^k$ si può verificare che M è diagonalizzata dalla matrice A nel seguente modo:

$$M = ADA^t$$

dove la matrice A è data da

$$A = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ X_1 & X_2 & X_3 & \ldots & X_u \\ X_1^2 & X_2^2 & X_3^2 & \ldots & X_u^2 \\ \vdots & \vdots & \vdots & & \vdots \\ X_1^{u-1} & X_2^{u-1} & X_3^{u-1} & \ldots & X_u^{u-1} \end{pmatrix}$$

e la matrice D è data da

$$D = \begin{pmatrix} Y_1 X_1^b & 0 & 0 & \ldots & 0 \\ 0 & Y_2 X_2^b & 0 & \ldots & 0 \\ 0 & 0 & Y_3 X_3^b & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & Y_u X_u^b \end{pmatrix} = (Y_j X_j^b \delta_{ij})_{i,j}$$

[8] Da [4] teorema 7.2.2 pag. 170.

Infatti

$$\begin{aligned}(ADA^t)_{i,j} &= \sum_{k=1}^{u} X_k^{i-1}(\sum_{l=1}^{u} X_k^b Y_k \delta_{kl} X_l^{j-1}) \\ &= \sum_{l=1}^{u} X_k^{i-1} X_k^b Y_k X_k^{j-1} \\ &= \sum_{l=1}^{u} Y_k X_k^{b+i+j-1} \\ &= (M)_{ij}\end{aligned}$$

Per il teorema di Binet $det(M) = det(A)det(D)det(A)$; il determinante di D è non nullo se le coppie (Y_k, X_k) sono tutte composte da elementi non nulli, cosa vera per $u = w$.
A è una matrice di Vandermonde che ha determinante nullo se e solo se le sue colonne sono differenti fra loro e non nulle, cosa che accade se $u = w$, mentre non accade per $u > w$. □

Dimostrazione. (del teorema 5.6.5) Proponiamo una strada per ricavare i coefficienti del polinomio locatore; dimostriamo cioè che $\lambda_1, \ldots, \lambda_w$ soddisfano il sistema

$$S_{j+w} + S_{j+w-1}\lambda_1 + \cdots + S_j\lambda_w = 0 \qquad b \leq j \leq b + w - 1 \tag{5.2}$$

In conseguenza di ciò dato che 5.2 è un sistema con w equazioni e w incognite, il cui determinante associato è non nullo dal lemma precedente, allora risolvendolo con qualche metodo possiamo ricavare i coefficienti del polinomio locatore e dimostrare il teorema.
Per ogni $k = 1, \ldots, w$ poniamo $x = 1/X_k$ in $\lambda(x) = 0$ e moltiplichiamo ambo i membri per $Y_k X_k^{j+w}$ per ogni j $b \leq j \leq b + w - 1$:

$$Y_k X_k^{j+w} + \lambda_1 Y_k X_k^{j+w-1} + \lambda_2 Y_k X_k^{j+w-2} + \cdots + \lambda_w Y_k X_k^j = 0$$

Esplicitando per $k = 1, \ldots, w$ si ottiene

$$\begin{cases} Y_1 X_1^{j+w} + \lambda_1 Y_1 X_1^{j+w-1} + \lambda_2 Y_1 X_1^{j+w-2} + \cdots + \lambda_w Y_1 X_1^j = 0 \\ Y_2 X_2^{j+w} + \lambda_1 Y_2 X_2^{j+w-1} + \lambda_2 Y_2 X_2^{j+w-2} + \cdots + \lambda_w Y_2 X_2^j = 0 \\ Y_3 X_3^{j+w} + \lambda_1 Y_3 X_3^{j+w-1} + \lambda_2 Y_3 X_3^{j+w-2} + \cdots + \lambda_w Y_3 X_3^j = 0 \\ \quad \vdots \\ Y_w X_w^{j+w} + \lambda_1 Y_w X_w^{j+w-1} + \lambda_2 Y_w X_w^{j+w-2} + \cdots + \lambda_w Y_w X_w^j = 0 \end{cases}$$

dal quale otteniamo per somma

$$\begin{aligned} &Y_1 X_1^{j+w} + Y_2 X_2^{j+w} + \cdots + Y_w X_w^{j+w} + \\ &+ \lambda_1 (Y_1 X_1^{j+w-1} + Y_2 X_2^{j+w-1} + \cdots + Y_w X_w^{j+w-1}) + \\ &+ \lambda_2 (Y_1 X_1^{j+w-2} + Y_2 X_2^{j+w-2} + \cdots + Y_w X_w^{j+w-2}) + \\ &+ \cdots + \\ &+ \lambda_w (Y_1 X_1^j + Y_2 X_2^j + \cdots + Y_w X_w^j) = 0 \end{aligned}$$

Cioè

$$\sum_{K=1}^{w} Y_k X_k^{j+w} + \lambda_1 \sum_{K=1}^{w} Y_k X_k^{j+w-1} + \lambda_2 \sum_{K=1}^{w} Y_k X_k^{j+w-2} + \cdots + \lambda_1 \sum_{K=1}^{w} Y_k X_k^{j} = 0$$

Da cui ricordando che

$$S_h = \sum_{k=1}^{w} Y_k X_k^h$$

segue la tesi. □

Sostituendo nel sistema lineare 5.1 le X_k trovate risolvendo l'equazione polinomiale $\lambda(x) = 0$ i cui coefficienti sono noti grazie al teorema precedente, si possono ricavare le Y_k risolvendolo e quindi ottenere il vettore $e(x)$ e quindi la parola del codice originariamente trasmessa mediante sottrazione:

$$c(x) = v(x) - e(x)$$

Rivediamo la successione dei vari passaggi da effettuare dopo aver ricevuto $v(x)$:

- Per prima cosa è necessario trovare il valore w, che corrisponde alla quantità di errori entrati durante la trasmissione del messaggio, cioè al peso di $e(x)$. Usiamo il lemma 5.6.1: poniamo come valore iniziale $w = \lfloor \delta/2 \rfloor$ e calcoliamo $det(M)$. Se $\det(M) = 0$ allora si diminuisce w di 1 e si ricalcola il determinante, finché non si trova $det(M) \neq 0$.
- Calcoliamo la matrice M ed i coefficienti del polinomio locatore risolvendo:

$$M(1, \lambda_1, \lambda_2, \ldots, \lambda_w)^t = (0, 0, 0, \ldots, 0)^t$$

- Risolviamo l'equazione polinomiale $\lambda(x) = 0$ alle cui soluzioni corrispondono $X_1, \ldots, X_w$.
- Risolviamo il sistema di equazioni non lineari 5.1 ottenendo così le coppie $\{(Y_k, X_k)\}_{k=1}^{w}$ che forniscono magnitudine e locazione di ciascun errore.

Il procedimento appena descritto non è computazionalmente efficiente, dato che bisogna risolvere un sistema di equazioni lineari per trovare i coefficienti del polinomio locatore e di equazioni non lineari per trovare $Y_1, \ldots, Y_w$. Esiste una alternativa che implica l'uso dell'algoritmo di Berklmap-Massey e dell'algoritmo di Forney[9].

[9] Un approfondimento di questo tema si trova ad esempio in [4] pag.183.

Capitolo 6

La trasformata di Winograd nella teoria dei codici correttori

In questo capitolo presentiamo due applicazioni della trasformata di Winograd alla teoria dei codici correttori proposte nell'articolo [8] da pagina 43 a pagina 46.

6.1 Matrice Γ come generatrice dei codici ciclici

La prima applicazione ha origine con la seguente domanda:
quali informazioni possiamo ricavare dalla matrice Γ sui codici ciclici di $\mathcal{R}_{r,q}$?
Vedremo che la trasformata di Winograd è matrice generatrice e di controllo di determinati codici ciclici.

6.1.1 Matrice $\Gamma^{(v)}$ come matrice di controllo

Sia $a(x) = \mathbf{a}$ divisore di $x^r - 1$ in $\mathcal{R}_{r,q}$, $\mathfrak{a}$ ideale generato da $a(x)$, quindi codice ciclico e sia $c(x) = \mathbf{c}$ una sua parola. Consideriamo in parallelo la tesi del teorema 4.2.4 (indicata con (1)) e la tesi del lemma 3.3.2 (indicata con (2)) adattate a questo contesto:

$$\mathbf{c} \in \mathfrak{a} \iff H\mathbf{c}^t = \mathbf{0}^t \qquad (1)$$

$$c(x) \in (M^{(v)}(x)) \iff \Gamma^{(v)}\mathbf{c}^t = \mathbf{0}^t \qquad (2)$$

Se il divisore $a(x)$ è irriducibile e quindi coincidente con $M^{(v)}(x)$, allora il v-esimo blocco della trasformata di Winograd soddisfa le caratteristiche di una matrice di controllo del codice:

Teorema 6.1.1. *Sia $v \in \mathscr{L}$, allora il codice ciclico massimale $\mathfrak{a} = (M^{(v)}(x))$ in $\mathcal{R}_{r,q}$ ha come matrice di controllo $\Gamma^{(v)}$, v-esimo blocco della trasformata di Winograd.*

Dimostrazione. Sia $c(x)$ parola del codice, allora dal lemma 3.3.2 segue che $\Gamma^{(v)}\mathbf{c}^t = \mathbf{0}^t$. Viceversa se vale la precedente equazione allora $c(x) \in \mathfrak{a}$. Dal teorema 4.2.4 allora $\Gamma^{(v)}$ è matrice di controllo per il codice $(M^{(v)}(x))$ □

6.1.2 $\Gamma^{(v)}$ come matrice di generatrice

Invece di considerare il teorema 4.2.4 sulla matrice di controllo consideriamo il teorema sulla matrice generatrice 4.2.3 (indicato con $(1')$):

$$\mathbf{c} \in \mathfrak{a}^{\perp} \iff G\mathbf{c}^t = \mathbf{0}^t \qquad (1')$$

$$c(x) \in (M^{(v)}(x)) \iff \Gamma^{(v)}\mathbf{c}^t = \mathbf{0}^t \qquad (2)$$

Segue allora [1] che $\Gamma^{(v)}$ è matrice generatrice di un codice generato da un divisore particolare.

Teorema 6.1.2. *Sia $v \in \mathscr{L}$, allora il codice ciclico minimale $(\hat{M}^{(-v)}(x))$ in $\mathcal{R}_{r,q}$ ha come matrice generatrice $\Gamma^{(v)}$ v-esimo blocco della trasformata di Winograd.*

Dimostrazione. Sia $\mathfrak{a} = (\hat{M}^{(v)}(x))$ (differente dal codice ciclico dell'ipotesi per il segno). Conseguenza del teorema 5.4.2 il polinomio generatore del codice $\mathfrak{a}^{\perp}$ ortogonale ad $\mathfrak{a}$ è dato da $x^d M^{(-v)}(x)$, quindi dalla (2) vale la biimplicazione

$$c(x) \in (\hat{M}^{(v)}(x))^{\perp} = (M^{(-v)}(x)) \iff \Gamma^{(-v)}\mathbf{c}^t = \mathbf{0}^t$$

Cambiando i segni abbiamo che

$$c(x) \in (\hat{M}^{(-v)}(x))^{\perp} \iff \Gamma^{(v)}\mathbf{c}^t = \mathbf{0}^t$$

Dalla $(1')$ quindi $\Gamma^{(v)}$ è matrice generatrice di $(\hat{M}^{(-v)}(x))$. □

Quindi il v-esimo blocco di Γ non è solo matrice di controllo del codice ciclico massimale $(M^{(v)}(x))$, ma è anche matrice generatrice del codice ciclico minimale $(\hat{M}^{(-v)}(x))$, cioè del polinomio di controllo con i segni opposti. Possiamo evitare di preoccuparci della variazione dei segni se le orbite sono autoconiugate, cioè se $-1 \in O(1)$.

