Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

Corso di Laurea Magistrale in Matematica · Curriculum applicativo

CAMPI FINITI E SEGNALE GPS

Relatore:
Chiar.mo Prof.
Rüdiger Achilles

Correlatore:
Chiar.ma Prof.
Mirella Manaresi

Presentata da:
Virginia Brecciaroli

Seconda Sessione
Anno Accademico 2012/2013

Se non puoi essere un'autostrada, sii un sentiero,
se non puoi essere il sole, sii una stella.
Sii sempre il meglio di ciò che sei.
Cerca di scoprire il disegno che sei chiamato ad essere,
poi mettiti con passione a realizzarlo nella vita.
Martin Luther King

Introduzione

La teoria dei campi finiti viene usata sia in matematica pura che in quella applicata, in particolare è uno strumento importante in: teoria dei numeri, geometria algebrica, geometria aritmetica, teoria di Galois, crittografia, teoria dei codici e correzione quantistica degli errori.
Lo scopo di questa tesi è quello di studiare in maniera approfondita una delle tante affascinanti applicazioni di questa teoria: il segnale GPS. A questo scopo, si è scelto di studiare i registri a scorrimento a retroazione lineare (linear feedback shift registers, LFSR). Questi dispositivi possono essere implementati in hardware e questo li rende utili in applicazioni che richiedono la generazione molto rapida di numeri pseudo-casuali, come nella tecnica radio Direct Sequence Spread Spectrum usata ad esempio nell'UMTS. In particolare, si è scelto di studiare le vantaggiose proprietà degli LFSR utilizzati nei sistemi GPS per trasmettere rapidamente una sequenza, che indica degli istanti relativi ad alta precisione, sfruttandone il determinismo.
La tesi è stata quindi suddivisa in tre parti:

1. Breve introduzione al funzionamento del GPS

2. Campi finiti

3. Registri a scorrimento a retroazione lineare.

1. Nel primo capitolo si spiega in modo sintetico come funziona il Sistema di Posizionamento Globale, in breve GPS. Si parte da cenni di storia, dalla presentazione di alternative al sistema statunitense e dall'illustrazione del sistema che, attraverso una rete satellitare in orbita, fornisce

ad un ricevitore GPS informazioni sulle sue coordinate geografiche ed orario. In seguito, si descrive come avviene la localizzazione e cioè, intersecando con la superficie terrestre tre sfere, i cui raggi sono la distanza satellite-ricevitore. Questa distanza viene calcolata tramite la trasmissione di un segnale radio da parte di ciascun satellite e l'elaborazione dei segnali ricevuti da parte del ricevitore, che è in grado di misurare il tempo impiegato dal segnale ad arrivare. Questo è possibile grazie ad un registro a scorrimento che genera una sequenza di bit (di zeri e di uno), periodica, poco correlata con traslazioni di se stessa, permettendo così al ricevitore di identificare il satellite da cui proviene il segnale e di sincronizzarsi con esso. Il grado attuale di accuratezza è dell'ordine dei metri, in dipendenza dalle condizioni meteorologiche, dalla disponibilità e dalla posizione dei satelliti rispetto al ricevitore, dalla qualità e dal tipo di ricevitore, dagli effetti di radiopropagazione del segnale radio in ionosfera e troposfera (come la riflessione) e dagli effetti della relatività. Infine, vengono descritti il messaggio di navigazione, le informazioni in esso contenute e come viene generato.

2. Il secondo capitolo è dedicato alla teoria dei campi finiti, che fornisce tutti gli strumenti necessari allo studio degli LFSR trattati nel terzo ed ultimo capitolo. Nel paragrafo 2.1 si studiano i campi primi con p elementi, dove p è un primo. Questi campi sono denotati con $\mathbb{F}_p$, non contengono sottocampi propri e ne esiste uno solo per ogni p. Dopo alcune definizioni di base ed esempi pratici, si definisce l'aritmetica di $\mathbb{F}_p$ e si forniscono dei metodi di inversione, tra cui quello che sfrutta l'algoritmo euclideo.
   Si procede nel paragrafo 2.2 con la trattazione della struttura dei campi finiti: si definiscono la caratteristica e l'ordine, si mostrano l'esistenza e l'unicità. Successivamente, si parla di sottocampi e loro proprietà e si mostra che il gruppo moltiplicativo di un campo finito è ciclico. Questo è un risultato importante, poiché nei registri vengono sfruttate le potenze di un generatore di tale gruppo.

