Codici e Geometria*

Francesco Mazzocca **

Abstract: The *theory of linear codes* is a mathematical theory that sometimes allows to automatically correct errors that may occur during the transmission of information. Actually, it is a series of natural applications of results of finite fields and combinatorial geometries; very interestingly, its methods are also the basis of deep results in design theory such as, for instance, the *non-existence of a projective plane order* 10.

Sunto: Una teoria matematica che permette, in alcuni casi, di correggere automaticamente gli errori che possono verificarsi durante la trasmissione di informazioni è la *teoria dei codici lineari,* della quale intendiamo esporre i primi elementi. Questa si presenta come una serie di naturali applicazioni delle teorie dei campi finiti e delle geometrie combinatorie; inoltre, cosa molto interessante, i suoi metodi sono alla base di risultati profondi in teoria dei disegni come, ad esempio, la *non esistenza di un piano proiettivo d'ordine* 10.

1 Introduzione

Codificare e decodificare messaggi per permettere una loro rapida trasmissione è un'antica esigenza dell'umanità. I metodi usati nel passato sono stati i più svariati; dai *tam-tam* delle popolazioni indigene africane ai *segnali di fumo* degli indiani d'America. Oggi, l'era della trasmissione dell'informazione in *tempo reale,* i mezzi che abbiamo a disposizione per inviare e ricevere messaggi hanno raggiunto un grado di sofisticazione molto elevato e il loro funzionamento si basa sull'alta tecnologia e su precise teorie matematiche, fisiche ed informatiche.

L'articolo di *Claude Elwood Shannon* [15] segna l'inizio della Teoria dell'Informazione. Qui per la prima volta si traduce in termini matematici ciò che comunemente si intende per *informazione*. È, inoltre, precisato il problema fondamentale della comunicazione: *"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point"*. Questo significa che da un *messaggio ricevuto*, anche se in errore, deve potersi risalire al *messaggio inviato.*

*Testo della relazione tenuta a Caserta il 27 ottobre 2011 in occasione del Congresso Nazionale della Mathesis.

**Dipartimento di Matematica della Seconda Università degli Studi di Napoli - francesco.mazzocca@unina2.it

Uno dei risultati teorici fondamentali di Shannon può sintetizzarsi nel modo seguente: *si può sempre comunicare in modo efficiente a patto di liberare prima i messaggi da tutta l'informazione in eccesso* (compressione dei dati) *e di aggiungere poi in modo controllato altra informazione* (ridondanza) *che permetta di scoprire o correggere eventuali errori dipendenti dalla trasmissione.*

La *compressione di dati* e l'*aggiunta di ridondanza controllata* costituiscono due problemi che possono essere studiati usando dei modelli matematici. La *teoria dei codici lineari* è particolarmente utile allo studio del secondo problema.

Una delle caratteristiche essenziali richieste ad un sistema di comunicazione è dunque l'affidabilità: *un messaggio trasmesso deve poter essere decifrato correttamente dal destinatario.* Purtroppo, non esistendo sistemi di comunicazione perfetti, la probabilità che nel corso di una trasmissione si verifichino degli errori non può mai ridursi a zero. Gli effetti negativi dovuti a questa situazione possono essere ridotti sostanzialmente in due modi:

1. Intervenire direttamente sui canali di trasmissione mediante l'utilizzo di nuove tecnologie, modificando in parte quelli in uso o costruendone dei nuovi, allo scopo di ridurre la possibilità che si verifichino errori nei canali stessi (si pensi per esempio all'alta affidabilità dei sistemi che fanno uso di fibre ottiche o di raggi laser).

2. Adottare codici che, tenendo conto del grado di affidabilità del sistema di comunicazione in uso, permettano di scoprire e correggere automaticamente eventuali errori.

Il primo approccio, vicino all'ingegneria e alla fisica, necessita ovviamente di disponibilità finanziarie non indifferenti. Già a livello di ricerca i progetti hanno costi molto alti e non è detto che eventuali risultati positivi siano convenienti per immediate applicazioni. Basti pensare che le scelte legate a modifiche, anche parziali, delle architetture dei sistemi di comunicazione a larga diffusione, al di là dei problemi economici e tecnologici, possono avere notevoli implicazioni anche di carattere politico. Chi non è giovanissimo certamente ricorda i problemi che ebbe il Governo italiano nel 1975 quando, per l'introduzione della TV a colori, si trovò a dover scegliere tra il sistema francese SECAM (Séquential Coleur à Mémoire) e quello tedesco PAL (Phase Alternation Line).

Il secondo approccio, di tipo matematico-informatico, pur avendo bisogno nella pratica di supporti tecnologici, non presenta gli inconvenienti del primo ed è molto meno costoso. In altre parole, *correggere gli errori è più conveniente che prevenirli* intervenendo direttamente sull'hardware dei sistemi di comunicazione. L'approccio in questione ha le sue basi teoriche nei già citati risultati di

Shannon di [15]; qui, tra l'altro, è provata l'esistenza di codici che rendono possibile la massima efficienza nella correzione degli errori, tenendo conto del canale utilizzato.

Chiariamo subito il discorso con un esempio.

Esempio 1.1. Supponiamo di disporre di un ipotetico canale di trasmissione con la proprietà di modificare al più una lettera su ogni parola binaria di lunghezza non superiore a cinque. Supponiamo, inoltre, di dover trasmettere dei messaggi scelti fra i seguenti quattro: NORD, SUD, EST, OVEST. In questo caso il modo più naturale e veloce per trasmettere, ma anche il più ingenuo, è quello di ridurre al minimo possibile il numero di lettere nelle parole (compressione dei dati) che servono per trasmettere i nostri messaggi, codificandoli con il codice binario

$$C_1 = \{00, 01, 10, 11\};$$

per esempio ponendo

$$\text{NORD}\equiv 00,\ \text{SUD}\equiv 01,\ \text{EST}\equiv 10,\ \text{OVEST}\equiv 11.$$

In questo modo, se una parola viene modificata, per esempio 10 in 11, il destinatario riceve il messaggio OVEST invece di EST, non avendo il decodificatore alcun elemento per scoprire che la parola ricevuta è diversa da quella trasmessa.

Proviamo ora ad aggiungere della *ridondanza controllata* all'informazione compressa del codice C_1. Per esempio usiamo il codice C_2 con parole di lunghezza tre

$$C_2 = \{000, 011, 101, 110\}$$

e poniamo

$$\text{NORD}\equiv 000,\ \text{SUD}\equiv 011,\ \text{EST}\equiv 101,\ \text{OVEST}\equiv 110.$$

In questo caso, se su una parola si commette un solo errore, questo può essere scoperto ma non corretto. Per esempio, se 000 si trasforma in 100, il decodificatore si trova di fronte ad una parola non appartenente al codice C_2 e riconosce l'errore ma, contenendo C_2 più di una parola che può trasformarsi in 100 col cambio di una sola lettera, non è in grado di risalire alla parola effettivamente trasmessa.

A questo punto è facile rendersi conto che, se vogliamo avere la possibilità di scoprire e correggere almeno un errore, abbiamo bisogno di un codice con parole di lunghezza almeno cinque, per esempio

$$C_3 = \{00000 \equiv \text{NORD},\ 01101 \equiv \text{SUD},\ 10110 \equiv \text{EST},\ 11011 \equiv \text{OVEST}\}.$$

È infatti immediato provare che, se una parola $\boldsymbol{x} \in C_3$ si trasforma in una $\boldsymbol{y}$ mediante lo scambio di una sola lettera, non esistono parole di C_3 diverse da $\boldsymbol{x}$ con questa stessa proprietà.

Osserviamo che nei codici C_2 e C_3 le prime due lettere delle parole individuano completamente il *messaggio* (informazione compressa) mentre le rimanenti sono le cosiddette *lettere di controllo* (ridondanza controllata). Inoltre, nel passaggio dal codice C_1 al codice C_3 si arriva ad un sistema di comunicazione affidabile. Il prezzo pagato per l'affidabilità è chiaramente la minore velocità della trasmissione; infatti, essendo le parole di C_3 più lunghe di quelle di C_1, la trasmissione stessa sarà necessariamente più lenta. □

I risultati di Shannon, purtroppo, sono puramente teorici. Essi riguardano solo l'*esistenza* e non la *effettiva costruzione* di codici correttori ottimali; nella pratica la ricerca di questi codici è spesso un problema molto arduo. Il primo a occuparsi della costruzione di codici ottimali, almeno due anni prima della pubblicazione del citato articolo di Shannon, è stato il matematico americano *Richard Hamming,* che nel 1948 fondò la *teoria della correzione degli errori* con la scoperta di una classe di codici correttori binari ([7],[8]) che ora portano il suo nome. Nell'esempio che segue riportiamo uno dei primi codici correttori scoperti da Hamming.

ESEMPIO 1.2. I dati (cioè le informazioni) sono rappresentati dalle successioni binarie di lunghezza t^2 e ogni dato è codificato da una successione binaria di lunghezza $(t+1)^2$; si aggiungono cioè $2t+1$ simboli di controllo. Per codificare un dato si opera nel seguente modo:
• si dispongono gli elementi di una successione binaria di lunghezza t^2 in una tabella quadrata d'ordine t,
• si orla tale tabella con le somme modulo 2 degli elementi di ciascuna linea e, per ultimo, si aggiunge la somma modulo 2 di tutti gli elementi della tabella di partenza.
• La successione di lunghezza $(t+1)^2$ ottenuta scrivendo di seguito le righe della nuova tabella sarà la codifica del dato di partenza.

Se in questa codifica si cambia valore in un solo posto, i controlli individuano il posto in cui è stato cambiato il valore iniziale: *il nostro codice autocorregge un errore*!

Per esempio, nel caso $n = 3$, la codifica del dato 010011100 avviene nel seguente modo:
• Si forma la tabella

$$\begin{matrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{matrix}$$

• Si orla la tabella ottenuta con la somma modulo 2 delle sue linee e di tutti i suoi elementi

$$\begin{matrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{matrix}$$

• La codifica cercata è: 0101011010011010 . □

I codici di Hamming si rivelarono particolarmente adatti ai cosiddetti canali simmetrici binari e furono subito molto utilizzati, specialmente nel caso di canali che facevano uso di onde elettromagnetiche nell'etere e di impulsi elettrici nei fili. Successivamente essi hanno dato origine alla *teoria dei codici lineari* che, al di là dell'importanza per le applicazioni, ha ormai assunto un ruolo di rilievo nell'ambito delle teorie combinatorie e oggi diverse aree della geometria combinatoria e della geometria algebrica su campi finiti hanno come applicazione proprio la costruzione di codici lineari che correggono un prefissato numero di errori ([1],[4],[6],[9],[14],[18]).

Nei paragrafi che seguono intendiamo approfondire, tralasciando naturalmente quasi tutti i dettagli dimostrativi, i concetti ed i risultati esposti nell'introduzione. Per ragioni di spazio, purtroppo, non avremo la possibilità di mostrare nessuna applicazione della geometria algebrica su campi finiti alla teoria dei codici lineari; rimandiamo comunque a [6] per eventuali approfondimenti. Avvertiamo che, data la natura divulgativa della relazione, spesso saremo costretti ad essere alquanto imprecisi e poco formali; di ciò ci scusiamo soprattutto con i Lettori più esperti di questi argomenti.

2 Richiami su piani finiti e disegni

Una delle difficoltà comune a quasi tutti i campi delle teorie combinatorie è che, variando anche di poco i parametri numerici di uno stesso problema, spesso bisogna cambiare completamente tecnica per trovare una soluzione. Ciò rende particolarmente stimolante la ricerca di metodi per superare questa sorta di *sporadicità* in modo da sintetizzare in teorie generali i risultati già noti e quelli nuovi che si spera di ottenere. Uno di tali metodi si basa sullo studio di famiglie di insiemi finiti con qualche proprietà di *regolarità.* È appunto in quest'ambito che si inquadrano le teorie dei *piani proiettivi finiti* e dei *disegni* ([1],[10],[11],[18]), delle quali intendiamo richiamare alcuni risultati che useremo per lo studio dei codici lineari.

2.1 Piani proiettivi

Sia $\pi = (\mathcal{P}, \mathcal{L})$ una coppia costituita da un insieme non vuoto $\mathcal{P}$ e da un insieme $\mathcal{L}$ di sottoinsiemi di $\mathcal{P}$. Gli elementi di $\mathcal{P}$ e $\mathcal{L}$ si chiamano rispettivamente *punti* e *rette.* L'insieme di tutte le rette contenenti un fissato punto P si dice *fascio di rette* di centro P. Un punto e una retta che si appartengono si dicono anche *incidenti.*

DEFINIZIONE 2.1. La coppia $\pi = (\mathcal{P}, \mathcal{L})$ prende il nome di *piano proiettivo* se sono verificate le seguenti proprietà:

(1) *due punti distinti appartengono ad un'unica retta;*

(2) *due rette distinte hanno esattamente un punto in comune;*

(3) *esistono quattro punti a tre a tre non appartenenti ad una stessa retta.*

Punti appartenenti ad una stessa retta si dicono *allineati.* L'unica retta contenente due fissati punti P e T sarà denotata con PT. □

In un piano proiettivo $\pi = (\mathcal{P}, \mathcal{L})$ valgono le seguenti proprietà:

(i) due rette sono equipotenti;

(ii) una retta e un fascio di rette sono equipotenti;

(iii) due fasci di rette sono equipotenti;

(iv) ogni retta contiene almeno tre punti.