Corollario 6.1.1. *Se le orbite dell'azione di $G \cong Gal(\mathbb{F}_q(\xi), \mathbb{F}_q)$, $G \trianglelefteq \mathbb{Z}_r^{\star}$ su $\mathbb{Z}_r$ sono autoconiugate allora il codice massimale $(M^{(v)}(x))$ ed il codice minimale $(\hat{M}^{(v)}(x))$ sono legate dalla matrice $\Gamma^{(v)}$ che genera il primo ed è matrice di controllo per il secondo.*

Dimostrazione. È conseguenza del teorema precedente e del fatto che se le orbite sono autoconiugate allora $M^{(v)}(x) = M^{(-v)}(x)$. □

Esempio 6.1.1. *In $\mathcal{R}_{9,2}$, ricordando che $G = \mathbb{Z}_9^{\star}$ e che $\mathscr{L} = \{0, 1, 3\}$ abbiamo la fattorizzazione*

$$\begin{aligned} x^9 - 1 &= M^{(0)}(x)M^{(1)}(x)M^{(3)}(x) \\ &= (x-1)(x^6 + x^3 + 1)(x^2 + x + 1) \end{aligned}$$

[1] Ricrodando che $(a(x))^{\perp} = (\hat{a}(x)^{\perp})$

Ci sono 6 *possibili codici non banali,* 3 *minimali e* 3 *massimali. La trasformata di Winograd è data da:*

$$\Gamma = \begin{pmatrix} \Gamma^{(0)} \\ \\ \Gamma^{(1)} \\ \\ \Gamma^{(3)} \end{pmatrix} = \left(\begin{array}{ccccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{array}\right)$$

Consideriamo la parola $c(x)$ *del* $(9,3)$*-codice* $(M^{(1)}(x))$ *definita da*

$$\begin{aligned} c(x) &= (x^2+1)M^{(1)}(x) \\ &= 1 + x^2 + x^3 + x^5 + x^6 + x^8 \\ &= (1,0,1,1,0,1,1,0,1) \end{aligned}$$

Possiamo verificare che

$$\Gamma^{(1)}\mathbf{c}^t = \mathbf{0}^t$$

dal fatto che $\Gamma^{(1)}$ *è matrice di controllo di* $(M^{(1)}(x))$*. Inoltre dato che le orbite sono autoconiugate* $\Gamma^{(1)}$ *è matrice generatrice di* $(\hat{M}^{(1)}(x)) = (M^{(0)}(x)M^{(3)}(x))$*.*

Fino a qui abbiamo esaminato i codici massimali e minimali. Cosa dire dei codici generati da un divisore $a(x)$ di $x^r - 1$ generico?

6.1.3 Composizioni dei blocchi $\Gamma^{(v)}$

Sia $a(x) = \mathbf{a}$ divisore di $x^r - 1$ in $\mathcal{R}_{r,q}$ composto dal prodotto di divisori irriducibili definiti da un sottoinsieme dell'insieme delle etichette $A = \{v_1, \ldots, v_k\} \subseteq \mathscr{L}$:

$$a(x) = M^{(v_1)}(x) \cdot \cdots \cdot M^{(v_k)}(x)$$

In questo caso per costruire la matrice di controllo si dovrà prendere la matrice determinata dai blocchi ordinati[2] $\Gamma^{(v_j)}$. Inoltre, sempre generalizzando i risultati dei paragrafi precedenti la matrice costituita dai k blocchi $\Gamma^{(v_j)}$ è matrice generatrice del codice generato dal polinomio $\hat{b}(x)$ per

$$b(x) = \prod_{v \in \mathscr{L} \setminus A} M^{(-v)}(x)$$

Abbimo quindi

[2]La questione dell'ordine sembra superflua di fronte alla commutatività del prodotto dei generatori. Però nel passare dal polinomio in $\mathcal{R}_{r,q}$ al vettore circolante concatenato corrispondente è necessario mantenere un ordine nei generatori che definiscono ogni sottovettore circolate. Questo ordine si dovrà rispecchiare nel comporre la matrice di controllo dai blocchi $\Gamma^{(v_j)}$.

Corollario 6.1.2. *Sia $k \leq l$, allora per ogni scelta di k etichette ordinate $v_1, \ldots, v_k$ che costituiscono il sottoinsieme A di $\mathscr{L}$, la matrice costruita dai blocchi*

$$\Gamma^{(A)} = \begin{pmatrix} \Gamma^{(v_1)} \\ \Gamma^{(v_2)} \\ \vdots \\ \Gamma^{(v_k)} \end{pmatrix}$$

è matrice di controllo del codice

$$\mathfrak{a} = (M^{(v_1)}(x) \cdot \cdots \cdot M^{(v_k)}(x))$$

ed è matrice generatrice del codice generato dal polinomio $\hat{b}(x)$ per

$$b(x) = (M^{(-v_1)}(x) \cdot \cdots \cdot M^{(-v_k)}(x))$$

Dimostrazione. Il polinomio $m(x)$ è isomorfo al vettore $(m) \in \mathcal{V}^{\mathscr{L}}_{r,q}$ composto dai sottovettori circolanti corrispondenti all'immagine di $m(x) \mod M^{(v)}(x)$ tramite ψ_2 nella posizione v-esima.

$$\mathbf{m}_v = \psi_2(m(x) \mod M^{(v)}(x)) \in \mathcal{V}^c_{m(v),q}$$
$$\mathbf{m} = concat(\mathbf{m}_0, \mathbf{m}_1, \ldots, \mathbf{m}_{max(\mathscr{L})})$$