Nel paragrafo 2.3 si studiano i campi finiti $\mathbb{F}_{p^n}$ ottenuti dall'anello polinomiale a coefficienti in $\mathbb{F}_p$ attraverso la riduzione modulo un polinomio irriducibile $r(x)$. A questo scopo, vengono illustrate le operazioni in tale anello polinomiale, nelle notazioni complete e compatte, proponendo esempi pratici. Una volta costruito tale campo attraverso $r(x)$, si mostra che è isomorfo all'estensione di campo di $\mathbb{F}_p$ ottenuta aggiungendo una radice u di $r(x)$. Anche qui, vengono presentati dei metodi di calcolo seguiti da esempi pratici, poi si apre una piccola parentesi per mostrare l'importanza che hanno le potenze di u nell'implementazione di un LFSR. Infine, si prosegue con la fattorizzazione in $\mathbb{F}_{p^n}$: costruendo la tabella delle radici del polinomio universale, sfruttando il criterio del sottocampo attraverso i logaritmi, applicando l'algoritmo di Berlekamp.

Il paragrafo 2.4 viene dedicato ai polinomi irriducibili. Anzitutto, viene fornita una formula per calcolarne il numero fissando il grado. In seguito, vengono illustrati metodi semplici per la determinazione, come il metodo della radice ed il metodo della divisione; poi vengono presentati altri metodi come l'algoritmo del polinomio universale, il test della derivata per fattori multipli e l'algoritmo di tutti gli irriducibili. Si introducono inoltre, il polinomio minimo ed i coniugati di una radice di un polinomio irriducibile, definizioni necessarie alla trattazione della teoria di Galois (automorfismo ed orbite di Frobenius, gruppo e corrispondenza di Galois). Questa teoria, in alcuni casi, viene utilizzata per calcolare la correlazione tra due sequenze di un LFSR. Infine, si definiscono traccia e polinomio caratteristico e si mostrano le rispettive proprietà.

Il paragrafo 2.5 viene dedicato ai polinomi primitivi, importanti negli LFSR, poiché generano una sequenza di periodo massimo. Qui, vengono forniti dei metodi per trovarli e la formula per calcolarne il numero.

3. Il terzo capitolo è dedicato ai registri a scorrimento a retroazione li-

neare, gli LFSR. Nel paragrafo 3.1 si studiano le sequenze ricorrenti lineari, la cui generazione può essere implementata su un LFSR, di cui vengono illustrate le componenti e proposti esempi. Poi, viene studiata la periodicità di una sequenza e si dimostra una condizione sufficiente. In seguito, si definiscono:

- La sequenza risposta impulsiva, univocamente determinata dai suoi valori iniziali, la quale fornisce il massimo valore per il periodo minimo.
- Il polinomio caratteristico, associato ad una sequenza: viene mostrato come i termini di una sequenza possano essere rappresentati esplicitamente attraverso le radici del polinomio caratteristico. Inoltre, se il polinomio caratteristico è irriducibile, gli elementi della sequenza possono essere rappresentati in termini di un'appropriata funzione traccia.
- La sequenza di periodo massimo: data una sequenza di ordine k in $\mathbb{F}_p$, il periodo minimo può essere al più $p^k - 1$; il periodo è massimo, ossia uguale a $p^k - 1$, se il polinomio caratteristico è primitivo e se la sequenza ha un vettore dello stato iniziale non nullo.

Nel paragrafo 3.2 si studiano gli LFSR che generano il segnale GPS. Si definiscono anzitutto, la correlazione tra due sequenze e le proprietà da essa soddisfatte. Si mostra esplicitamente che i campi $\mathbb{F}_{2^k}$ ed $\mathbb{F}_2^k$ sono isomorfi, quindi è possibile rappresentare univocamente ogni polinomio di $\mathbb{F}_{2^k}$ in k bit; la somma tra polinomi può essere eseguita efficientemente come semplice XOR bit-a-bit; la moltiplicazione per piccoli coefficienti richiede al massimo uno shift a sinistra e uno XOR. Si conclude con la dimostrazione del Teorema 3.2.2, che rappresenta il risultato più importante di tutto il capitolo. Questo teorema afferma che possiamo inizializzare un LFSR in modo tale da fargli generare una sequenza di periodo massimo poco correlata con ogni traslazione di se stessa.