ESEMPIO 2.2. Sia K un campo e si consideri la struttura geometrica, che denotiamo con $PG(2, K)$, i cui punti e rette siano rispettivamente i sottospazi di dimensione 1 e 2 dello spazio vettoriale K^3 delle terne ordinate di elementi di K. È facile rendersi conto che $PG(2, K)$ è un piano proiettivo; esso prende il nome di *piano proiettivo su* K. □

Un piano proiettivo si dice *finito* se è finito l'insieme dei suoi punti. In questo caso, se $n + 1$ è il numero dei punti di una retta, l'intero n si dice *ordine* del piano. Se $\pi = (\mathcal{P}, \mathcal{L})$ è un piano proiettivo finito d'ordine n, risulta

$$|\mathcal{P}| = |\mathcal{L}| = n^2 + n + 1.$$

Il piano proiettivo $PG(2, q)$ sul campo di Galois F_q con q elementi[1] è un piano proiettivo finito il cui ordine è evidentemente q. Ne segue che *esiste almeno un*

[1] Ricordiamo che esiste, ed è unico a meno di isomorfismi, un campo finito F_q con q elementi se, e solo se, q è potenza di un numero primo p. Il campo F_p, con p numero primo, non è altro che il campo Z_p dei resti modulo p.

piano proiettivo d'ordine una qualunque potenza di un numero primo. Il più piccolo piano proiettivo è $PG(2,2)$ e consiste di 7 punti e 7 rette. Esso ha ordine 2, prende il nome di *piano di Fano* e si può rappresentare con la Figura 2.

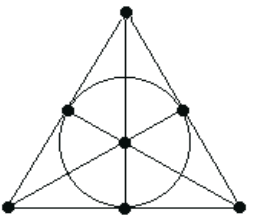

Figura 1: Il piano di Fano

Al momento sono note molte classi di piani proiettivi d'ordine $q = p^h$, p primo e $h > 1$, e non del tipo $PG(2,q)$. Non è ancora noto se esistano o meno piani proiettivi finiti d'ordine primo p che non siano del tipo $PG(2,p)$. Inoltre, *non si conoscono esempi di piani proiettivi finiti il cui ordine non sia la potenza di un primo.* Il problema dell'esistenza di tali piani è uno dei più difficili e affascinanti della geometria combinatoria. Al riguardo, nonostante l'enorme sviluppo della matematica del discreto negli ultimi cinquanta anni, il risultato più profondo rimane ancora quello pubblicato nel 1949 da ***R.H.Bruck*** e ***H.J.Ryser*** [3], i quali provarono che, *se esiste un piano proiettivo d'ordine* $n \equiv 1, 2 (mod\, 4)$, *allora* n *deve necessariamente essere la somma di due quadrati interi.* Esistono quindi infiniti interi che, in conseguenza del precedente teorema, non possono essere ordini di piani proiettivi. Per esempio, il numero di tali interi minori di 2000 è 558 e quelli minori di cento sono: 6, 14, 21, 22, 30, 33, 38, 42, 46, 54, 57, 62, 66, 69, 70, 77, 78, 86, 93, 94. I primi interi non esclusi dal teorema di Bruck e Ryser sono $n = 10$ e $n = 12$. Nel primo caso il problema è stato risolto in senso negativo [12] con tecniche di teoria dei codici lineari, che esporremo nel seguito. Il risultato, ottenuto grazie alla possibilità di utilizzare potenti elaboratori (della metà degli anni ottanta) per una ricerca esaustiva, è stato pubblicato nel 1989 e si deve a ***C.W.Lam, S.Swiercz e L.Thiel.*** Naturalmente, sarebbe molto bello trovare una dimostrazione della non esistenza di un piano proiettivo d'ordine 10 senza l'uso di strumenti di calcolo. Per $n = 12$ il problema è ancora completamente aperto e non siamo a conoscenza di risultati che facciano prevedere una sua soluzione in tempi brevi.

2.2 Disegni

Siano $\mathcal{P}$ un insieme finito con v elementi e $\mathcal{B} = \{B_1, B_2, \ldots, B_b\}$ un insieme di sottoinsiemi non vuoti di $\mathcal{P}$, ciascuno di cardinalità k. Gli elementi di $\mathcal{P}$ e di $\mathcal{B}$ si dicono rispettivamente *punti* e *blocchi.* Un punto e un blocco che si appartengano e due blocchi ad intersezione non vuota si dicono *incidenti.*

Definizione 2.3. La coppia $\mathbf{D} = (\mathcal{P}, \mathcal{B})$ prende il nome di $t-$*disegno con parametri* (v, k, λ), o di $t-(v, k, \lambda)$ *disegno,* se ogni sottoinsieme di $\mathcal{P}$ con t punti è contenuto in esattamente λ blocchi, λ e t essendo interi positivi tali che

$$v \geq k \geq t \geq 1.$$

Talvolta un $t-$disegno è anche chiamato semplicemente *disegno.* Un $t-(v, k, \lambda)$ disegno si dice *banale* o *completo* se ogni sottoinsieme di $\mathcal{P}$ d'ordine k è un blocco. Un $t-$disegno per cui $v = b$ si dice *simmetrico.* □

Nel seguito, tranne esplicito avviso, ci riferiremo esclusivamente a disegni non banali e quindi supporremo costantemente

$$v > k > t \quad \text{e} \quad \binom{v}{k} > b.$$

Esempio 2.4. Sia $PG(n, q)$ lo spazio proiettivo $n-$dimensionale su un campo di Galois di ordine q con $n > 1$, allora i punti e i sottospazi $d-$dimensionali di $PG(n, q)$, con $1 \leq d < n$, costituiscono un $2-$disegno. Denotiamo questo disegno con $PG_d(n, q)$, se è $d > 1$, e con $PG(n, q)$, se è $d = 1$. I parametri v, k, λ di $PG_d(n, q)$ sono i seguenti:

$$v = q^n + q^{n-1} + \cdots + q + 1 = \frac{q^{n+1} - 1}{q - 1},$$

$$k = q^d + q^{d-1} + \cdots + q + 1 = \frac{q^{d+1} - 1}{q - 1},$$

$$\lambda = \begin{cases} \frac{(q^{n-1}-1)(q^{n-2}-1)\cdots(q^{n-d+1}-1)}{(q^{d-1}-1)(q^{d-2}-1)\cdots(q-1)} & \text{se } d > 1 \\ 1 & \text{se } d = 1 \end{cases}.$$

Si osservi che $PG_d(n, q)$ è un disegno simmetrico se, e solo se, risulta $d = n-1$. In modo analogo si definisce il disegno $AG_d(n, q)$ dei sottospazi di dimensione d dello spazio affine $n-$dimensionale su un campo di Galois di ordine q. □

ESEMPIO 2.5. È facile verificare che i punti e le rette di un piano proiettivo finito d'ordine n definiscono un $2-$disegno simmetrico con parametri

$$v = b = n^2 + n + 1 \ , \ \ k = n + 1 \ , \ \ \lambda = 1.$$

Inoltre, ogni $2-$disegno avente gli stessi parametri di un piano proiettivo finito d'ordine n è un piano proiettivo. □

ESEMPIO 2.6. È possibile provare che esiste un solo $5-$disegno con parametri $(24, 8, 1)$, che si denota con $\mathcal{M}_{24}$. Esso è uno dei cosiddetti *disegni di Witt,* o di *Mathieu,* e si costruisce a partire dai punti e dalle rette del piano $PG(2, 4)$ mediante un procedimento detto di *estensione.* Vedremo che questo disegno ha un ruolo importante nella teoria dei codici lineari. □

Il problema di costruire $t-$disegni con parametri assegnati, e più in generale di descriverli tutti a meno di isomorfismi, è uno dei più interessanti e difficili di tutta la teoria. Basti pensare che fino al 1983 è stata in dubbio l'esistenza di $t-$disegni non banali con $t > 5$. Il primo esempio di $6-$disegno non banale, scoperto appunto nel 1983, si deve a *S.S. Magliveras* e *D.W. Leavitt* [13]. Solo nel 1987 *L. Teirlink* provò sorprendentemente che, per ogni intero positivo t, esiste un $t-$disegno non banale [16],[17].

Le due proposizioni che seguono danno alcune relazioni elementari tra i parametri di un disegno, che costituiscono anche delle condizioni necessarie di esistenza.

PROPOSIZIONE 2.7. *Sia* $\mathbf{D} = (\mathcal{P}, \mathcal{B})$ *un* $t - (v, k, \lambda)$ *disegno. Se* i *è un intero tale che* $1 \leq i \leq t$, *allora* $\mathbf{D}$ *è anche un* $i - (v, k, \lambda_i)$ *disegno con*

$$\lambda_i = \frac{(v - t + 1)(v - t + 2) \cdots (v - i)}{(k - t + 1)(k - t + 2) \cdots (k - i)} \lambda \, .$$

Posto $r = \lambda_1$, *risulta:*

$$(v - i)\lambda_{i+1} = (k - i)\lambda_i, \ \textit{per ogni} \ i < t \, , \ e \ \ vr = bk.$$

Nel caso $t = 2$ *si ha anche*

$$r(k - 1) = \lambda(v - 1) \, .$$

Se si ordinano linearmente gli insiemi dei punti e dei blocchi di un $t - (v, k, \lambda)$ disegno $\mathbf{D}$, $\mathcal{P} = \{p_1, p_2, ..., p_v\}$ e $\mathcal{B} = \{B_1, B_2, ..., B_b\}$, si può considerare la matrice sul campo razionale $A = (a_{ij})$ di tipo $b \times v$ definita da

$$a_{ij} = \begin{cases} 1, & se \quad p_j \in B_i \\ 0, & se \quad p_j \notin B_i \end{cases} \, .$$

Una tale matrice, che ovviamente dipende dagli ordinamenti scelti su $\mathcal{P}$ e $\mathcal{B}$, prende il nome di *matrice d'incidenza* di $\mathbf{D}$. È evidente che ogni disegno è completamente individuato da una sua matrice d'incidenza. Talvolta conviene considerare A come matrice su un campo diverso da quello dei razionali. Quando ciò si renderà necessario, verrà posto opportunamente in evidenza.

Esempio 2.8. Consideriamo il piano di Fano $PG(2,2)$, cioè il $2-(7,3,1)$ disegno simmetrico dei punti e delle rette del piano proiettivo sul campo di Galois d'ordine 2. Se numeriamo i punti e le rette come in Figura 2,

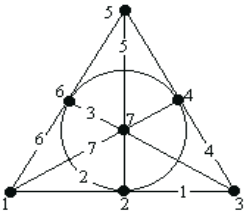

Figura 2: Il piano di Fano

la relativa matrice d'incidenza è

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \qquad \square$$

Denotiamo con I e J una matrice identità e una matrice con tutti gli elementi uguali ad 1, rispettivamente, il cui ordine sarà chiaro dal contesto.

Proposizione 2.9. *Se A è una matrice d'incidenza di un $2-(v,k,\lambda)$ disegno, allora risulta:*

$$AJ = rJ,\ JA = kJ,$$

$$AA^t = (r-\lambda)I + \lambda J, \tag{1}$$

$$det(AA^t) = rk(r-\lambda)^{v-1}. \tag{2}$$

Inoltre, se il $2-$disegno è simmetrico non banale, A è non singolare.

3 Generalità sui codici

3.1 Prime definizioni ed esempi

Un insieme finito F con q elementi, nel linguaggio della teoria dei codici, prende il nome di *alfabeto finito con q lettere,* le lettere essendo gli elementi di F. Una *parola su F di lunghezza n* è una successione finita $a_1a_2...a_n$ di lettere di F.

Nel seguito, per comodità di scrittura, identificheremo la parola $a_1a_2...a_n$ con l'$n-$pla corrispondente $\boldsymbol{a} = (a_1, a_2, ..., a_n)$; in tal modo ogni parola di lunghezza n potrà essere considerata come un elemento di F^n.

DEFINIZIONE 3.1. Un *codice* C su un alfabeto F è un qualsiasi sottoinsieme finito e non vuoto di parole su F. Esso prende il nome di *codice a blocchi* se le sue parole hanno tutte la stessa lunghezza; nel caso contrario si dice *a lunghezza variabile.* La comune lunghezza delle parole di un codice a blocchi si chiama *lunghezza del codice.* □

Riportiamo alcuni esempi di codici molto comuni i cui nomi almeno sono sicuramente noti al Lettore.

ESEMPIO 3.2. Il *Codice Fiscale Italiano* è un codice su un alfabeto di 36 lettere (le 26 dell'alfabeto inglese e le cifre decimali da 0 a 9). Le sue parole servono a codificare qualunque persona o ente abbia rapporti con il sistema fiscale italiano. Per maggiori informazioni si rimanda al sito WEB:
http://it.wikipedia.org/wiki/Codice_fiscale. □

ESEMPIO 3.3. Il *codice* $ISBN$ (*International Standard Book Number*) è un codice a blocchi di lunghezza 10 sull'alfabeto di undici lettere costituite dalle cifre decimali da 0 a 9 e dalla lettera X ed è usato per codificare i libri in commercio. Per maggiori informazioni si rimanda al sito WEB:
http://www.alice.it/bookshop/law.bks/codiinte.htm. □

ESEMPIO 3.4. L'*American Standard Code for Information Interchange,* noto come *Codice ASCII,* è il codice a blocchi sull'alfabeto $F = \{0, 1\}$ formato da tutte le parole di lunghezza sette e, quindi, contiene esattamente $2^7 = 128$ parole. Esso è stato costruito per codificare le lettere dell'alfabeto inglese maiuscole e minuscole, le cifre decimali da 0 a 9 e una serie di altri simboli e istruzioni allo scopo di permettere all'architettura interna di un computer di operare solo con i simboli 0 e 1. Se ad ogni parola di questo codice si aggiunge 0 o 1, a seconda che contenga un numero pari o dispari di 1 rispettivamente, si ottiene un codice a blocchi di lunghezza otto, detto *codice ASCII esteso.* Per maggiori informazioni si rimanda al sito WEB: *http://it.wikipedia.org/wiki/ASCII.* □

Nel seguito prenderemo in considerazione soltanto codici a blocchi. Useremo, pertanto, il termine *codice* come sinonimo di *codice a blocchi.*

Un codice su un alfabeto con q lettere che contenga esattamente M parole di lunghezza n prende il nome di $(n, M)-$*codice* $q-$*ario,* o semplicemente di (n, M)-*codice,* se q risulta chiaro dal contesto. Nei casi $q = 2, 3$ il codice si dice rispettivamente *binario* e *ternario.*

3.2 La trasmissione dell'informazione

Per *canale* (*discreto*) *di trasmissione,* o *di comunicazione,* in modo empirico e non formale, intendiamo un sistema fisico in grado di accettare in una *entrata* le lettere di un alfabeto $F = \{a_1, a_2, \cdots a_q\}$ e, in corrispondenza di ciascuna lettera accettata, emettere in una *uscita* lettere dello stesso alfabeto. Questo significa che in entrata il sistema possiede q possibili stati fisici in corrispondenza biunivoca con le lettere dell'alfabeto F e una situazione analoga si riproduce in uscita. Osserviamo esplicitamente che *escludiamo la possibilità che all'immissione di una lettera in input non corrisponda l'emissione di una lettera in output.*

Quando nel canale Σ immettiamo in successione le lettere $a_{i_1}, a_{i_2}, ..., a_{i_n}$ e in uscita troviamo nell'ordine le lettere $a_{j_1}, a_{j_2}, ..., a_{j_n}$ diremo che è stata trasmessa la parola $\boldsymbol{a}_i = (a_{i_1}, a_{i_2}, ..., a_{i_n})$ e che è stata ricevuta la parola $\boldsymbol{a}_j = (a_{j_1}, a_{j_2}, ..., a_{j_n})$; in queste ipotesi, il numero di componenti omologhe distinte tra $\boldsymbol{a}_i$ e $\boldsymbol{a}_j$ prende il nome di *numero di errori* commesso nella trasmissione della parola $\boldsymbol{a}_i$.