Ora, $\Gamma^{(A)}\mathbf{m}^t$ è il vettore nullo se e solo se per ogni j il blocco $\Gamma^{(v_j)}$ annulla il vettore

$$concat(\mathbf{0}, \ldots, \mathbf{0}, \mathbf{m}_{v_j}, \mathbf{0}, \ldots, \mathbf{0})$$

Ma questa è una conseguenza del teorema precedente.
In modo analogo si dimostra che $\Gamma^{(A)}$ è una matrice generatrice. □

Prima di passare alla seconda applicazione esaminiamo un esempio:

Esempio 6.1.2. *In $\mathcal{R}_{7,2}$ abbiamo $G \trianglelefteq \mathbb{Z}_7^\star$, $G = \{1, 2, 4\}$, $\mathscr{L} = \{0, 1, 3\}$ e la seguente fattorizzazione:*

$$\begin{aligned} x^7 - 1 &= M^{(0)}(x)M^{(1)}(x)M^{(3)}(x) \\ &= (x-1)(x^3 + x + 1)(x^3 + x^2 + 1) \end{aligned}$$

Consideriamo il $(7, 3)$-codice $\mathfrak{a}$ generato da $M^{(0)}(x)M^{(1)}(x)$: la sua matrice di controllo è la composizione delle matrici $\Gamma^{(0)}$ e $\Gamma^{(1)}$:

$$\begin{pmatrix} \Gamma^{(0)} \\ \Gamma^{(1)} \end{pmatrix} = \left(\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right)$$

La stessa composizione è matrice generatrice del $(7,4)$*-codice* $(M^{(-3)}(x)) = (M^{(1)}(x))$ *(in questo caso le orbite non sono autoconiugate).*
Sia $m(x) = 1 + x + x^3 = (1,1,0,1)$ *messaggio che vogliamo inviare usando il codice* $(M^{(1)}(x))$*, allora la sua codifica è:*

$$\mathbf{c} = (1,1,0,1) \begin{pmatrix} \Gamma^{(0)} \\ \\ \Gamma^{(1)} \end{pmatrix}$$

e quindi $c(x) = x + x^5 + x^6$.
Se invece riceviamo la parola $c(x) = 1 + x + x^2 + x^5 = (1,1,1,0,0,1,0)$ *e vogliamo verificare la sua appartenenza al codice* $(M^{(0)}(x)M^{(1)}(x))$ *ne calcoliamo il prodotto per la sua matrice di controllo:*

$$\left(\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Dato che il risultato è nullo segue che $c(x) \in (M^{(0)}(x)M^{(1)}(x))$.

6.2 Matrice Δ nella decodifica

In che modo possiamo usare Δ per decodificare un messaggio?
Sia $\mathfrak{a}$ codice ciclico in $\mathcal{R}_{r,q}$. Il suo polinomio generatore $a(x)$ può essere rappresentato su algebre isomorfe in diversi modi equivalenti:

1. $\gamma(a(x)) = (a(x) \mod M^{(v)}(x))_{v \in \mathscr{L}} \in \mathcal{Q}_{r,q}$
2. $\eta(a(x)) = (a(\xi^v))_{v \in \mathscr{L}} \in \mathcal{P}_{r,q}$
3. $\psi_2(a(x)) = \mathbf{b} \in \mathcal{V}^c_{r,q}$
4. $\psi_6(\gamma(a(x))) = \mathbf{a} \in \mathcal{V}^{\mathscr{L}}_{r,q}$
5. $\psi_5(\eta(a(x))) = \mathbf{a} \in \mathcal{V}^{\mathscr{L}}_{r,q}$

dove osserviamo che $\psi_5(\eta(a(x))) = \psi_6(\gamma(a(x)))$ è il vettore circolante concatenato diverso dal vettore circolante $\mathbf{b}$. Dato che $a(x)$ è un divisore, allora nella rappresentazione in $\mathcal{V}^{\mathscr{L}}_{r,q}$ è il vettore circolante concatenato nel quale i blocchi corrispondenti alle posizioni dei generatori di $\mathfrak{a}$ sono sottovettori nulli. Possiamo quindi ridurre il numero di spazio occupato da un messaggio inviato.

6.2.1 Sottovettori privi di informazione

Sia $a(x)$ generatore di un codice ciclico definito dal prodotto di k divisori di $\mathcal{R}_{r,q}$:

$$a(x) = M^{(v_1)}(x) \cdot ... \cdot M^{(v_k)}(x)$$

allora per $c(x)$ parola del codice, il vettore $\psi_6(\gamma(c(x))) = \mathbf{c}$ ha nel v_j-esimo posto un sottovettore nullo di dimensione $m(v)$:

$$\begin{aligned}\psi_2(c(x)) &= (c_0, c_1, \ldots, c_{r-1})\\ \gamma(c(x)) &= (c(x) \mod M^{(v)}(x))_{v\in\mathscr{L}}\\ \psi_6(\gamma(c(x))) &= (\psi_2(c(x) \mod M^{(v)}(x)))_{v\in\mathscr{L}}\end{aligned}$$

Quindi $\psi_6(\gamma(c(x)))_{v_j} = \mathbf{0} \in \mathcal{V}^c_{m(v),q}$ per ogni j compreso fra 1 e k.
Il messaggio $c(x)$ può essere quindi scritto ed inviato omettendo i sottovettori nulli e scrivendo solo gli $l-k$ vettori circolanti non nulli; restringendo cioè il suo dominio di appartenenza ai soli campi determinati da polinomi che non compaiono fra i fattori di $a(x)$.

$$\mathbf{c} \in \mathcal{V}^{\mathscr{L}}_{r,q} \mapsto \mathbf{c} \;\Big|\; \coprod_{v\in\mathscr{L}\setminus\{v_j\}} \mathcal{V}^c_{m(v),q}$$

Definiamo **sottovettori privi di informazione** i sottovettori nulli omessi in questo procedimento. Diamo all'algebra dei vettori concatenati di dimensione $\sum_{v\in\mathscr{L}\setminus\{v_j\}} m(v)$ senza blocchi privi di informazione

$$\coprod_{v\in\mathscr{L}\setminus\{v_j\}} \mathcal{V}^c_{m(v),q}$$

il nome di **spazio dei sottovettori di informazione**.