La tesi si conclude con l'appendice A, dove vengono mostrati due metodi per la costruzione di una tabella delle radici del polinomio universale.

Indice

Capitolo 1

Breve introduzione al funzionamento del GPS

1.1 Sistema di posizionamento globale

1.1.1 La storia ed il sistema

Il progetto GPS è stato sviluppato nel 1973, successivamente creato dal Dipartimento della Difesa statunitense ed è operativo dal 1994. Oltre al GPS, attualmente sono in uso o in fase di sviluppo altri sistemi.

- Il russo Global Navigation Satellite System (GLONASS) è stato impiegato solamente dai militari russi e sovietici, fin quando è stato reso pienamente disponibile anche ai civili nel 2007. Alcuni moderni smartphone, come l'iPhone 4S, il Samsung Galaxy S3 ed il Samsung Galaxy Ace 2, presentano un'antenna in grado di riceve sia i segnali GPS sia i segnali GLONASS.

- La Cina ha realizzato il Sistema di posizionamento Beidou, per uso civile esteso a tutta l'Asia, ed il Sistema di navigazione COMPASS (previsto per il 2020).

- L'India ha pianificato il sistema di navigazione regionale IRNSS, che copre India ed Oceano Indiano.

- L'Unione europea ha in progetto il completamento di una propria rete di satelliti, il Sistema di posizionamento Galileo, per scopi civili e militari. Questo progetto ha un'evidente valenza strategica in quanto la rete statunitense è proprietà dei soli Stati Uniti d'America ed è gestita da autorità militari, che, in particolari condizioni, potrebbero decidere di ridurne la precisione o bloccare selettivamente l'accesso al sistema: la condivisione dell'investimento e della proprietà da parte degli stati utilizzatori garantisce continuità, accessibilità e interoperabilità del servizio europeo.

Il sistema. Il sistema di posizionamento è così composto:

1) *segmento spaziale*: comprende da 24 a 32 satelliti;
2) *segmento di controllo*: composto da una stazione di controllo principale, una stazione di controllo alternativa, varie antenne dedicate e condivise e stazioni di monitoraggio;
3) *segmento utente*: composto dai ricevitori GPS.

Il segmento spaziale. Originariamente il sistema disponeva di 24 satelliti ed erano progettati in modo tale che almeno 21 sarebbero stati funzionanti per il 98% del tempo. Attualmente sono in orbita 31 satelliti attivi nella costellazione GPS (più alcuni satelliti dismessi, alcuni dei quali riattivabili in caso di necessità). I satelliti sono posizionati a 20.184 km circa dalla superficie della Terra. Sono distribuiti lungo sei piani orbitali, ognuno inclinato con un angolo di 55° rispetto al piano equatoriale. Ci sono almeno 4 satelliti per piano orbitale approssimativamente equidistanti l'uno dall'altro. Ogni satellite completa un'orbita circolare attorno alla Terra in 11 ore e 58 minuti. I satelliti sono situati in modo tale che, in qualsiasi momento ed in qualsiasi punto della Terra ci troviamo, possiamo osservarne almeno 4 di essi.

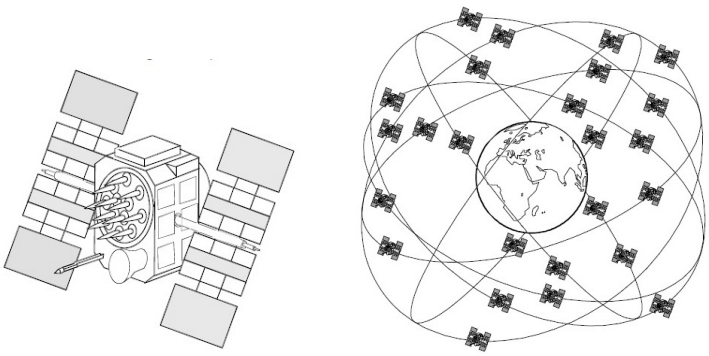

Figura 1.1: Si veda [24, Figure 35-38]. Un satellite GPS e la costellazione di satelliti distribuita attorno alla Terra su sei piani orbitali.