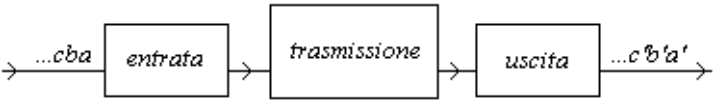

Figura 3: Canale di trasmissione

Per ogni $a_i, a_j \in F$, denotiamo con $P(a_j, a_i)$ la probabilità che immettendo nel canale Σ la lettera a_i si trovi corrispondentemente in uscita la lettera a_j e supponiamo che tale probabilità dipenda soltanto dalla coppia (a_i, a_j). In queste ipotesi, il canale di trasmissione si dice *senza memoria* e, posto

$$p_{ij} := P(a_j, a_i), \tag{3}$$

la matrice

$$P := (p_{ij})$$

si chiama *matrice del canale rispetto all'alfabeto* F. Quando l'alfabeto è chiaro dal contesto si parla semplicemente di *matrice di* Σ. Naturalmente, poiché ogni riga di P contiene le probabilità di tutte le lettere che in uscita possono corrispondere all'immissione della lettera relativa alla riga scelta, la somma degli elementi su ogni riga di P è uguale ad 1, cioè

$$0 \leq p_{ij} \leq 1 \quad e \quad \sum_{j=1}^{q} p_{ij} = 1\,;$$

proprietà che si esprimono anche dicendo che P è una *matrice stocastica.*

ESEMPIO 3.5. Nell'ambito dei canali di comunicazione senza memoria sono particolarmente importanti i *canali simmetrici*. Un canale di questo tipo è definito dalla proprietà che la probabilità p che una lettera a_i in input sia trasformata in output in una lettera diversa a_j non dipende da a_i e a_j, ma è la stessa per tutte le coppie di lettere distinte; in altre parole, nella matrice del canale P risulta $p_{ij} = p$, per ogni coppia (i, j), con $i \neq j$. Il numero p si chiama *probabilità d'errore* del canale. La matrice di un canale simmetrico rispetto ad un alfabeto con m lettere è, dunque, del tipo

$$P = \begin{bmatrix} 1-(m-1)p & p & \dots & p \\ p & 1-(m-1)p & \dots & p \\ \vdots & \vdots & \ddots & \vdots \\ p & p & \dots & 1-(m-1)p \end{bmatrix}.$$

In particolare, sono molto usati i *canali simmetrici binari*. Questi sono caratterizzati dall'operare con un alfabeto binario, per esempio $\{0, 1\}$, e ciascuna lettera in input ha la stessa probabilità p di essere trasformata nell'altra in output e, quindi, la matrice del canale è data da

$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}. \qquad \square$$

Per rendere più chiaro il problema della correzione automatica degli errori, descriviamo a grandi linee e in modo non formale le trasformazioni cui viene sottoposto un messaggio immesso in un sistema di comunicazione che, molto schematicamente, supporremo composto da:

• una *stazione emittente* **E**;

• un *canale di trasmissione senza memoria* Σ, con *codice* C e *matrice* P;

• una *stazione ricevente* **R**.

La stazione **E** può inviare ad **R** messaggi scelti in un insieme prefissato $\mathcal{M}$ (*insieme dei messaggi con informazione compressa*); i messaggi in $\mathcal{M}$ possono essere identificati con parole del codice C (*insieme dei messaggi con aggiunta di ridondanza controllata*) mediante una funzione iniettiva γ tra $\mathcal{M}$ e C (*funzione di codifica*).

Supponiamo che l'emittente **E** debba inviare un messaggio *M*, scelto in $\mathcal{M}$, alla stazione ricevente **R**. Prima di essere trasmesso, *M* deve essere trasformato in una parola $\boldsymbol{x}$ del codice C. Questa operazione, detta *codifica*, si realizza mediante un algoritmo che calcola automaticamente il valore su *M* della funzione di codifica γ; nel nostro caso $\gamma(M) = \boldsymbol{x}$. A questo punto, la parola $\boldsymbol{x}$ viene immessa nel canale Σ e viene ricevuta in uscita una parola $\boldsymbol{y}$ che, a causa di disturbi del canale, può essere diversa da $\boldsymbol{x}$.

La parola $\boldsymbol{y}$ viene finalmente tradotta (*decodificata*) in un messaggio *M'*$\in \mathcal{M}$ mediante un *algoritmo di decodifica* che opera nel seguente modo:

• se $\boldsymbol{y} \in C$, pone *M'*$= \gamma^{-1}(\boldsymbol{y})$;

• se $\boldsymbol{y} \notin C$, individua la parola $\boldsymbol{z}$ di C in qualche modo più simile a $\boldsymbol{y}$ e pone *M'*$= \gamma^{-1}(\boldsymbol{z})$.

Per una buona trasmissione, dunque, nell'ipotesi $\boldsymbol{y} \neq \boldsymbol{x}$, non deve accadere che $\boldsymbol{y} \in C$, altrimenti non vi è alcun modo di capire che vi sono stati degli errori e il destinatario riceve inevitabilmente un messaggio sbagliato. Inoltre, alla fine del processo di trasmissione deve ottenersi $\boldsymbol{z} = \boldsymbol{x}$, cioè $\boldsymbol{x}$ deve potersi riconoscere a partire da $\boldsymbol{y}$. Queste condizioni, che riterremo le *condizioni di affidabilità* del sistema, nell'ipotesi che C abbia lunghezza n, sono per esempio soddisfatte se sono verificate le seguenti proprietà:

• $P(\boldsymbol{a}, \boldsymbol{a}) \neq 0$, per ogni parola $\boldsymbol{a} \in C$;

• $P(\boldsymbol{b}, \boldsymbol{a}) = 0$ per ogni $\boldsymbol{a}, \boldsymbol{b}$ parole distinte di C;

• se $\boldsymbol{a} \in C$ e $P(\boldsymbol{z}, \boldsymbol{a}) \neq 0$, allora $P(\boldsymbol{z}, \boldsymbol{b}) = 0$ per ogni parola $\boldsymbol{b}$ di C diversa da $\boldsymbol{a}$.

Sotto queste ipotesi, infatti, se una parola $\boldsymbol{y}$ non appartiene a C, esiste al più una parola $\boldsymbol{x}$ di C tale che $P(\boldsymbol{y}, \boldsymbol{x}) \neq 0$ e di conseguenza, almeno teoricamente, è possibile descrivere un algoritmo per la correzione degli errori e, quindi, per la decodifica. A tal fine, basta osservare che gli insiemi

$$B_{\boldsymbol{a}} = \{\boldsymbol{z} \in F^n \ : \ P(\boldsymbol{z}, \boldsymbol{a}) \neq 0\} \ ,$$

al variare di $\boldsymbol{a} \in C$, sono a due a due disgiunti e, quindi, se si riceve $\boldsymbol{y} \in B_{\boldsymbol{x}}$, con

$\boldsymbol{x} \in C$, si può dedurre che $\boldsymbol{x}$ è la parola di C inizialmente trasmessa.

Nel seguito diremo che un sistema di comunicazione è *affidabile* se sono verificate le precedenti tre condizioni. In tali ipotesi, il processo della trasmissione di un messaggio può rappresentarsi con la Figura 4.

3.3 Distanza di Hamming e correzione degli errori

Esponiamo alcuni elementi della teoria dei codici correttori ([2],[9]) che, come abbiamo osservato, si può considerare il supporto teorico di base dei problemi descritti nell'introduzione. A tale scopo introduciamo un concetto di distanza fra parole dovuto a *R.Hamming.* Ciò permetterà anche di illustrare e giustificare un principio di decodifica, detto *nearest neighbour decoding,* che è alla base della teoria che intendiamo esporre. Tale principio si presta ad essere usato con successo nei sistemi di comunicazione che utilizzano canali di trasmissione simmetrici (*cfr.*Esempio3.5) e codici lineari, dei quali parleremo più avanti.

DEFINIZIONE 3.6. Se F^n, $n > 0$, è l'insieme di tutte le parole di lunghezza n su un alfabeto F e se $\boldsymbol{x}, \boldsymbol{y}$ sono due tali parole, si definisce *distanza di Hamming* tra $\boldsymbol{x}$ e $\boldsymbol{y}$, e si denota con $d(\boldsymbol{x}, \boldsymbol{y})$, il numero di posizioni in cui $\boldsymbol{x} = (x_1, x_2, \cdots x_n)$ e $\boldsymbol{y} = (y_1, y_2, \cdots y_n)$ presentano lettere differenti, cioè

$$d(\boldsymbol{x}, \boldsymbol{y}) = |\{i \ : \ x_i \neq y_i, \ \ 1 \leq i \leq n\}|.$$

La funzione d è una metrica su F^n, detta *metrica di Hamming.* □

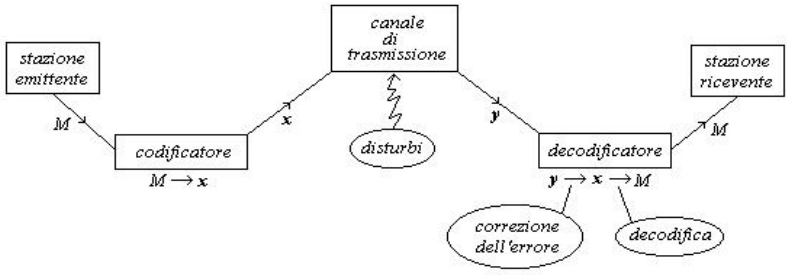

Figura 4: Sistema di comunicazione affidabile

Seguendo ora lo schema della Figura 4, supponiamo che una parola $\boldsymbol{x}$ di un fissato codice C venga trasmessa da una emittente e modificata in una parola $\boldsymbol{y} \neq \boldsymbol{x}$ dal canale di trasmissione. Se $\boldsymbol{y}$ è ancora una parola di C, è chiaro che non esiste

alcuna possibilità di scoprire lo scambio di parole. Se invece $\boldsymbol{y}$ non appartiene al codice C, allora è evidente che c'è stato un errore e di conseguenza si pone il problema di correggerlo, cioè di risalire dalla parola ricevuta $\boldsymbol{y}$ a quella $\boldsymbol{x}$ effettivamente trasmessa.

Ad esempio, supponiamo di avere a disposizione un canale di trasmissione T per il quale sia molto alta la probabilità che il massimo numero di lettere di una parola che si possono modificare nel corso di una trasmissione sia più piccolo della metà della minima distanza fra due qualsiasi parole distinte del codice C. In queste ipotesi, se la parola ricevuta $\boldsymbol{y}$ non appartiene a C, esiste generalmente un'unica parola $\boldsymbol{z} \in C$ a distanza minima da $\boldsymbol{y}$[2]; così il decodificatore sostituisce automaticamente $\boldsymbol{y}$ con $\boldsymbol{z}$ e la probabilità che sia $\boldsymbol{z} = \boldsymbol{x}$ è estremamente alta. Il principio appena esposto, secondo il quale si decodifica la parola del codice a distanza minima da quella ricevuta, prende il nome di *nearest neighbour decoding.*

Poiché nella realtà non esistono canali di trasmissione immuni da disturbi, quanto finora detto suggerisce di non scegliere mai il codice C uguale ad F^n. In altre parole, *un buon codice deve essere un sottoinsieme proprio di* F^n con la proprietà che ogni sua parola, al fine di una buona decodifica, oltre a contenere il minimo numero di lettere necessarie per la codifica del messaggio associato, contenga delle ulteriori lettere di *controllo.* Queste aiutano a ricostruire la parola trasmessa nel caso non si siano verificati troppi errori durante la trasmissione.

Nel seguito, tranne avviso contrario, riterremo fissato un $(n, M)-$codice C su un alfabeto F con q lettere. Inoltre, spesso identificheremo C con la matrice su F le cui righe sono le parole di C, preventivamente ordinate. Una tale matrice ha M righe ed n colonne e si dice *associata* a C. Due matrici associate ad uno stesso codice differiscono, quindi, per una permutazione delle righe.