Esempio 6.2.1. *Proseguiamo l'esempio 6.1.2. Vogliamo inviare la parola già codificata* $c(x) = 1 + x + x^2 + x^5 = (1,1,1,0,0,1,0)$ *appartenente al codice* $(M^{(0)}(x)M^{(1)}(x))$*, allora la scomponiamo tramite* γ *nel vettore di polinomi:*

$$\gamma(c(x)) = (0, 0, x^2)$$

che tramite ψ_5 *diventa il vettore circolante a blocchi*

$$\psi_5(\gamma(c(x))) = (0|0,0,0|0,0,1)$$

quindi possiamo inviare solo $(0,0,1)$ *per comunicare la parola. Il rcevente aggiungerà i sottovettori privi di informazione, applicherà la trasformata di Winograd inversa* Δ *e procederà alla decodifica.*

Nel precedente esempio la matrice Δ ha avuto un ruolo essenziale nella decodifica; vediamo alcuni dettagli nel prossimo paragrafo.

6.2.2 Codifica con Δ, decodifica con Γ

Vogliamo utilizzare l'idea dei sottovettori privi di informazione come punto di partenza per codificare e decodificare una parola di un codice in $\mathcal{R}_{r,q}$. Per fare ciò partiamo dalla struttura $\mathcal{V}^{\mathscr{L}}_{r,q}$. Per quanto detto, una volta scelto il polinomio generatore, che determina un sottoinsieme dell'insieme delle etichette $A = \{v_1, \ldots, v_k\}$ possiamo scrivere il messaggio da inviare nei $l-k$ blocchi di $\mathcal{V}^{\mathscr{L}}_{r,q}$ che non sono sottovettori privi di informazione. Applicando la matrice Δ trasportiamo questo vettore nello spazio $\mathcal{V}^c_{r,q}$ al quale corrisponde tramite ψ_2 il polinomio $c(x)$ che è la parola di $\mathcal{R}_{r,q}$ che vogliamo inviare.
Riassumendo

1. Scegliamo $a(x)$ divisore di $x^r - 1$ che determina univocamente l'insieme $A = \{v_1, \dots v_k\}$ sottoinsieme dell'insieme delle etichette corrispondenti ai polinomi $M^{(v)}(x)$ che compaiono nella fattorizzazione di $a(x)$.

2. Scriviamo negli $l - k$ blocchi di dimensione complessiva $\sum_{v \in \mathscr{L} \setminus \{v_j\}} m(v)$ il messaggio da inviare. Questo è un elemento di $\coprod_{v \in \mathscr{L} \setminus \{v_j\}} \mathcal{V}^c_{m(v),q}$ che indichiamo con $\tilde{\mathbf{c}}$.

3. Aggiungiamo ad $\tilde{\mathbf{c}}$ i k blocchi privi di informazione, ottenendo $\mathbf{m} \in \mathcal{V}^{\mathscr{L}}_{r,q}$.

4. Calcoliamo $\Delta \mathbf{c}^t \in \mathcal{V}^c_{r,q}$ al qual corrisponde il polinomio $c(x) \in \mathcal{R}_{r,q}$ tramite ψ_2.

A questo punto la matrice Γ può essere usata per decodificare il messaggio $c(x)$, infatti $\Gamma \mathbf{c}^t$ fornisce le sindromi dei codici massimali che contengono il codice $(a(x))$. Queste sindromi devono essere blocchi nulli nelle posizioni v_j del vettore decodificato.

Capitolo 7

Appendice

7.1 Ricerca dei polinomi minimi sui campi finiti

Nella pratica, per trovare i polinomi minimi sui campi finiti, esiste un procedimento che non implica la ricerca del gruppo H isomorfo al gruppo di Galois $Gal(\mathbb{F}(\xi), \mathbb{F})$ presentato nel capitolo 1, e che risulta essere quindi più maneggevole per le implementazioni[1]. Si riconduce alla seguente definizione che riformula la definizione di elementi coniugati, classi ciclotomiche e polinomio minimo.

Definizione 7.1.1. *Sia $\xi \in \mathbb{F}_{q^m}$ non nullo e sia t il più piccolo intero positivo tale che $\xi^{p^t} = \xi$. Allora l'insieme*

$$C(\xi) = \{\xi, \xi^p, \xi^{p^2}, \ldots, \xi^{p^{t-1}}\}$$

è detto **classe ciclotomica** *e due elementi di tale insieme sono detti* **coniugati**. *Si definisce* **polinomio minimo** *di ξ il più piccolo polinomio in $\mathbb{F}_q$ che ammette ξ come radice.*

Con questa definizione si dimostra[2] che il polinomio minimo di un elemento di $\mathbb{F}_q$ non nullo è il più piccolo polinomio che contiene tutti gli elementi di una stessa classe ciclotomica:

Teorema 7.1.1. *Se ξ è un elemento non nullo di $\mathbb{F}_q$ allora il suo polinomio minimo $M_\xi(x)$ è un polinomio irriducibile in $\mathbb{F}_q[x]$ ed è definito come*

$$M_\xi(x) = \prod_{\beta \in C(\xi)} (x - \beta)$$

Dimostrazione. Se per assurdo $M_\xi(x)$ polinomio minimo di ξ non è un polinomio irriducibile allora può essere scritto come prodotto di due polinomi in $\mathbb{F}_q[x]$ $M_\xi(x) = f_1(x)f_2(x)$ con $0 \leq deg(f_j(x)) < deg(M_\xi(x))$. Dato che $M_\xi(\xi) = f_1(\xi)f_2(\xi) = 0$ e dato che $\mathbb{F}_q$ è un campo, allora uno dei due polinomi della fattorizzazione di $M_\xi(x)$ ammette ξ come radice in contraddizione con la minimalità di $M_\xi(x)$.