1.1.2 Principio di funzionamento

Determinare la posizione

Intersecando tre sfere il cui raggio è la distanza dal satellite (che conosciamo) con la superficie terrestre si può individuare un punto su di essa. Il principio di funzionamento si basa su un metodo di posizionamento sferico (triangolazione), che parte dalla misura del tempo impiegato da un segnale radio a percorrere la distanza satellite-ricevitore. Poiché il ricevitore non conosce quando è stato trasmesso il segnale dal satellite, per il calcolo della differenza dei tempi il segnale inviato dal satellite è di tipo orario, grazie all'orologio atomico presente sul satellite: il ricevitore calcola l'esatta distanza di propagazione dal satellite a partire dalla differenza tra l'orario pervenuto e quello del proprio orologio sincronizzato con quello a bordo del satellite, tenendo conto della velocità di propagazione del segnale. L'orologio a bordo dei ricevitori GPS è però molto meno sofisticato di quello a bordo dei satelliti e deve essere corretto frequentemente non essendo altrettanto accurato sul lungo periodo. Se il ricevitore avesse anch'esso un orologio atomico al cesio perfettamente sincronizzato con quello dei satelliti sarebbero sufficienti le informazioni fornite da 3 satelliti, ma nella realtà non è così e dunque il ricevitore deve risolvere un sistema di 4 incognite (latitudine, longitudine,

altitudine e tempo) e per riuscirci necessita dunque di 4 equazioni.

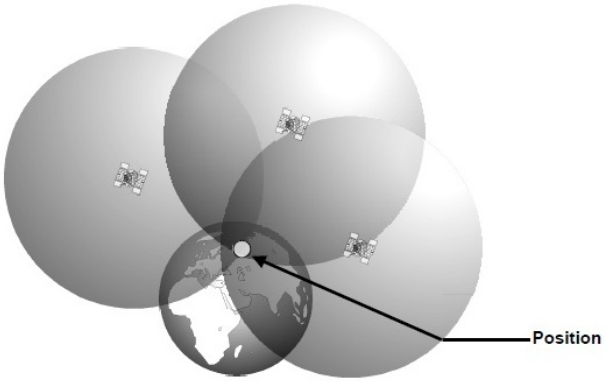

Figura 1.2: Si veda [24, Figure 8]. La posizione viene determinata intersecando, con la superficie terrestre, tre sfere che hanno per centro i satelliti e raggio la distanza di ciascun satellite dalla Terra.

Canali di frequenza e precisione. Ciascun satellite emette continuamente un segnale su due canali separati, comuni a tutta la rete di satelliti: L1, l'unico disponibile all'uso civile, ed L2 per l'uso esclusivamente militare. Vengono quindi usate due codifiche differenti: una codifica pubblica a bassa risoluzione di navigazione (si veda 1.1.3) ed una codifica criptata utilizzata dai militari degli Stati Uniti. Negli ultimi 5-10 anni alcuni modelli di ricevitori GPS per uso civile in campo ingegneristico hanno la possibilità di usufruire del secondo canale L2, permettendo così di raggiungere un margine di precisione centimetrico. Fino al 2000 la precisione del GPS per usi civili era intenzionalmente degradata per decisione del governo statunitense (Selective Availability).

Messaggio di navigazione. Sulle due frequenze portanti L1 C/A ed L2 P/Y, modulate in fase, viene modulato il messaggio di navigazione (il segnale), che ha una velocità di trasmissione pari a 50 bit per secondo con una modulazione numerica di tipo binario, contenente:

- sincronizzazione e tempo della trasmissione del satellite;

- effemeridi satellite (tabelle contenenti precisi dati orbitali);
- almanacco (parametri orbitali approssimati di ogni satellite);
- tempo di correzione dell'informazione per determinare il tempo esatto del satellite;
- effetti di ritardo del segnale dovuti alla ionosfera;
- correzione segnali per calcolare il tempo di trasmissione del segnale;
- stato della costellazione (grado di funzionalità del satellite).