Definizione 3.7. Si chiama *distanza minima* di un $(n, M)-$codice C l'intero $d(C)$ dato dalla più piccola distanza fra due parole distinte di C, cioè

$$d(C) := min\{d(\boldsymbol{x}, \boldsymbol{y}) \ : \ \boldsymbol{x}, \boldsymbol{y} \in C \ , \ \boldsymbol{x} \neq \boldsymbol{y}\}.$$

Se non vi è possibilità di equivoci scriveremo d in luogo di $d(C)$. Un $(n, M)-$-codice di distanza minima d si dice anche un (n, M, d)-*codice* e gli interi n, M, d si dicono *parametri* del codice. □

Diciamo che il codice C è $k-$*sistematico* o semplicemente *sistematico,* se in una delle sue matrici associate esistono k colonne di posto $i_1, i_2, ..., i_k$ tali che, per

[2]Si osservi che questa proprietà è falsa se il numero di lettere di $\boldsymbol{x}$ modificate nel corso della trasmissione è maggiore della metà della minima distanza fra due qualsiasi parole distinte del codice C.

ogni $k-$pla $(\alpha_1, \alpha_2, ..., \alpha_k)$ di lettere di F, esiste un'unica parola $(a_1, a_2, \ldots, a_n)$ di C per cui risulta

$$a_{i_1} = \alpha_1,\ a_{i_2} = \alpha_2, ... \ , a_{i_k} = \alpha_k.$$

Gli interi $i_1, i_2, ..., i_k$ si chiamano anche *posti di informazione.* In queste ipotesi l'intero $n - k$ prende il nome di *ridondanza* di C e si dicono *ridondanti* o *di controllo* le lettere delle parole di C che occupano posizioni diverse da $i_1, i_2, ..., i_k$.

Per ogni parola $\boldsymbol{x} \in F^n$ e per ogni intero positivo r, consideriamo la *sfera di centro $\boldsymbol{x}$ e raggio r* (*sfera di Hamming*), cioè l'insieme

$$S(\boldsymbol{x}, r) := \{\boldsymbol{y} \in F^n \ : \ d(\boldsymbol{x}, \boldsymbol{y}) \leq r\}.$$

L'insieme

$$\overline{S}(\boldsymbol{x}, r) := \{\boldsymbol{y} \in F^n \ : \ d(\boldsymbol{x}, \boldsymbol{y}) = r\}$$

prende il nome di *superficie sferica di centro $\boldsymbol{x}$ e raggio r* e risulta

$$\begin{cases} \overline{S}(\boldsymbol{x}, r) \cap \overline{S}(\boldsymbol{x}, r') = \emptyset, \quad r \neq r', \quad r, r' \leq n \\ S(\boldsymbol{x}, r) = \bigcup\limits_{0 \leq s \leq r} \overline{S}(\boldsymbol{x}, s) \end{cases} \tag{4}$$

PROPOSIZIONE 3.8. *Una sfera di raggio r, $0 \leq r \leq n$, in F^n contiene esattamente*

$$\binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{r}(q-1)^r$$

parole.

DIMOSTRAZIONE. Le parole a distanza s da una fissata parola $\boldsymbol{x}$ si ottengono modificando arbitrariamente s lettere di $\boldsymbol{x}$ e quindi sono esattamente

$$\binom{n}{s}(q-1)^s.$$

Dalle (4) segue allora l'asserto. □

DEFINIZIONE 3.9. Si dice che un $(n, M)-$codice C *scopre k errori,* ove k è un intero positivo, se la sfera $S(\boldsymbol{x}, k)$ ha in comune con C la sola parola $\boldsymbol{x}$, per ogni $\boldsymbol{x} \in C$. Si dice poi che C *corregge k errori,* se scopre k errori e due qualsiasi sfere di raggio k con centri in parole distinte di C sono ad intersezione vuota. □

Queste definizioni sono del tutto naturali se si pensa alle osservazioni fatte precedentemente, specialmente nell'Esempio 1.1, e la proposizione che segue è una loro immediata conseguenza.

Proposizione 3.10. *Un $(n, M)-$codice C scopre k errori se, e soltanto se, risulta $d \geq k+1$ e corregge h errori se, e soltanto se, risulta $d \geq 2h+1$.*

Di solito il massimo numero di errori che un codice può correggere viene denotato con e e, in questo caso, il codice si dice $e-$*correttore.* La Prop.3.10 assicura che, se d è la minima distanza di C, allora è

$$d = 2e+1 \quad \text{o} \quad d = 2e+2,$$

a seconda che d sia dispari o pari, rispettivamente.

Proposizione 3.11. (Disuguaglianza di Hamming) *Per ogni $(n, M)-$codice C che sia $e-$correttore, risulta*

$$M\left[\binom{n}{0} + \binom{n}{1}(q-1) + \cdots + \binom{n}{e}(q-1)^e\right] \leq q^n. \tag{5}$$

Dimostrazione. La Prop.3.10 assicura che sfere di centro due parole distinte di C e raggio e sono fra loro disgiunte. Dalla Prop.3.8 segue allora l'asserto. □

Definizione 3.12. I codici i cui parametri verificano l'uguaglianza nella (5) si dicono *perfetti.* □

Esempio 3.13. Il codice $C = F^n$, i codici contenenti una sola parola e i codici di ripetizione binari di lunghezza dispari, definiti nel seguito dalla (7), sono perfetti; essi vengono detti *codici perfetti banali.* □

Osservazione 3.14. È facile verificare che un $(n, M, d)-$codice C che sia $e-$correttore è perfetto se, e soltanto se, le sfere di raggio e e centro le parole di C formano una partizione di F^n. Ne segue che, se C è perfetto, allora d deve essere dispari. Inoltre, la proprietà di un codice di essere perfetto dipende esclusivamente dai suoi parametri. Questo significa che, se C è perfetto, ogni codice con gli stessi parametri di C è anch'esso perfetto. □

Terminiamo il paragrafo con un esempio di codice perfetto non banale.

Esempio 3.15. Sia

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \\ \boldsymbol{a}_4 \\ \boldsymbol{a}_5 \\ \boldsymbol{a}_6 \\ \boldsymbol{a}_7 \end{matrix}$$

una matrice d'incidenza del piano di Fano $PG(2,2)$ e indichiamo con $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, $\boldsymbol{a}_4$, $\boldsymbol{a}_5$, $\boldsymbol{a}_6$, $\boldsymbol{a}_7$ le sue righe. Aggiungiamo ad A le righe $\mathbf{0} = (0,0,0,0,0,0,0)$, $\mathbf{1} = (1,1,1,1,1,1,1)$ e le righe $\boldsymbol{b}_i$ che si ottengono dalle $\boldsymbol{a}_i$ scambiando tra loro 1 e 0. Otteniamo così la matrice

$$H(3,2) = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{matrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \\ \boldsymbol{a}_4 \\ \boldsymbol{a}_5 \\ \boldsymbol{a}_6 \\ \boldsymbol{a}_7 \\ \mathbf{0} \\ \mathbf{1} \\ \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \boldsymbol{b}_3 \\ \boldsymbol{b}_4 \\ \boldsymbol{b}_5 \\ \boldsymbol{b}_6 \\ \boldsymbol{b}_7 \end{matrix}$$

È facile verificare che le righe di H costituiscono le parole di un $(7,16,3)$-$-$codice. Più precisamente, per $i \neq j$, si ha: $d(\mathbf{0},\boldsymbol{a}_i) = 3$, $d(\mathbf{0},\boldsymbol{b}_i) = 4$, $d(\mathbf{0},\mathbf{1}) = 7$, $d(\boldsymbol{a}_i,\mathbf{1}) = 4$, $d(\boldsymbol{b}_i,\mathbf{1}) = 3$, $d(\boldsymbol{a}_i,\boldsymbol{a}_j) = 4$, $d(\boldsymbol{a}_i,\boldsymbol{b}_i) = 7$, $d(\boldsymbol{a}_i,\boldsymbol{b}_j) = 3$, $d(\boldsymbol{b}_i,\boldsymbol{b}_j) = 4$. Il codice appena definito si chiama *codice binario associato al piano di Fano* ed è un semplice esercizio verificare che è perfetto. È questo un esempio di codice costruito a partire da un disegno. Notiamo che se 0 e 1 si pensano come gli elementi del campo di Galois F_2 con 2 elementi, allora le parole del codice formano un sottospazio vettoriale di dimensione 4 dello spazio vettoriale F_2^7. □

3.4 Il problema fondamentale della teoria dei codici

Trovare codici che, tenuto conto del grado di affidabilità del sistema di comunicazione, assicurino con la massima efficienza possibile una trasmissione fedele dell'informazione è uno dei problemi centrali nella teoria delle comunicazioni. Quanto esposto nei precedenti paragrafi giustifica il fatto che ad un buon codice si richiede che abbia:

• lunghezza n *abbastanza piccola* per permettere una trasmissione veloce delle sue parole;

• un numero M di parole *abbastanza grande* per codificare una buona quantità di informazioni;

• distanza minima *abbastanza grande* per correggere il maggior numero possibile di errori (*cfr.*Prop.3.10).

Chiaramente queste richieste sono tra loro contrastanti e di conseguenza non è possibile ottimizzare uno dei parametri senza avere preventivamente fissato gli altri due. Per esempio, fissati n e d, il calcolo del più grande intero $M = A_q(n, d)$, per cui esiste un (n,M, d)-codice $q-$ario, è noto come *problema fondamentale della teoria dei codici.* Questo problema è tra i più difficili della teoria e, al momento, si conoscono pochi risultati al riguardo. Due casi molto semplici sono i seguenti:

$$\begin{cases} A_q(n,1) = |F^n| = q^n \\ \\ A_q(n,n) = |F| = q \end{cases} . \tag{6}$$

La condizione $d = 1$, infatti, impone soltanto che le parole di C siano tutte distinte fra loro e quindi il massimo valore di M si ottiene per $C = F^n$, cioè la prima delle (6). Se invece abbiamo $d = n$, allora le lettere che figurano in una fissata posizione nelle parole di C devono essere a due a due distinte e quindi $A_q(n, n) \leq q$. D'altra parte il codice

$$C(F, n) = \{(a, a, ..., a) \quad : \quad a \in F\} \tag{7}$$

contiene esattamente q parole e gode della proprietà richiesta; abbiamo così la seconda delle (6). Il codice $C(F, n)$ si chiama *codice di ripetizione $q-$ario di lunghezza n* o anche $(n, q, n)-$*codice di ripetizione su* F.
Un altro risultato noto è il seguente

$$A_q(4, 3) = q^2,$$

che si ottiene usando la *teoria dei quadrati latini.*

4 Codici lineari

D'ora in avanti supporremo che q sia potenza di un numero primo p e che l'alfabeto $F = F_q$ sia il campo di Galois con q elementi, per cui F^n è lo spazio vettoriale di dimensione n su F. Un *codice lineare* è, per definizione, un sottospazio vettoriale di F^n e ogni matrice ad esso associata è una matrice sul campo F. Il codice $ASCII$, il codice $ASCII$ esteso e quello associato al piano di Fano sono esempi di codici lineari su F_2. Se C è un codice lineare di dimensione k e distanza minima d, parleremo di $[n, k, d]-$*codice,* o di $[n, k]-$*codice* e diremo che n, k e d sono i suoi *parametri.* Un tale codice è chiaramente un $(n, q^k, d) - codice$.

4.1 Prime definizioni ed esempi

Nel seguito riterremo fissato un $[n, k, d]-$codice C e denoteremo sempre con $\mathbf{0}$ il vettore nullo (*parola nulla*).

DEFINIZIONE 4.1. Si chiama *peso* di una parola $\boldsymbol{a} \in F^n$, e si denota con $w(\boldsymbol{a})$, il numero delle componenti di $\boldsymbol{a}$ diverse da zero, cioè la distanza di $\boldsymbol{a}$ dalla parola nulla. Il minimo $w(C)$ dei pesi delle parole di C diverse da $\mathbf{0}$ è per definizione il *peso minimo* di C, cioè

$$w(C) := min\{w(\boldsymbol{a}) : \boldsymbol{a} \in C, \boldsymbol{a} \neq \mathbf{0}\}.$$

Quando non vi è possibilità di equivoci, scriveremo w in luogo di $w(C)$. □

Se due parole $\boldsymbol{a}$ e $\boldsymbol{b}$ di C hanno distanza h, allora la parola $\boldsymbol{a} - \boldsymbol{b}$, che è ancora in C, ha peso h. Ne segue che in ogni $[n, k, d]-$codice risulta

$$d = w. \tag{8}$$

Per trovare la distanza minima del codice C basta calcolare i pesi delle $M - 1 = q^k - 1$ parole di C diverse da $\mathbf{0}$. Questo è un primo vantaggio offerto dalla proprietà di linearità di un codice. Senza questa ipotesi, infatti, per determinare la distanza minima di un $(n, M)-$codice occorre calcolare $M(M-1)/2$ distanze e tale numero è maggiore di $M-1$ non appena è $M > 2$. Un secondo e importante vantaggio di un codice lineare è che esso può essere descritto completamente da una sua base. Pertanto, nello studio di un codice lineare C, risultano di particolare importanza le matrici le cui righe costituiscono una base di C. Tali matrici si chiamano *matrici generatrici* del codice e, se questo ha parametri $[n, k, d]$, sono di tipo $k \times n$.

ESEMPIO 4.2. Il codice

$$C = \{(0,0,0), (1,1,0), (1,0,1), (0,1,1)\}$$

è un $[3, 2, 2]-$codice binario con matrice generatrice

$$G = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}. \qquad \square$$

ESEMPIO 4.3. Il $[7, 4, 3]-$codice associato al piano di Fano $PG(2, 2)$ (*cfr.* Esempio 3.15) ha matrice generatrice

$$G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}. \qquad \square$$

ESEMPIO 4.4. Il codice di ripetizione $q-$ario di lunghezza n è un $[n, 1, n]$-codice con matrice generatrice

$$G = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix}. \qquad \square$$

DEFINIZIONE 4.5. Diciamo che due $[n, k, d]-$codici sullo stesso campo sono *equivalenti* se due matrici ad essi rispettivamente associate possono ottenersi l'una dall'altra mediante una successione finita di operazioni dei seguenti tipi:
(A) *scambio di due colonne* (*questa operazione equivale a scambiare tra loro in ogni parola del codice le lettere che si trovano in due posizioni fissate*)*;*
(B) *moltiplicazione degli elementi di una fissata colonna per uno scalare non nullo.* $\square$

Lo studio dei codici si fa a meno di equivalenze. Con semplici argomenti di algebra lineare si può provare la seguente proposizione.