[1] Come proposto ad esempio in [6] pag.83.
[2] Da [6], teoremi 4.36 e 4.38.

Rimane da dimostrare che il polinomio minimo è proprio determinato dal prodotto $\prod_{\beta \in C(\xi)}(x-\beta)$.
Sia $M_\xi(x) = \sum_{j=0}^{t} m_j x^j$ definito come polinomio minimo di ξ. Dato che se ξ è una sua radice, lo deve necessariamente essere anche ξ^p:

$$M_\xi(\xi^p) = \sum_{j=0}^{t} m_j (\xi^p)^j = \sum_{j=0}^{t} m_j^p (\xi^p)^j \tag{7.1}$$

$$= \sum_{j=0}^{t} (m_j \xi^j)^p = (\sum_{j=0}^{t} m_j \xi^j)^p \tag{7.2}$$

$$= (M_\xi(\xi))^p = 0 \tag{7.3}$$

inoltre per la sua minimalità $M_\xi(x)$ non ammette altre radici. Quindi

$$M_\xi(x) = \prod_{\beta \in C(\xi)} (x-\beta)$$

La 7.1 è valida solo se si può affermare che $m_j^p = m_j$ cosa che accade solo dimostrando che $m_j \in \mathbb{F}_q$. Questo completa la dimostrazione:
da un lato segue che

$$(M_\xi(x))^p = \prod_{\beta \in C(\xi)} (x-\beta)^p = \prod_{\beta \in C(\xi)} (x^p - \beta^p)$$

$$= \prod_{\beta \in C(\xi)} (x^p - \beta) = (M_\xi(x^p)) = \sum_{j=0}^{t} m_j x^{jp}$$

e d'altra parte

$$(M_\xi(x))^p = \sum_{j=0}^{t} (m_j x^j)^p = \sum_{j=0}^{t} m_j^p x^{jp}$$

Pertanto $m_j^p = m_j$ ed $m_j \in \mathbb{F}_q$. □

7.2 Ricorrenze lineari

Accenniamo al rapporto fra l'algebra $\mathcal{R}_{r,q}$ e le ricorrenze lineari cominciando con un esempio:

Esempio 7.2.1. *Dato il polinomio $s(x) = x^4+x+1$ nell'algebra $\mathcal{R}_{7,2}$, indichiamo con $\mathscr{F}_{s(x)}$ un vettore di lunghezza infinita verso destra costruito ripetendo $\psi_2(s(x))$ senza alterazioni:*

$$\mathscr{F}_{s(x)} = (1,1,0,0,1,0,0|1,1,0,0,1,0,0|1,1,0,0,1,0,0|1,1,0,\dots)$$

Indicando con $\mathscr{F}_{s(x)}(j)$, per j intero non negativo, il suo j-esimo elemento, allora $\mathscr{F}_{s(x)}$ è caratterizzato dall'equazione

$$\mathscr{F}_{s(x)}(j) = \mathscr{F}_{s(x)}(j-7) \qquad \forall j \geq 7$$

Variando $s(x)$ in $\mathcal{R}_{7,2}$, l'insieme costituito da tutti i vettori di lunghezza infinita $\mathscr{F}_{s(x)}$ è uno spazio vettoriale isomorfo a $\mathbb{F}_2^7$ ed è un'algebra isomorfa a $\mathcal{R}_{7,2}$ se dotata del prodotto di convoluzione in modo naturale.
Da un altro punto di vista, il vettore $\psi_2(s(x))$ definisce una funzione da $\mathbb{Z}_7$ in $\mathbb{F}_2$ che associa ad ogni j in $\mathbb{Z}_7$ il j-esimo elemento del vettore $\psi_2(s(x))$, indicato con s_j. Tale funzione può essere considerata anche da $\mathbb{Z}$ in $\mathbb{F}_2$; in questo caso a $j \in \mathbb{Z}$ è associato il j-esimo elemento di una stringa infinita anche verso sinistra costruita ripetendo il vettore $\psi_2(s(x))$.
Se indichiamo tale stringa con $\mathscr{I}_{s(x)}$ ed il suo j-esimo elemento con $\mathscr{I}_{s(x)}(j)$ allora abbiamo

$$\begin{aligned}&\mathscr{I}_{s(x)} = (\dots 0,0|1,1,0,0,1,0,0|1,1,0,0,1,0,0|1,1,0,\dots)\\&j \longmapsto \mathscr{I}_{s(x)}(j) = \mathbf{s}_j \mod 7\end{aligned}$$

Come prima, considerando l'insieme costituito da tutte le funzioni sulla stringa infinita $\mathscr{I}_{s(x)}$ al variare di $s(x)$ otteniamo uno spazio vettoriale isomorfo a $\mathbb{F}_2^7$ ed isomorfa a $\mathcal{R}_{7,2}$ come algebra se considerato con il prodotto di convoluzione sui vettori che generano le stringhe infinite. Infatti ogni funzione è univocamente determinata dal vettore $\psi_2(s(x))$ che ne definisce il dominio.