Ogni satellite trasmette continuamente un unico segnale ad esso associato della durata di 1 ms: è come una firma di identificazione. Questa firma è composta da un rumore pseudo-casuale (Codice-PRN = Pseudo Random Noise - Code), ossia un codice di 1023 zeri ed uno. Il periodo del segnale è quindi fissato ed il punto iniziale del ciclo può essere determinato attraverso l'uso dell'almanacco.

Figura 1.3: Si veda [24, Figure 40]. Rumore pseudo casuale

Segmento utente. Importanti funzioni del ricevitore di bordo sono:

- identificare il satellite attraverso la banca dati di codici che quest'ultimo ha in suo possesso;
- calcolare il Δt, ovvero il tempo impiegato dal segnale, che viaggia alla velocità della luce c, per arrivare dal satellite al ricevitore.

Il Δt viene ricavato dalla misura dello scorrimento necessario ad adattare la sequenza dei bit ricevuta dal satellite a quella identica replicata dal ricevitore. Questo Δt determina la distanza esatta tra i satelliti ed il ricevitore. Vediamo cosa accade nel dettaglio. L'orologio atomico permette ad ogni satellite di rimanere sincronizzato con i punti iniziali contenuti nell'almanacco. Quando un ricevitore registra un segnale da un satellite, inizia immediatamente a confrontarlo con quello da esso generato e che combacia perfettamente con quello ricevuto. In generale, questi segnali non combaceranno immediatamente. Quindi, il ricevitore fa scorrere la copia da esso generata finché non combacia con il segnale ricevuto e questo, attraverso dei calcoli, determina la correlazione tra i due.
Ogni satellite trasmette l'almanacco dell'intera costellazione, ma esclusivamente le effemeridi relative a se stesso. In tal modo il ricevitore GPS, mentre effettua il conteggio Doppler[1], riceve i parametri dell'orbita da cui deriva la posizione del satellite: viene così a disporre di tutti gli elementi necessari a definire la posizione nello spazio.

Correzione degli errori. Vediamo ora quali sono le cause principali degli errori. Gli orologi satellitari sono affetti dalle *conseguenze della teoria della relatività*. Infatti, a causa degli effetti combinati della velocità relativa, che rallenta il tempo sul satellite, e della minore curvatura dello spazio-tempo a livello dell'orbita del satellite, che lo accelera, il tempo sul satellite scorre ad un ritmo leggermente più veloce che a terra, rendendo necessaria una correzione automatica da parte dell'elettronica di bordo. Oltre agli errori, compensati, derivanti dagli effetti relativistici, esistono altri tipi di errori del GPS di tipo atmosferico e di tipo elettronico. È importante sottolineare che, quello che permette al GPS di raggiungere la precisione metrica, sono proprio le correzioni della relatività generale e quelle della relatività ristretta. In assenza di queste correzioni si otterrebbero incertezze dell'ordine del chi-

[1]Il conteggio Doppler è una misura diretta della variazione di distanza tra il ricevitore ed il satellite. Per i dettagli si veda il paragrafo 5.8 del seguente file:
http://navigaz.uniparthenope.it/sez_nav/downloads/navigazionesat/capitolo05.pdf

lometro. Gli errori dei GPS sono influenzati dalla diluizione geometrica della precisione e dipendono dagli errori di misura del tempo Δt, da errori numerici, dagli effetti atmosferici, dagli errori delle effemeridi e altri effetti. La più grande fonte di disturbo della dinamica orbitale dei satelliti è la variabilità della pressione della radiazione solare.

Raffinamenti. Le caratteristiche chiave del sistema GPS (accuratezza, integrità, disponibilità) possono essere incrementate grazie all'uso di sistemi di GNSS Augmentation. Tali sistemi possono basarsi su satelliti geostazionari come il WAAS (statunitense) o l'EGNOS (europeo), oppure su collegamenti radio terrestri per distribuire agli utenti le informazioni correttive da applicare durante il calcolo della posizione. Nel caso di collegamenti radio terrestri ci si riferisce a sistemi Ground-based augmentation system (GBAS). La modalità DGPS-IP sfrutta invece la rete Internet per l'invio di informazioni di correzione.