PROPOSIZIONE 4.6. *Due matrici G e G' su F generano codici lineari equivalenti se, e soltanto se, si ottengono l'una dall'altra mediante un numero finito di operazioni elementari del tipo seguente:* (R1) *scambio di due righe,* (R2) *moltiplicazione di una riga per uno scalare non nullo,* (R3) *sostituzione di una riga con la somma di quest'ultima e di un'altra riga,* (C1) *scambio di due colonne,* (C2) *moltiplicazione di una colonna per uno scalare non nullo.*

È utile osservare che una matrice generatrice G di un $[n, k, d]-$codice C, mediante operazioni di cui alla Prop.4.6 , può sempre mettersi nella forma

$$[I_k, A], \tag{9}$$

ove I_k è la matrice identità d'ordine k e A una matrice di tipo $k \times (n-k)$. La (9), che rappresenta la matrice generatrice di un codice equivalente a C, prende il nome di *forma standard* di G. Usando questa forma standard è facile provare che *ogni codice lineare di dimensione k è $k-$sistematico.*

Ricordiamo che il *prodotto scalare* (*standard*) di due vettori $\boldsymbol{a}$, $\boldsymbol{b}$ di F_q^n, che denotiamo con $\boldsymbol{ab}$, è definito da

$$\boldsymbol{ab} = (a_1, a_2, ..., a_n)(b_1, b_2, ..., b_n) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

e verifica le seguenti proprietà:

$$\begin{cases} \boldsymbol{ab} = \boldsymbol{ba} \\ \\ (\lambda\boldsymbol{a} + \mu\boldsymbol{b})\boldsymbol{c} = \lambda(\boldsymbol{ac}) + \mu(\boldsymbol{bc}) \end{cases}$$

per ogni $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in F_q^n$ e $\lambda, \mu \in F$. Due vettori $\boldsymbol{a}, \boldsymbol{b}$ per cui è $\boldsymbol{ab} = 0$ si dicono *ortogonali* e, per ogni sottoinsieme A di F_q^n, $A^\perp$ denota il sottospazio ortogonale ad A, cioè il sottospazio dei vettori ortogonali a tutti i vettori di A. Se W è un sottospazio $k-$dimensionale di F_q^n con base

$$\{\boldsymbol{g}_i = (g_{i1}, g_{i2}, ..., g_{in}) \ , \ i = 1, 2, ..., k\},$$

allora $W^\perp$ è il sottospazio costituito dai vettori $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ tali che

$$\begin{cases} g_{11}x_1 + g_{12}x_2 + \cdots + g_{1n}x_n = 0 \\ g_{21}x_1 + g_{22}x_2 + \cdots + g_{2n}x_n = 0 \\ \vdots \\ g_{k1}x_1 + g_{k2}x_2 + \cdots + g_{kn}x_n = 0 \end{cases}$$

e quindi, avendo la matrice (g_{ij}) rango k, risulta

$$dim(W) + dim(W^\perp) = n. \tag{10}$$

La (10) assicura che, se C è un $[n, k]-$codice, il sottospazio $C^\perp$ ortogonale a C è un codice con parametri $[n, n-k]$. Esso prende il nome di *codice duale* di C e, detta G una matrice generatrice di C, risulta

$$\boldsymbol{a} \in C^\perp \quad \Leftrightarrow \quad \boldsymbol{a}G^t = \boldsymbol{0} \ .$$

Una matrice generatrice H di $C^\perp$ prende il nome di *matrice di controllo* (*di parità*) di C e gode delle seguenti proprietà:

$$\begin{cases} GH^t = 0, \\ \\ C = \{\boldsymbol{x} \in F_q^n \ : \ \boldsymbol{x}H^t = 0\}. \end{cases} \tag{11}$$

La (11) mostra che il codice C può essere completamente descritto da una sua matrice controllo di parità, cosa molto utile nelle applicazioni. Se supponiamo $G = [I_k, A]$ in forma standard, allora è facile rendersi conto che

$$H = [-A^t \ , \ I_{n-k}]$$

è una matrice controllo di parità di C.

DEFINIZIONE 4.7. Quando C è contenuto in $C^\perp$, diciamo che è un codice *autoortogonale.* Se un codice autoortogonale C coincide con $C^\perp$, diciamo che C è *autoduale.* □

Per un codice autoortogonale, la (10) dice che

$$n = dim(C) + dim(C^{\perp}) \geq 2dim(C),$$

da cui abbiamo

$$\begin{cases} C \text{ autoortogonale} & \Rightarrow \quad dim(C) \leq \frac{n}{2}, \\ \\ C \text{ autoduale} & \Rightarrow \quad dim(C) = \frac{n}{2}. \end{cases} \tag{12}$$

Per ogni parola $\boldsymbol{a} = (a_1, a_2, ..., a_n)$ su F_q, diciamo *controllo di parità* di $\boldsymbol{a}$ l'opposto della somma delle sue componenti, cioè $a_0 = -(a_1 + a_2 + \cdots + a_n)$. Fissato allora l' $[n, k, d]-$codice C, possiamo considerare l' $[n+1, k]-$codice $\overline{C}$ definito da

$$\overline{C} = \{\boldsymbol{a}' = (a_0, a_1, a_2, ..., a_n) \ : \ \boldsymbol{a} \in C\},$$

il quale prende il nome di *codice esteso* di C. Un'utile proprietà di $\overline{C}$ è che tutte le sue parole hanno controllo di parità nullo. Per esempio, il codice $ASCII$ esteso è proprio il codice esteso di $C = F_2^7$. Osserviamo che, se C ha matrice controllo di parità H, allora

$$\overline{H} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & & & \\ \vdots & & H & \\ 0 & & & \end{bmatrix}$$

è una matrice controllo di parità per $\overline{C}$.

4.2 Codifica e decodifica di un codice lineare

Dedichiamo questo paragrafo ad una breve esposizione riguardante la codifica e la decodifica di un $[n, k]-$codice lineare C $e-$correttore su F_q. A tale scopo premettiamo alcune osservazioni e definizioni.

Distribuiamo tutte le parole di F_q^n in una matrice $\Sigma = (\sigma_{ij})$ con q^{n-k} righe e q^k colonne in modo che siano soddisfatte le seguenti proprietà:

(i) *la prima riga contiene tutte le parole di C ed è $\sigma_{11} = \mathbf{0}$;*

(ii) *per ogni indice di riga i, la parola $\mathbf{a}_i = \sigma_{i1}$ è di peso minimo rispetto a quelle contenute nella riga scelta e nelle righe successive;*

(iii) *per ogni coppia (i, j) di indici è $\sigma_{ij} = \mathbf{a}_i + \sigma_{1j}$.*

Una matrice Σ così costruita prende il nome di *tabella standard* di C ed è chiaro che la sua riga $i-$esima, per ogni indice i, contiene tutte le parole del laterale

$$\boldsymbol{a}_i + C = \{\boldsymbol{a}_i + \boldsymbol{c} \ : \ \boldsymbol{c} \in C\}$$

di C. Inoltre, ogni laterale di C ha le sue parole distribuite su una riga di Σ. La parola di un laterale di C che occupa la prima posizione nella corrispondente riga di Σ prende il nome di *direttrice* del laterale.

Un semplice algoritmo per costruire una tabella standard Σ di C è il seguente:

primo passo: distribuire le parole di C sulla prima riga di Σ con l'unica condizione $\sigma_{11} = \mathbf{0}$;
secondo passo: scegliere una parola $\boldsymbol{a}_2$ di peso minimo in $F_q^n \setminus C$ e porre $\sigma_{21} = \boldsymbol{a}_2$;
terzo passo: distribuire sulla seconda riga di Σ le parole di $\boldsymbol{a}_2 + C$ in modo che sia $\sigma_{2j} = \boldsymbol{a}_2 + \sigma_{1j}$;
quarto passo: scegliere una parola $\boldsymbol{a}_3$ di peso minimo in $F_q^n \setminus \{C \cup (\boldsymbol{a}_2 + C)\}$ e porre $\sigma_{31} = \boldsymbol{a}_3$;
quinto passo: distribuire sulla terza riga di Σ le parole di $\boldsymbol{a}_3 + C$ in modo che sia $\sigma_{3j} = \boldsymbol{a}_3 + \sigma_{1j}$;
..... continuare in questo modo fino all'esaurimento delle parole di F_q^n.

DEFINIZIONE 4.8. Per ogni vettore $\boldsymbol{a} \in F_q^n$, diciamo *sindrome* di $\boldsymbol{a}$ il vettore $S(\boldsymbol{a}) \in F_q^{n-k}$ definito da

$$S(\boldsymbol{a}) = \boldsymbol{a}H^t,$$

H essendo una matrice controllo di parità di C. □

PROPOSIZIONE 4.9. *Se C è un $[n,k]-$codice, risulta*

$$\boldsymbol{a} \in C \Leftrightarrow S(\boldsymbol{a}) = \mathbf{0}.$$

Inoltre, due vettori $\boldsymbol{a}, \boldsymbol{b} \in F_q^n$ hanno la stessa sindrome se, e soltanto se, appartengono ad uno stesso laterale di C in F_q^n. Ne segue che le sindromi sono in corrispondenza biunivoca con i laterali di C in F_q^n.

DIMOSTRAZIONE. La prima parte è ovvia. Per la seconda basta osservare che, detti $\boldsymbol{a}$ e $\boldsymbol{b}$ due vettori di F_q^n, risulta

$$\boldsymbol{a} + C = \boldsymbol{b} + C \Leftrightarrow \boldsymbol{b} - \boldsymbol{a} \in C \Leftrightarrow (\boldsymbol{b} - \boldsymbol{a})H^t = \mathbf{0} \Leftrightarrow$$
$$\boldsymbol{a}H^t = \boldsymbol{b}H^t \Leftrightarrow S(\boldsymbol{a}) = S(\boldsymbol{b}).$$

□

Torniamo ora al problema della codifica e della decodifica. L'operazione di codifica consiste sostanzialmente nel porre in corrispondenza biunivoca q^k messaggi assegnati con le parole di C e ovviamente non è restrittivo supporre che

l'insieme dei messaggi sia l'insieme di tutti i vettori di F_q^k, lo spazio vettoriale $k-$dimensionale su F_q. Se G è una matrice generatrice di C, della quale denotiamo con $\boldsymbol{g}_i$ i vettori riga, si sceglie come *funzione di codifica* l'applicazione

$$\boldsymbol{a} = (a_1, a_2, ..., a_k) \in F_q^k \to a_1\boldsymbol{g}_1 + a_2\boldsymbol{g}_2 + \cdots + a_k\boldsymbol{g}_k \in C,$$

che è un isomorfismo di spazi vettoriali. Allora, avendosi

$$a_1\boldsymbol{g}_1 + a_2\boldsymbol{g}_2 + \cdots + a_k\boldsymbol{g}_k = \boldsymbol{a}G,$$

l'*algoritmo di codifica* è semplicemente il prodotto (righe per colonne) di vettori numerici di lunghezza k per la matrice G. Di solito la matrice G è data in forma standard $G = [I_k, A]$. In questo caso, le prime k lettere di $\boldsymbol{a}G$ coincidono ordinatamente con le componenti di $\boldsymbol{a}$, rappresentano cioè il messaggio, mentre le rimanenti $n - k$ sono quelle che abbiamo chiamato lettere di controllo. Sia dunque

$$\boldsymbol{x} = \boldsymbol{a}G$$

la parola di C con la quale è stato codificato il messaggio $\boldsymbol{a}$ e supponiamo che questa venga trasmessa e ricevuta in errore; il decodificatore riceva cioè una parola $\boldsymbol{y} \neq \boldsymbol{x}$. Supponiamo inoltre $d(\boldsymbol{x}, \boldsymbol{y}) \leq e$. In questo caso $\boldsymbol{y} \notin C$ e il decodificatore deve risalire in modo automatico a $\boldsymbol{x}$ secondo il principio del *nearest neighbour decoding;* deve quindi usare un algoritmo di decodifica che gli permetta di trovare la parola $\boldsymbol{z}$ di C a distanza minima da $\boldsymbol{y}$. Ricordiamo che, essendo $d(\boldsymbol{x}, \boldsymbol{y}) \leq e$, risulta $\boldsymbol{z} = \boldsymbol{x}$. Assegnata una tabella standard di C, un possibile schema di decodifica è il seguente: *se i è l'indice della riga della tabella cui $\boldsymbol{y}$ appartiene, si decodifica $\boldsymbol{y}$ come $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{a}_i$*. Poiché il peso di $\boldsymbol{a}_i = \boldsymbol{y} - \boldsymbol{z}$, che è uguale alla distanza fra $\boldsymbol{y}$ e $\boldsymbol{z}$, è per costruzione il più piccolo possibile, al variare di $\boldsymbol{z} \in C$, siamo sicuri di aver usato uno schema di decodifica secondo il principio del *nearest neighbour decoding.* L'algoritmo corrispondente a questo schema è semplice a descriversi e consta dei seguenti passi:

primo passo: scorrere la tabella standard, iniziando dal primo elemento della prima riga e continuando in successione, fino a trovare la parola ricevuta $\boldsymbol{y}$;
secondo passo: decodificare $\boldsymbol{y}$ come la prima parola della colonna della tabella cui $\boldsymbol{y}$ appartiene.