Giocando su vettori e stringhe, nel precedente esempio abbiamo costruito due algebre

$$\begin{aligned}\mathscr{F} &= \{\mathscr{F}_{s(x)} \mid s(x) \in \mathcal{R}_{7,2}\} = \{f : \mathbb{Z}_7 \to \mathbb{F}_2 \mid f(j) = \mathscr{F}_{s(x)}(j)\}\\ \mathscr{I} &= \{\mathscr{I}_{s(x)} \mid s(x) \in \mathcal{R}_{7,2}\} = \{f : \mathbb{Z} \to \mathbb{F}_2 \mid f(j) = \mathscr{I}_{s(x)}(j)\}\end{aligned}$$

le quali, oltre ad essere due ulteriori varianti di $\mathcal{R}_{7,2}$, sono un caso particolare di successioni lineari ricorrenti 7-periodiche e di funzioni lineari 7-periodiche, che introdurremo in questo paragrafo[3].

Definizione 7.2.1. *Siano $\mathbf{s} = (s_0, s_1, \dots, s_{r-1})$ vettore dei* **valori iniziali** *ed $\mathbf{a} = (a_0, a_1, \dots, a_{r-1})$ vettore dei* **coefficienti***, elementi di $\mathcal{V}_{r,q}^c$, allora una successione $F_j = F_n(\mathbf{a}, \mathbf{s})$ che soddisfa la relazione di ricorrenza*

$$\begin{cases} F_j(\mathbf{a},\mathbf{s}) = s_j & 0 \le j \le r-1 \\ F_j(\mathbf{a},\mathbf{s}) = \sum_{k=0}^{r-1} a_k F_{j-r+k}(\mathbf{a},\mathbf{s}) & n \ge r \end{cases}$$

è detta **sequenza lineare ricorrente** *di ordine r.*

Ad ogni sequenza lineare ricorrente si può associare un polinomio ed una matrice caratteristica che permettono di avere a disposizione degli strumenti in più per il loro studio.

Definizione 7.2.2. *Sia $F_j(\mathbf{a}, \mathbf{s})$ sequenza lineare ricorrente, allora il polinomio $c(x) \in \mathbb{F}_q[x]$ definito come*

$$c(x) = x^r - a_{r-1}x^{r-1} - a_{r-2}x^{r-2} - \dots - a_1 x - a_0$$

[3]Per una esposizione completa delle successioni lineari ricorrenti sui campi finiti: [17] pagine 190 e seguenti.

è detto **polinomio caratteristico** *di* F_j.
Mentre la matrice definita sul vettore $\mathbf{a}$ *da*

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ 0 & 0 & 1 & \cdots & 0 & a_3 \\ & \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{r-1} \end{pmatrix}$$

è detta **matrice caratteristica** *o matrice compagna.*

Osserviamo che per $a_0 \neq 0$ la matrice A è invertibile ed appartiene al gruppo lineare. La matrice caratteristica di una ricorrenza lineare genera la ricorrenza lineare, infatti definendo $v_k = v_k(\mathbf{a}, \mathbf{s}) = (F_k, F_{k+1}, \ldots, F_{k+r-1})$ il vettore costituito da una successione di r elementi della sequenza lineare ricorrente F_j a partire dall'elemento k-esimo si verifica che

$$v_k = v_0 A^k$$

dove per definizione v_0 coincide con il vettore dei valori iniziali.
Grazie alla definizione di matrice compagna si può dimostrare[4] che se la sequenza lineare ricorrente $F_j(\mathbf{a}, \mathbf{s})$ è omogenea, allora è periodica, cioè esiste un intero positivo f tale che $F_{j+f} = F_j$ per ogni j positivo.
Il più piccolo f che soddisfa l'equazione precedente è detto **periodo** di F_n, e la sequenza lineare è detta f**-periodica**.

Definizione 7.2.3. *Lo spazio vettoriale costituito dall'insieme delle ricorrenze lineari aventi* $c(x)$ *come polinomio caratteristico è indicato con* $Rec(c(x))$*:*

$$Rec(c(x)) = \{F_j(\psi_2(c(x)), \mathbf{s}) \mid \mathbf{s} \in \mathcal{V}_{r,q}^c\}$$

Come presentato nell'esempio introduttivo, ogni polinomio $a(x)$ di $\mathcal{R}_{r,q}$ il cui vettore circolante associato è dato da $\psi_2(a(x)) = (a_0, a_1, \ldots, a_{r-1})$ definisce una funzione da $\mathbb{Z}_r$ in $\mathbb{F}_q$:

$$\begin{aligned} F\big|_{\mathbb{Z}_r} : \mathbb{Z}_r &\longrightarrow \mathbb{F}_q \\ j &\longmapsto a_j \end{aligned}$$

il cui dominio può essere esteso ai numeri interi, considerando la composizione con la proiezione π da $\mathbb{Z}$ in $\mathbb{Z}_r$:

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\pi} & \mathbb{Z}_r \\ & \searrow^{F} & \downarrow F\big|_{\mathbb{Z}_r} \\ & & \mathbb{F}_q \end{array}$$

$$\begin{aligned} F = F\big|_{\mathbb{Z}_r} \circ \pi : \mathbb{Z} &\longrightarrow \mathbb{F}_q \\ j &\longmapsto a_{j \mod r} \end{aligned}$$

[4] Per brevità rimandiamo i dettagli al già citato [17].

Quindi la funzione $F(j) = F_j$ è r-periodica e può essere vista come una ricorrenza lineare con polinomio caratteristico $x^r - 1$. Le due strutture si equivalgono. Nell'esempio 7.2.1 abbiamo considerato la ricorrenza lineare avente come polinomio caratteristico $x^7 - 1$. In generale le ricorrenze lineari aventi come polinomio caratteristico $x^r - 1$ danno luogo a sequenze r-periodiche nelle quali il vettore $s(x)$ si ripete indefinitamente e senza variazioni. La struttura $Rec(x^r - 1)$ è uno spazio vettoriale r-dimensionale isomorfo a $\mathbb{F}_q^r$ e considerando il prodotto di convoluzione fra i vettori dei valori iniziali è un'algebra isomorfa a $\mathcal{R}_{r,q}$:

Teorema 7.2.1. *La funzione*

$$\begin{aligned} \psi_2 : \mathcal{R}_{r,q} &\longrightarrow Rec(x^r - 1) \\ s(x) &\longmapsto F_j(\psi_2(x^r - 1), \psi_2(s(x))) \end{aligned}$$

è un isomorfismo di algebre

Dimostrazione. I due spazi vettoriali sono entrambi isomorfi a $\mathbb{F}_2^7$, inoltre per $s(x)$ e $t(x)$ in $\mathcal{R}_{r,q}$ si ha che

$$\psi_2(s(x)t(x)) = \psi_2(s(x))\psi_2(t(x))$$

per come è stato definito il prodotto su $Rec(x^r - 1)$. □

È stato dimostrato ([8] pag. 21) che un ideale generato da $a(x)$ divisore di $x^r - 1$ coincide con lo spazio delle ricorrenze lineari r-periodicheaventi polinomio caratteristico

$$h(x) = \frac{\hat{a}(x)^{\perp}}{a_0}$$

ed inoltre l'ideale minimale $(M^{(v)}(x))$ è costituito da tutte le ricorrenze lineari r-periodiche aventi polinomio caratteristico $M^{(-v)}(x)$.
Una ulteriore indagine potrebbe essere sviluppata per rispondere a questa domanda: Cosa accade se anziché considerare $Rec(x^r - 1)$, consideriamo $Rec(a(x))$ per $a(x)$ divisore di $\mathcal{R}$?

Bibliografia

[1] Michael Artin, *Algebra*, Prentice Hall of India, New Delhi 2007

[2] Luigia Berardi, *Algebra e teoria dei codici correttori*, Franco angeli Editore 1994.

[3] E.R. Berkelamp, *Factoring Polynomials over Finite Fileds*, The Bell System Technical Journal, October 1967, pag. 1853-1859.

[4] Richard E. Blahut, *Theory and Practice of Error Control Codes*, Addison Wesley publishing Company, 1984.

[5] Ian F. Blake, Ronald C. Mullin, *The Mathematical Theory of Coding*, Academic Press 1975.

[6] Giulia Maria Piacentini Cattaneo, *Algebra, un approccio algoritmico*, Zanichelli 2007, prima ed. 1996.

[7] D.G. Cantor, H. Zasenhaus, *A new Algorithm for Factoring Polynomials over Finite Fileds*, Mathematics of Computation, volume 36, numero 154, aprile 1981, pag. 587-592.

[8] U. Cerruti, F. Vaccarino *From Cyclotomic Extensions to Generalized Ramanuyan's Sum through the Winograd Transform*, pre-print.

[9] K.M. CHerung, F. Pollara *Phobos Lander Coding System: Software and Analysis*, TDA progess report, April-June 1988, pag 274-286.

[10] Lindsday N. Childs, *A Concrete Introduction to Higher Algebra*, Springer-Verlag Gmbh, Third Edition 2009.

[11] Philip J. Davis, *Circulant Matrices*, John Wiley and Sons, 1979.

[12] Angela Di Febbraro, Alessandro Giua, *Sistemi ad eventi discreti*, McGraw Hill, ed 2011, prima ed 2002.

[13] I.N. Herstein, *Algebra*, editori riuniti, University Press 2010, prima ed. 1982.

[14] Nathan Jacobson, *Basic Algebra I*, W.H. Freeman and Company, prima ed. 1985.

[15] Thomas Koshy, *Elementary Number Theory with applications*, Accademic Press, Elzevier, 2007.

[16] A. Languasco, A. Zaccagnini, *Introduzione alla crittografia*, Ulrico Hoepli editore 2004.

[17] R. Lidl, H. Niederreiter *Introduction to Finite Fields and Applications*, Cambridge university press 1994, prima edizione 1986.

[18] J.H. van Lint, *Introduction to Coding Theory*, Springer-Verlag, GTM 86 Third Edition 1999.

[19] J.S. Milne, *Fields and Galois Theory*, electronic version, Creative Commons license, 27 May 1998.

`http://www.jmilne.org/math/`

[20] Timothy Murphy, *Course 373, Finite Fields*, electronic version, Creative Commons license.

`pet.ece.iisc.ernet.in/sathish/FiniteFields.pdf`

[21] Andrea Montabone, *Matrici Circolanti ed Applicazioni*, Tesi di Laurea Magistrale, Università degli studi di Torino, Ottobre 2011. Relatore: Prof. U. Cerruti.

[22] R. L. Miller, T. K. Truong e I. S. Reed, *Efficient Program for decoding the* $(255, 223)$ *Reed-Solomon Code over* $GF(2^8)$, IEEE num. 127 1980 pag 136-142.

[23] Daniel Perrin, *Algebraic Geometry, an introduction*, Springer 2008.

[24] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Publishing company 1977.

[25] Claude E. Shannon, *A Mathematica Theory of Comunication*, Bell System Technical Journal 27 379-423, July/October 1948.

[26] Victor Shoup, *A Computational Introduction to Number Theory and Algebra*, electronic version, Creative Commons license, 2008, version 2.

`http://shoup.net/ntb/`

[27] R. Sivaramakrishnan, *Classical Theory of Arithmetic Functions*, Taylor and Francis, 1989.

[28] M. Stoka, *Corso di Geometria*, CEDAM, 1995.

[29] Alun Wyn-jones, *Circulants*, no editor: electronic version, Creative Commons license, January 2008.

`http://www.circulants.org/circ/`

[30] S. Winograd, *Arithmetic Complexity of Computations*, SIAM 1980.

[31] S. Winograd, *On Computing the Discrete Fourier Transform*, Mathematics of Computation, Vol. 32, Num. 141, Gennaio 1978, pag. 175-199.