Osserviamo esplicitamente che lo schema di decodifica descritto si fonda sostanzialmente su due fatti:

(1) *l'errore $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{x}$, che il decodificatore non conosce e deve scoprire, e la parola $\boldsymbol{y}$ ricevuta sono nello stesso laterale di C*;

(2) *la speranza che durante la trasmissione non si siano verificati troppi errori; cioè il peso di $\boldsymbol{e}$ sia abbastanza piccolo in modo che $\boldsymbol{e}$ abbia buona probabilità di coincidere con la direttrice del laterale $\boldsymbol{y} + C$.*

Quando il numero delle parole di C è molto grande, il primo passo del nostro algoritmo di decodifica può richiedere molto tempo, così l'intero sistema di comunicazione corre il rischio di essere troppo lento. Se ciò accade, conviene servirsi di sistemi di decodifica più veloci. Uno di questi si basa sulla cosiddetta *decodifica a sindromi,* che funziona nel seguente modo:

(1) *si estende una tabella standard di C aggiungendo la colonna delle sindromi, cioè la colonna il cui elemento generico è la sindrome delle parole del laterale corrispondente alla riga cui l'elemento stesso appartiene* (*cfr.*Prop.4.9);

(2) *si calcola la sindrome $S(\boldsymbol{y})$ di $\boldsymbol{y}$ e, scorrendo la colonna delle sindromi, si trova l'indice i della riga cui $S(\boldsymbol{y})$ e $\boldsymbol{y}$ appartengono;*

(3) *y si decodifica come $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{a}_i$.*

Questo schema di decodifica necessita quindi di una matrice M con due sole colonne, la prima delle quali coincida con la prima colonna di una tabella standard Σ di C, la seconda con la colonna delle sindromi di Σ. Allora anche in questo caso l'algoritmo di decodifica è molto semplice e, detta H una matrice controllo di parità di C, consiste dei seguenti passi:

primo passo: calcolare la sindrome $S(\boldsymbol{y}) = \boldsymbol{y}H^t$ della parola ricevuta $\boldsymbol{y}$;
secondo passo: scorrere la colonna delle sindromi fino a trovare $S(\boldsymbol{y})$;
terzo passo: decodificare $\boldsymbol{y}$ come la differenza $\boldsymbol{z}$ tra $\boldsymbol{y}$ e la parola che si trova a sinistra di $S(\boldsymbol{y})$ nella matrice M.

Si noti che la parola $\boldsymbol{z}$ ottenuta alla fine dell'algoritmo è la stessa che si otterrebbe usando il primo schema di decodifica descritto. Si noti ancora che, al fine della decodifica di $\boldsymbol{y}$, il primo algoritmo deve scorrere una tabella con q^{n-k} righe e q^k colonne, mentre il secondo soltanto la colonna delle sindromi, che ha q^{n-k} elementi. E' chiaro quindi che, se C è abbastanza grande, il secondo algoritmo è molto più veloce del primo.

ESEMPIO 4.10. Consideriamo il $[4,2]-$codice binario

$$C = \{(0,0,0,0),(1,0,1,1),(0,1,0,1),(1,1,1,0)\}$$

avente come matrice controllo

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}.$$

Una tabella standard di C, ampliata mediante la colonna delle sindromi, è data da

$$\begin{array}{cccc|c} (0,0,0,0) & (1,0,1,1) & (0,1,0,1) & (1,1,1,0) & (0,0) \\ (1,0,0,0) & (0,0,1,1) & (1,1,0,1) & (0,1,1,0) & (1,1) \\ (0,1,0,0) & (1,1,1,1) & (0,0,0,1) & (1,0,1,0) & (0,1) \\ (0,0,1,0) & (1,0,0,1) & (0,1,1,1) & (1,1,0,0) & (1,0) \end{array}.$$

□

Gli algoritmi descritti sono applicabili a tutti i codici lineari. C'è da osservare che essi possono essere modificati e resi più efficienti in presenza di particolari classi di tali codici.

4.3 Il problema fondamentale della teoria dei codici lineari

Iniziamo con due risultati, il secondo immediata conseguenza del primo, che sono alla base di quelli che esporremo in questo paragrafo.

PROPOSIZIONE 4.11. *Siano C un $[n, n-m]-$codice su F_q, $H = (a_{ij})$ una sua matrice controllo di parità e denotiamo con $\boldsymbol{a}^1, \boldsymbol{a}^2, ..., \boldsymbol{a}^n$ i vettori colonna della matrice H. Valgono le seguenti proprietà:*
(1) *Se $\boldsymbol{x}$ è una parola di C di peso t e se $x_{i_1}, x_{i_2}, .., x_{i_t}$ sono le sue componenti diverse da zero, allora le colonne $\boldsymbol{a}^{i_1}, \boldsymbol{a}^{i_2}, .., \boldsymbol{a}^{i_t}$ di H sono linearmente dipendenti di F_q^m.*
(2) *Se $x_{i_1}, x_{i_2}, .., x_{i_t}$ sono elementi di F_q non tutti nulli tali che*

$$\sum_{j=1}^{t} x_{i_j} \boldsymbol{a}^{i_j} = \boldsymbol{0},$$

allora il vettore $\boldsymbol{a} \in F_q^n$ di componenti

$$a_s = \begin{cases} 0 \quad \textit{se } s \neq i_1, i_2, .., i_t, \\ \\ x_{i_j} \quad \textit{se } s = i_j \textit{ , con } j = i_1, i_2, .., i_t. \end{cases}$$

è una parola del codice C di peso al più t.

DIMOSTRAZIONE. Nell'ipotesi (1), osservato che

$$\boldsymbol{x}H^t = \left(\sum_{j=1}^{n} x_j a_{1j}, \sum_{j=1}^{n} x_j a_{2j}, .., \sum_{j=1}^{n} x_j a_{mj} \right)$$

$$= \left(\sum_{j=1}^{t} x_{i_j} a_{1i_j}, \sum_{j=1}^{t} x_{i_j} a_{2i_j}, .., \sum_{j=1}^{t} x_{i_j} a_{mi_j} \right) ,$$

l'asserto segue dalle seguenti implicazioni:

$$\boldsymbol{x}H^t = \boldsymbol{0} \Rightarrow \left(\sum_{j=1}^{t} x_{i_j} a_{1i_j}, \sum_{j=1}^{t} x_{i_j} a_{2i_j}, .., \sum_{j=1}^{t} x_{i_j} a_{mi_j} \right) = \boldsymbol{0} \Rightarrow$$

$$x_{i_1}\boldsymbol{a}^{i_1} + x_{i_2}\boldsymbol{a}^{i_2} + \cdots + x_{i_t}\boldsymbol{a}^{i_t} = \mathbf{0}.$$

Con un ragionamento analogo si prova la seconda parte della proposizione. □

PROPOSIZIONE 4.12. *Siano C un $[n, n-m]-$codice su F_q e H una sua matrice controllo di parità. Allora C ha distanza minima d se, e soltanto se, le colonne di H generano F_q^m e verificano le seguenti due proprietà:*

$$\begin{cases} \textit{ogni sottoinsieme di } d-1 \textit{ colonne è indipendente,} \\ \\ \textit{esistono } d \textit{ colonne dipendenti.} \end{cases} \tag{13}$$

Inoltre, per ogni fissato $[n, n-m, d]-$codice su F_q, risulta

$$d \leq m+1. \tag{14}$$

La Prop.4.12 evidenzia l'importanza degli insiemi Γ di n vettori di F_q^m con le seguenti proprietà:
• Γ è un generatore di F_q^m,
• tutti gli insiemi di $d-1$ vettori di Γ sono indipendenti,
• in Γ esistono d vettori dipendenti.
Gli insiemi di questo tipo, detti $(n, d-1)-$*insiemi,* sono oggetti di studio della geometria combinatoria fin dagli anni '50 del secolo scorso, indipendentemente dalla teoria dei codici, e sono alla base della *teoria delle calotte* negli spazi proiettivi su campi di Galois. Con questa terminologia, La Prop.4.12 può essere riformulata nel seguente modo.

PROPOSIZIONE 4.13. *Sia C un $[n, n-m]-$codice su F_q e H una sua matrice controllo di parità. Allora C ha distanza minima $d > 2$ se, e soltanto se, le colonne di H formano un $(n, d-1)-$insieme di F_q^m.*

La ricerca del massimo numero $max_{d-1}(m, q)$ di vettori di un $(n, d-1)-$insieme di F_q^m è noto come *packing problem* ed è uno dei più difficili e studiati problemi della geometria su campi di Galois. Nella tabella che segue sono riportati i valori noti di $max_{d-1}(m, q)$.

m	q	$d-1$	$max_{d-1}(m,q)$	
		2	$\frac{q^m-1}{q-1}$	
	2	3	2^{m-1}	*Bose* 1947
3	*pari*	3	$q+2$	*Bose* 1947
3	*dispari*	3	$q+1$	*Bose* 1947
4	*dispari*	3	q^2+1	*Bose* 1947
4	*pari* > 2	3	q^2+1	*Qvist* 1952
5	3	3	20	*Pellegrino* 1970
6	3	3	56	*Hill* 1973

Fissati m, d, q, con $d \leq m+1$, il problema di calcolare il più grande intero n per cui esiste un codice su F_q con parametri $[n, n-m, d]$ è noto come *problema fondamentale della teoria dei codici lineari.* Tale problema, nel caso $d \geq 2$, è equivalente al *packing problem* in forza della Prop.4.13.

PROPOSIZIONE 4.14. *Fissati m, d, q, con $d \leq m+1$, il massimo intero n per cui esiste un codice su F_q con parametri $[n, n-m, d]$ è uguale al massimo numero di vettori di un $(n, d-1)-$insieme di F_q^m, cioè a $max_{d-1}(m, q)$.*

ESEMPIO 4.15. Le colonne della matrice

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & -1 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & 1 & 0 \end{bmatrix}$$

sono a quattro a quattro indipendenti sul campo $F_3 = \{0, 1, -1\}$ dei resti modulo 3. Allora il codice lineare su F_3 avente H come matrice di controllo, in forza della Prop.4.12, ha parametri $[11, 6, 5]$ ed è perfetto:

$$3^6 \left[\binom{11}{0} + \binom{11}{1} 2 + \binom{11}{2} 2^2 \right] = 3^6 \left[1 + 22 + 220\right] = 3^6 3^5 = 3^{11}.$$

Tale codice è noto come *codice ternario di Golay* e si denota con $\mathcal{G}_{11}$. □

I codici lineari per i quali la (14) è un'uguaglianza, cioè i codici con parametri $[n, n-m, m+1]$ si chiamano *codici MDS* (MDS sta per *maximum distance separable*).

ESEMPIO 4.16. Siano $a_1, a_2, \ldots, a_{q-1}$ gli elementi non nulli del campo F_q e in F_q^m, con $2 \leq m \leq q$, consideriamo l'insieme di $q+1$ vettori (*curva razionale normale*)

$$X = \left\{(1, t, t^2, ..., t^{m-1}) \ : \ t \in F_q\right\} \cup \left\{(0, \ldots, 0, 1)\right\}.$$

La matrice avente per colonne i vettori di X

$$H = \begin{bmatrix} 1 & 1 & \ldots & 1 & 1 & 0 \\ a_1 & a_2 & \ldots & a_{q-1} & 0 & 0 \\ a_1^2 & a_2^2 & \ldots & a_{q-1}^2 & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ a_1^{m-1} & a_2^{m-1} & \ldots & a_{q-1}^{m-1} & 0 & 1 \end{bmatrix}$$

ha le colonne a m a m indipendenti ed ha rango m. Allora il codice lineare su F_q avente H come matrice di controllo ha parametri

$$[q+1, q+1-m, m+1]$$

ed è un codice MDS. □

4.4 I codici di Hamming

L'intero $max_2(m,q)$ è il massimo numero di vettori non nulli di F_q^m a due a due indipendenti (non proporzionali), cioè non appartenenti ad uno stesso sottospazio vettoriale $1-$dimensionale. Ne segue che un $(n,2)-$insieme d'ordine $max_2(m,q)$ si ottiene prendendo un vettore non nullo in ciascuno dei sottospazi $1-$dimensionali di $V(n,q)$ e, quindi, $max_2(m,q)$ è uguale al numero dei sottospazi $1-$dimensionali di F_q^m.

Consideriamo la famiglia

$$\{V_i(1,q) \ : \ i = 1,2,...,n = q^{m-1} + q^{m-2} + \cdots + q + 1\}$$

di tutti i sottospazi $1-$dimensionali di F_q^m, $m > 2$, e in ognuno di essi scegliamo un vettore non nullo $\boldsymbol{a}^i = (a_{1i}, a_{2i}, ..., a_{mi})$. Denotata con

$$H = H_{m,q} = (a_{ij})$$

la matrice di tipo $m \times n$ avente come vettori colonna $\boldsymbol{a}^1, \boldsymbol{a}^2, ..., \boldsymbol{a}^n$, possiamo considerare l' $[n, n-m]-$codice lineare avente H come matrice controllo di parità. Tale codice, che denotiamo con $Ham(m,q)$, si chiama $(m,q)-$*codice di Hamming* ed è definito da

$$Ham(m,q) = \{\boldsymbol{a} \in F_q^n \ : \ \boldsymbol{a}H^t = \boldsymbol{0}\}.$$

Poiché H ha le colonne a due a due indipendenti e ne contiene tre dipendenti, abbiamo, in forza della Prop.4.12, che la distanza minima di $Ham(m,q)$ è 3 e, di conseguenza, esso è un $[n, n-m, 3]-$codice $1-$correttore. Ne segue che le sfere di centro le parole di $Ham(m,q)$ e raggio 1 sono a due a due disgiunte e, poiché ognuna di esse contiene esattamente $n(q-1)+1$ parole di F_q^n, risulta

$$q^{n-m}[1 + n(q-1)] = q^n = |F_q^n|,$$

il che assicura che $Ham(m,q)$ è un codice perfetto. Osserviamo che (*cfr.* Prop.4.11) una parola non nulla $\boldsymbol{a}$ di $H(m,q)$ ha peso c e presenta lettere diverse da zero nelle posizioni $i_1, i_2, ..., i_c$ se, e soltanto se, le colonne di posto $i_1, i_2, ..., i_c$ in H

sono linearmente dipendenti. Osserviamo, ancora, che $Ham(3,2)$ è equivalente al codice associato al piano di Fano. □

Nel caso $q = 2$, possiamo interpretare le colonne della matrice $H_{m,2}$ come la rappresentazione binaria degli interi da 1 a $2^m - 1$ e possiamo disporle in ordine crescente dell'intero che rappresentano. Sotto queste ipotesi, in un canale di trasmissione binario, che commette non più di un errore sulle parole di lunghezza $n = 2^m - 1$ e che lavora con il codice Ham(m, 2), si può usare il seguente algoritmo di decodifica, ove $\boldsymbol{y}$ è la parola ricevuta e $\boldsymbol{z}$ la decodifica di $\boldsymbol{y}$:

primo passo: Calcolare la sindrome $S(\boldsymbol{y})$.

secondo passo: Se $S(\boldsymbol{y}) = 0$, si ponga $\boldsymbol{z} = \boldsymbol{y}$.

terzo passo: Se $S(\boldsymbol{y}) \neq \mathbf{0}$, $\boldsymbol{z}$ si ponga uguale alla parola che si ottiene da $\boldsymbol{y}$ modificando la sua $j-$esima componente, j essendo il numero intero rappresentato in binario da $S(\boldsymbol{y})$.

Quello appena descritto è una variante dell'algoritmo di decodifica a sindromi ma, a differenza di quest'ultimo, è veloce perché, per trovare la riga della tabella standard di C cui $\boldsymbol{y}$ appartiene non è necessario scorrere la colonna delle sindromi: basta solo trovare la sindrome di $\boldsymbol{y}$ (un prodotto di matrici). Questo è uno dei motivi per cui i codici binari di Hamming sono molto utilizzati.

Esempio 4.17. Nel caso $m = 3$, per la decodifica veloce con $Ham(3,2)$, bisogna considerare la matrice di controllo

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

Se $\boldsymbol{y} = 1101011$ è la parola ricevuta, si calcola $S(\boldsymbol{y}) = 110$, che risulta la rappresentazione binaria di 6. Se ne deduce, allora, che è stato commesso un errore nella sesta posizione. La parola $\boldsymbol{y}$ viene così decodificata con $\boldsymbol{z} = 1101001$. □

4.5 L'enumeratore dei pesi

Uno dei problemi centrali nello studio di un codice lineare C di lunghezza n è il calcolo, per ogni $i = 1, 2, ..., n$, del numero $w_i = w_i(C)$ di tutte le parole di C di peso i e, a tale proposito, spesso si considera il polinomio

$$W(x,y) = W_C(x,y) = \sum_{\boldsymbol{a} \in C} x^{w(\boldsymbol{a})} y^{n - w(\boldsymbol{a})} = \sum_{i=0}^{n} w_i x^i y^{n-i}, \qquad (15)$$

che prende il nome di *enumeratore dei pesi* di C. Il teorema successivo fornisce una importante relazione fra $W(x,y)$ e l'enumeratore dei pesi $W^{\perp}(x,y)$ del codice duale di C.

PROPOSIZIONE 4.18. (F.J.MacWilliams, 1963). *Consideriamo $W(x,y)$ e $W^{\perp}(x,y)$, gli enumeratori dei pesi rispettivamente di un $[n,k]-$codice C su F_q e del suo duale. Allora risulta*

$$W^{\perp}(x,y) = q^{-k}W(y-x, y+(q-1)x)$$

e quindi, se C è autoduale

$$W(x,y) = q^{-n/2}W(y-x, y+(q-1)x).$$

In alcuni casi la conoscenza dei pesi delle parole di una matrice generatrice A di un codice lineare dà informazioni sui coefficienti del polinomio enumeratore dei pesi. Per esempio, se due parole $\boldsymbol{a}, \boldsymbol{b}$ di un codice binario C hanno peso pari $2s$ e $2t$, rispettivamente, detto m il numero degli indici j tali che $a_j = b_j = 1$, risulta

$$w(\boldsymbol{a}+\boldsymbol{b}) = 2s - m + 2t - m = 2(s+t-m).$$

Ne segue che, se le parole di A hanno tutte peso pari, allora C è un codice *pari,* cioè tutte le sue parole hanno peso pari. Analogamente si prova che, se C è autoortogonale e tutte le parole di A hanno peso divisibile per 4, la stessa proprietà è vera per tutte le parole di C. In questo caso C si dice *doppiamente pari.*

5 Codici lineari e disegni

5.1 Generalità

Sia $q = p^h$ una potenza di un numero primo p, $F = F_q$ il campo di Galois con q elementi e $\mathcal{P} = \{p_1, p_2, ..., p_v\}$ un insieme finito con v elementi.

Per ogni parola $\boldsymbol{a} = (a_1, a_2, ..., a_v)$ di lunghezza v su F, definiamo *supporto di* $\boldsymbol{a}$ l'insieme dei punti p_i di $\mathcal{P}$ tali che $a_i \neq 0$. Il supporto di una parola $\boldsymbol{a}$, che denoteremo con $supp(\boldsymbol{a})$, è quindi un sottoinsieme di $\mathcal{P}$ e chiaramente risulta

$$supp(\boldsymbol{a}) = \emptyset \quad \Leftrightarrow \quad \boldsymbol{a} = \boldsymbol{0},$$

$$|supp(\boldsymbol{a})| = w(\boldsymbol{a}),$$

$$supp(\boldsymbol{a}) = \mathcal{P} \quad \Leftrightarrow \quad w(\boldsymbol{a}) = v.$$

Osserviamo che, se denotiamo con $\star$ l'operazione di differenza simmetrica, risulta

$$q = 2 \Rightarrow \begin{cases} supp(\boldsymbol{a} + \boldsymbol{b}) = supp(\boldsymbol{a}) \star supp(\boldsymbol{b}) \\ d(\boldsymbol{a}, \boldsymbol{b}) = |supp(\boldsymbol{a}) \star supp(\boldsymbol{b})| \\ \boldsymbol{ab} = 0 \quad \Leftrightarrow \quad |supp(\boldsymbol{a}) \cap supp(\boldsymbol{b})| \text{ è pari} \end{cases} \tag{16}$$

per ogni due parole $\boldsymbol{a}$ e $\boldsymbol{b}$. Inoltre, sempre nell'ipotesi $q = 2$, ogni parola $\boldsymbol{a}$ può riguardarsi come la funzione caratteristica di $supp(\boldsymbol{a})$.

Il teorema che segue fornisce un primo esempio dello stretto legame esistente tra la teoria dei codici e quella dei disegni. Esso permette, infatti, di costruire un disegno a partire da un codice perfetto binario.

PROPOSIZIONE 5.1. *Sia C un $(v, M, d)-$codice binario $e-$correttore. Allora, se C è perfetto e non banale, i supporti delle sue parole di peso $d = 2e + 1$ formano i blocchi di un $(e + 1) - (v, d, 1)$ disegno.*

ESEMPIO 5.2. Ogni codice lineare binario C con gli stessi parametri di un codice di Hamming binario, cioè un $[2^m - 1, 2^m - 1 - m, 3]-$codice su F_2, è perfetto. Allora, usando la Prop.5.1, è possibile provare che il disegno associato a C è quello dei punti e delle rette di $PG(m - 1, 2)$, lo spazio proiettivo di dimensione $m - 1$ sul campo con 2 elementi. □

Per completezza espositiva riportiamo l'enunciato di uno dei più importanti teoremi che permettono di costruire $t-$disegni a partire da codici lineari.

PROPOSIZIONE 5.3. (E.F.Assmus - H.F.Mattson, 1969). *Sia C un $[n, k, d]-$codice su F_q e denotiamo con w_j il numero delle parole di peso j di C e con $d^\perp$ la distanza minima di $C^\perp$. Siano inoltre:*

- *t un intero minore di $d + 1$;*
- *v il più grande intero tale che*

$$v - \frac{v + q - 2}{q - 1} < d, \quad se \quad q \neq 2, \quad e \quad v = n, \quad se \quad q = 2.$$

Supponiamo, infine, che il polinomio

$$W^\perp(x, y) = \sum_{i=0}^{n} w_i^\perp x^i y^{n-i},$$

enumeratore dei pesi di $C^\perp$, abbia al più $d - t$ coefficienti diversi da zero tra $w_1^\perp, w_2^\perp, .., w_{n-t}^\perp$.

Allora, per ogni intero j tale che

$$w_j \neq 0 \ \ e \ \ d \leq j \leq v,$$

i supporti delle parole di C di peso j costituiscono la famiglia dei blocchi di un $t-$disegno. Analogamente, per ogni intero s tale che

$$w_s^{\perp} \neq 0 \ \ e \ \ d^{\perp} \leq s \leq min\{n-t, v^{\perp}\},$$

i supporti delle parole di $C^{\perp}$ di peso s costituiscono la famiglia dei blocchi di un $t-$disegno.

Il teorema precedente ha permesso di costruire diversi nuovi $5-$disegni. Stranamente, esso non ha dato luogo a nuovi $t-$disegni con $t > 5$.

5.2 Codice lineare associato ad un disegno

Vediamo ora come è possibile costruire dei codici a partire da un $t-$disegno. A tale scopo, supponiamo che $\mathbf{D} = (\mathcal{P}, \mathcal{B})$ sia un $t-(v,k,\lambda)$ disegno e, posto $\mathcal{B} = \{B_1, B_2, \ldots, B_b\}$, denotiamo con $A = (a_{ij})$ una matrice di tipo $b \times v$ che sia la matrice d'incidenza di $\mathbf{D}$. Le colonne e le righe di A sono dunque in corrispondenza biunivoca con i punti ed i blocchi di $\mathbf{D}$, rispettivamente. Il codice lineare $C_q(\mathbf{D})$ su F_q generato dalle righe di A prende il nome di *codice di* $\mathbf{D}$ su F_q ed è indipendente dalla matrice d'incidenza considerata.

Poiché A è ad elementi 0 e 1, la dimensione $dim_q(\mathbf{D})$ di $C_q(\mathbf{D})$ corrisponde al rango $rank_p(A)$ di A sul sottocampo fondamentale F_p di F_q e, per questo motivo, prende il nome di $p-$*rango* di A o del disegno $\mathbf{D}$.

Il codice intersezione di $C_q(\mathbf{D})$ e del suo duale $C_q(\mathbf{D})^{\perp}$ si chiama *q-involucro* di $\mathbf{D}$ e si denota con $Hull_q(\mathbf{D})$. In questo modo al $t-$disegno $\mathbf{D}$ rimangono associati i tre codici $C_q(\mathbf{D})$, $C_q(\mathbf{D})^{\perp}$ e $Hull_q(\mathbf{D})$, la cui conoscenza spesso permette di descrivere agevolmente molte proprietà di $\mathbf{D}$. Per esempio, è di grande utilità conoscere la distribuzione dei pesi di tali codici, cioè i coefficienti del polinomio (15). Viceversa, a volte è possibile trovare dei buoni codici partendo dalla conoscenza di particolari disegni.

Osserviamo esplicitamente che le righe della matrice A sono parole del codice $C_q(\mathbf{D})$ i cui supporti sono esattamente i blocchi di $\mathbf{D}$. Nel seguito denoteremo sempre con $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_b$ le righe di A e così, con le nostre notazioni, abbiamo

$$supp(\boldsymbol{a}_i) = B_i \ , \ \ \text{per ogni } i = 1, 2, \ldots, b.$$

I sottoinsiemi di $\mathcal{P}$ che risultano supporti di parole di $C_q(\mathbf{D})$, con abuso di linguaggio, saranno chiamati $q-$*parole* o semplicemente $parole$ nel caso $q = 2$. È

chiaro che, se X è una $q-$parola non vuota, esistono diverse parole di $C_q(\mathbf{D})$ aventi come supporto X. Nel caso $q = 2$ la corrispondenza tra parole di $C_2(\mathbf{D})$ e le parole di $\mathcal{P}$ è biunivoca. Pertanto, quando è $q = 2$, le parole di $\mathcal{P}$ sono tutti e soli gli insiemi di punti che risultano differenza simmetrica di blocchi (*cfr.*(16)). In altri termini questo significa che X è una parola di $\mathcal{P}$ se, e soltanto se, esiste un insieme $\mathcal{F}$ di blocchi con le seguenti proprietà:

(1) *ogni punto di X appartiene ad un numero dispari di blocchi di $\mathcal{F}$*,

(2) *ogni punto di $\mathcal{P} \setminus X$ appartiene ad un numero pari di blocchi di $\mathcal{F}$*.

Esempio 5.4. Il determinante, calcolato sui razionali, di una matrice d'incidenza del piano di Fano $PG(2,2)$ è uguale a 24(=$2^3 3$, *cfr.*(2)) e quindi risulta

$$dim_q(PG(2,2)) < 7 \ , \quad \text{per } p = 2, 3,$$

$$dim_q(PG(2,2)) = 7 \ , \quad \text{per ogni } p \neq 2, 3.$$

Il codice C di $PG(2,2)$ su F_2 è quello descritto nell'Esempio 3.15 e abbiamo già visto che esso è equivalente al codice di Hamming $Ham(3,2)$. C'è da osservare che i supporti delle parole di C di peso 3 sono tutte e sole le rette. Analogamente, i supporti delle parole di C di peso 4 sono tutti e soli i complementari delle rette che, in questo caso, coincidono con le iperovali (insiemi di 4 punti a 3 a 3 non allineati). Queste due famiglie individuano dunque rispettivamente $PG(2,2)$ e il disegno avente per blocchi i complementari delle rette di $PG(2,2)$.
Non è difficile verificare che il codice di $PG(2,2)$ su F_3 è il codice duale di $\{\mathbf{0}, \boldsymbol{j}, -\boldsymbol{j}\}$, ove $\boldsymbol{j} = (1,1,1,1,1,1,1,1)$. Quest'ultima proprietà è caso particolare di un teorema sui codici dei disegni simmetrici. Il codice di $PG(2,2)$ su F_q, con $q \neq 2, 3$, è F_q^7. □

Nello studio di un $t-$disegno $\mathbf{D}$, oltre alla conoscenza del codice di $\mathbf{D}$ su F_q è spesso determinante la conoscenza del suo involucro. Questo problema non si pone quando $C_q(\mathbf{D})$ è autoortogonale perché, in questo caso, risulta evidentemente $Hull_q(\mathbf{D}) = C_q(\mathbf{D})$. Per esempio, usando la (16) si prova immediatamente la seguente proposizione.

Proposizione 5.5. *Il codice binario associato ad un $t-(v,k,\lambda)$ disegno è autoortogonale se, e soltanto se, sono verificate le seguenti proprietà:*
(i) *k è pari;*
(ii) *ogni due blocchi distinti di $\mathbf{D}$ s'intersecano in un numero pari di punti.*

Esempio 5.6. Il codice binario associato al disegno di Mathieu $\mathcal{M}_{24}$ (*cfr.* Esempio 2.6) è di lunghezza 24, autoortogonale e di dimensione 12. Questo codice si denota con $\mathcal{G}_{24}$ ed è noto come *codice binario di Golay esteso.* A titolo di

informazione, ricordiamo che $\mathcal{G}_{24}$ è uno dei codici utilizzati dalla $NASA$ agli inizi degli anni $'80$ nel programma aerospaziale *Voyager* ed è stato usato per trasmettere sulla terra immagini di Giove e di Saturno. □

PROPOSIZIONE 5.7. *Sia* $\mathbf{D} = (\mathcal{P}, \mathcal{B})$ *un* $2-(v,k,\lambda)$ *disegno simmetrico e* q *la potenza di un primo* p *che non divide l'ordine* $n = k - \lambda$ *di* $\mathbf{D}$. *Allora*

$$rank_p(\mathbf{D}) \geq v - 1, \tag{17}$$

l'uguaglianza avendosi se, e solo se, p *divide* k. *In quest'ultimo caso* $C_q(\mathbf{D})$ *coincide con* $C(q,v)^{\perp}$, *ove*

$$C(q,v) = \{(a,a,\ldots,a) \ : \ a \in F_q\}$$

è il codice di ripetizione $q-$*ario di lunghezza* v *su* F_q.

5.3 Codice binario associato ad un piano proiettivo e non esistenza del piano d'ordine 10

Sia $\pi_n = (\mathcal{P}, \mathcal{B})$ un piano proiettivo d'ordine n e ricordiamo che una tale struttura è null'altro che un $2-(v, n+1, 1)$ disegno simmetrico con $v = n^2 + n + 1$. Denotate con $L_1, L_2, \ldots, L_v$ le rette di π_n, sia $\boldsymbol{a}_i$ la riga corrispondente alla retta L_i nella matrice A trasposta di incidenza di π_n.

Applicando a π_n la Prop.5.7, abbiamo

$$p \not| n(n+1) \Rightarrow rank_p(\pi_n) = dim_q(\pi_n) = n^2 + n + 1$$

e

$$p|(n+1) \Rightarrow \begin{cases} rank_p(\pi_n) = dim_q(\pi_n) = n^2 + n \\ \\ C_q(\pi_n) = C(q, n^2+n+1)^{\perp} \end{cases} . \tag{18}$$

Osserviamo esplicitamente che dalla (18) segue

$$n \text{ dispari } \Rightarrow \ rank_2(\pi_n) = dim_2(\pi_n) = n^2 + n \tag{19}$$

e $C_2(\pi_n)$ è il codice binario formato da tutte le parole di F_2^v aventi peso pari.

DEFINIZIONE 5.8. Un insieme $\mathcal{F}$ di rette di π_n si dice *pari* (risp. *dispari*) se ogni punto del piano appartiene ad un numero pari (risp. dispari) di rette di $\mathcal{F}$. Dualmente, un insieme X di punti di π_n si dice *pari* (risp. *dispari*) se ogni retta del piano interseca X in un numero pari (risp. dispari) di punti. □

Osservazione 5.9. Nel caso n dispari, come conseguenza della (19) e della successiva osservazione, abbiamo che un insieme pari e non vuoto di rette (risp. punti) di π_n deve necessariamente coincidere con $\mathcal{B}$ (risp. $\mathcal{P}$). □

Mettiamoci ora nell'ipotesi che n sia pari e, posto $C = C_2(\pi_n)$, consideriamo il codice esteso $\overline{C}$ di C. Poiché $n+1$ è dispari, la parola $\boldsymbol{a}'_i$ di $\overline{C}$ corrispondente ad $\boldsymbol{a}_i$ è data da $\boldsymbol{a}'_i = (\boldsymbol{a}_i, 1)$ e, poiché due rette distinte s'intersecano in esattamente un punto, abbiamo

$$\boldsymbol{a}'_i \boldsymbol{a}'_j = 0 \ , \quad \text{per ogni } i, j = 1, 2, \ldots, v.$$

Ne segue che $\overline{C}$ è autoortogonale e, in forza della (12), abbiamo

$$n^2 + n + 2 \geq 2dim(\overline{C}) = 2dim(C),$$

cioè

$$n \ \text{ pari } \Rightarrow \ rank_2(\pi_n) = dim_2(\pi_n) \leq \frac{n^2+n+2}{2}.$$

La Prop.5.7 e le considerazioni svolte nel precedente paragrafo mostrano che è interessante studiare il codice su F_q di un disegno simmetrico solo nel caso che questo abbia ordine n divisibile per p. Questo studio, che in generale presenta non poche difficoltà, è particolarmente interessante nel caso dei piani proiettivi d'ordine $n \equiv 2(mod\,4)$ e $q = 2$. I due risultati seguenti sono utili per comprendere il supporto teorico alla base della ricerca esaustiva che ha portato alla dimostrazione della non esistenza di un piano proiettivo d'ordine 10.

Proposizione 5.10. *Sia π_n un piano proiettivo finito d'ordine n, con $n \equiv 2(mod\,4)$. Allora risulta*

$$dim_2(\pi_n) = \frac{n^2+n+2}{2}.$$

Ne segue che il codice esteso di $C_2(\pi_n)$ è autoduale.

Proposizione 5.11. *Sia π_n un piano proiettivo finito d'ordine n, con $n \equiv 2(mod\,4)$. Allora la distanza minima di $C = C_2(\pi_n)$ è*

$$d = n + 1,$$

le parole di π_n di cardinalità $n+1$ essendo tutte e sole le rette di π_n. Inoltre, le parole di π_n di cardinalità $n+2$ sono tutte e sole le iperovali di π_n, cioè gli insiemi di $n+2$ punti a tre a tre non allineati.

Riportiamo ora la *storia* della dimostrazione della non esistenza di un piano proiettivo d'ordine 10. Tale dimostrazione è stata ottenuta sfruttando essenzialmente la potenza raggiunta negli anni $'80$ del secolo scorso dagli elaboratori elettronici ed è tuttora aperta la questione della ricerca di metodi adeguati per la soluzione del problema senza l'uso di strumenti di calcolo.

Denotiamo con C il codice su F_2 di un ipotetico piano proiettivo π d'ordine 10 e con w_i i coefficienti del relativo polinomio enumeratore dei pesi. Le parole $\boldsymbol{a}'_i$ del codice esteso $\overline{C}$ hanno peso $(= 12)$ divisibile per 4 e ciò, essendo $\overline{C}$ autoduale, implica che $\overline{C}$ è doppiamente pari.

Mediante l'uso di elaboratori elettronici sono stati provati i seguenti risultati:

$w_{12} = 0$ (C.W.H.Lam, L.Thiel, S.Swiercz, J.McKay, 1983),

$w_{15} = 0$ (R.H.F.Denniston, 1969 e indipendentemente F.J.MacWilliams, N.J.A.Sloane, J.C.Thompson, 1973),

$w_{16} = 0$ (C.W.H.Lam, S.Swiercz, L.Thiel, 1986).

Sono dunque noti i valori di w_i, per ogni $i = 0, 1, 2, .., 18$, che riportiamo nella seguente tabella.

w_0	$w_i \;\; i = 1, 2, .., 10$	w_{11}	$w_j \;\; j = 12, 13, .., 18$
1	0	111	0

Nel 1970, *E.F.Assmus* e *H.F.Mattson* hanno osservato che, usando i valori riportati nella tabella precedente e la relazione (4.18), è possibile trovare tutti i coefficienti del polinomio enumeratore dei pesi di C. In particolare, si ha

$$w_{19} = 24.675$$

e quindi, se esiste un piano proiettivo l'ordine 10, esso deve contenere delle configurazioni di 19 punti che risultino differenza simmetrica di rette. Queste configurazioni sono state studiate da *M.Hall Jr,* il quale nel 1980 ha provato che, se $\boldsymbol{a}$ è una parola di C di peso 19, allora, delle 111 rette di π, 6 intersecano $w(\boldsymbol{a})$ in 5 punti, 37 in 3 punti e 68 in un solo punto. In sostanza, le configurazioni in questione possono pensarsi come strutture geometriche aventi 19 *punti* e 43 *blocchi* tali che: 6 *blocchi hanno* 5 *punti,* 37 *blocchi hanno* 3 *punti e due punti distinti appartengono ad un unico blocco.* Inoltre, si può provare che, a meno di isomorfismi, il numero di tali strutture è 66 e ovviamente una almeno di queste, se esiste π, deve avere matrice di incidenza estendibile ad una matrice di incidenza di π. E' appunto questo il dato di partenza delle ricerche dei matematici che si sono occupati del problema del piano proiettivo d'ordine 10.

In un lavoro del 1985 *Lam, Crossfiefl e Thiel* hanno provato che 21 configurazioni delle 66 possibili hanno matrice di incidenza non estendibile a quella di un piano

proiettivo d'ordine 10. Finalmente, nel 1988, *Lam, Thiel e Swiercz* [12], usando dei programmi da loro elaborati (parte di questi hanno girato per 83 giorni su un *CRAY supercomputer* e per 160 giorni su cinque *VAX* collegati in rete) hanno esteso il precedente risultato alle rimanenti 45 configurazioni, giungendo cosí a provare che *non esiste un piano proiettivo d'ordine* 10.

Considerazioni finali e ringraziamenti

La matematica è una scienza fondamentale che, per i matematici, non ha bisogno di essere motivata da eventuali sue applicazioni. La maggior parte delle persone, invece, la apprezzano, la accettano e, a volte la sopportano, solo per l'evidente importanza delle sue applicazioni. È questo il motivo principale per il quale abbiamo scelto come argomento della presente relazione la *teoria dei codici lineari* e le sue applicazioni alla trasmissione dell'informazione. Questo capitolo della matematica, almeno nei suoi fondamenti, si presta ad essere presentato anche a studenti delle scuole superiori e ha il pregio di mostrare modelli matematici che, a differenza di quelli classici e giustamente più studiati, utilizzano l'algebra e la geometria su campi finiti piuttosto che l'analisi matematica e la geometria del continuo. Abbiamo parlato più volte su questi argomenti nell'ambito del progetto "*Lauree scientifiche*", limitandoci al caso binario, e abbiamo verificato personalmente che essi stimolano curiosità e interesse negli studenti.

Per finire, desideriamo ringraziare di cuore il Presidente *Emilio Ambrisi* e il *Comitato Scientifico* del Congresso Nazionale della Mathesis (Caserta, 27-29 ottobre 2011) per l'onore che ci hanno concesso nell'invitarci a tenere questa relazione.

Riferimenti bibliografici

[1] Assmus E.F., Key J.D., *Designs and their Codes, Cambridge Tracts in Mathematics,* 103, Cambridge University Press, 1992.

[2] Berardi L., *Algebra e teoria dei codici correttori,* Collana di matematica e statistica, Franco Angeli, 1994.

[3] Bruck R.H. , Ryser H.J., *The nonexistence of certain finite projective planes,* Canadian Journal of Mathematics, 1, 317-320, 1949.

[4] Cameron P.J., van Lint J.H., *Designs, Graphs, Codes and their Links,* London Mathematical Society, Student Texts 22, 1991.

[5] Dembowski P., *Finite geometries,* Springer-Verlag, 1968.

[6] Giuzzi L., *Codici correttori,* Collana UNITEXT, Springer, 2006

[7] Hamming R.W., Self-Correcting Codes, Case 20878, Memorandum 1130-RWH-MFW, *Bell Telephone*, 1947.

[8] Hamming R.W., *Error detecting and error correcting codes,* Bell Syst. Tech. J., 29, 147-160, 1950.

[9] Hill R., *A First Course in Coding Theory, Oxford Applied Mathematics and Computing Science Series,* Clarendon Press - Oxford, 1990.

[10] Hughes D.R., Piper F.C., *Projective planes,* Springer Verlag, 1973.

[11] Hughes D.R., Piper F.C., *Design Theory,* Cambridge University Press, 1988.

[12] Lam C.W.H., Thiel L., Swiercz S., *The non-existence of finite projective planes of order* 10, Canad. J. Math., 41, 1117-1123, 1989.

[13] Magliveras S.S., Leavitt D.W., *Simple six designs exist,* Proceedings of the 14th Southheastern Conference on Combinatorics, Graph Theory, Computing, Congr. Num. 40, Utilitas Math., Winnipeg, 195-205, 1983.

[14] Mazzocca F., *Lucidi del corso di codici lineari,* Dipartimento di Matematica della Seconda Università degli Studi di Napoli, 2011, (www.francesco.mazzocca.name) .

[15] Shannon C.E., A Mathematical Theory of Communication, *The Bell System Thecnical Journal,* vol.27, 379-656, July, October, 1948.

[16] Teirlinck L., *Nontrivial t-designs without repetead blocks exist for all t,* Discrete Mathematics, 65, 301-311, 1987.

[17] Teirlinck L., *Locally trivial t-designs and t-designs without repeated blocks,* Discrete Mathematics, 77, 345-356, 1989.

[18] Tonchev V.D., *Combinatorial Configurations,* Longman Scientific & Technical, 1988.