1.1.3 Generare il segnale

Schema a blocchi semplificato

La frequenza di risonanza di ognuno degli orologi atomici genera i seguenti impulsi a tempo e frequenze richieste per le operazioni (Figure 1.4 e 1.5):

- 50Hz di impulsi dati
- **codice C/A** (Coarse/Acquisition), ossia un Codice-PRN a 1.023 MHz, che modula i dati usando l'operazione XOR (exclusive-or) diffondendo i dati su una larghezza di banda di 2 MHz.
- Frequenza del canale civile L1.

Il dato modulato dal codice C/A a sua volta modula il canale L1 usando una codifica a scorrimento binario (BPSK = Binary-Phase-Shift-Keying). Ad ogni modifica del dato modulato vi è un cambiamento di 180° nella fase portante L1.

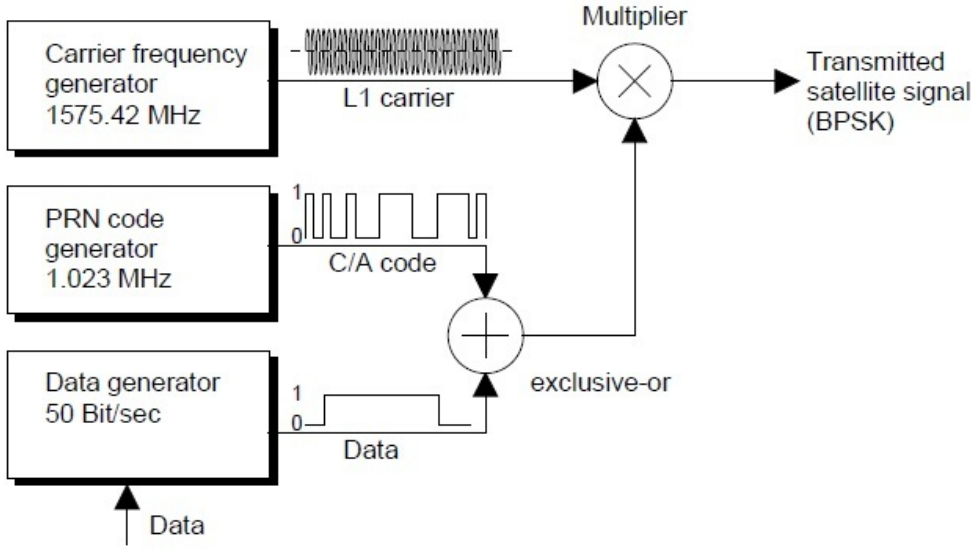

Figura 1.4: Si veda [24, Figure 41]. Schema a blocchi di un satellite semplificato

Il codice C/A gioca un ruolo importante nella modulazione: questo viene generato tramite un feedback shift register (si veda 3.2). Il generatore ha una frequenza di 1.023 MHz ed un periodo di 1023 taps (colpetti = bit) che corrisponde ad 1 ms. Il codice C/A, noto come codice di Gold[2], ha delle vantaggiose proprietà di correlazione. Quando il ricevitore sarà in grado di distinguere tutti i codici C/A attualmente in uso, allora si verificherà una corrispondenza totale, ossia il coefficiente di correlazione ($\mathrm{CF} = \mathrm{Cor}(\cdot,\cdot)/M$, si veda 3.2) sarà 1 e si otterrà un punto di correlazione. Il punto di correlazione viene usato per misurare l'effettivo tempo di transito del segnale e per identificare il satellite.
La qualità della correlazione è espressa qui come un coefficiente di correlazione CF. I valori del CF variano tra -1 ed 1 ed è 1 quando i segnali combaciano del tutto (in sequenza di bit ed in fase). Per la formula del coefficiente di correlazione si rimanda il lettore alla Proposizione 3.2.1, dove bisogna ricor-

[2]Robert Gold: matematico di fama internazionale e ricercatore nella teoria dei codici. http://www.rgcsystems.com/ppl1_gold.htm

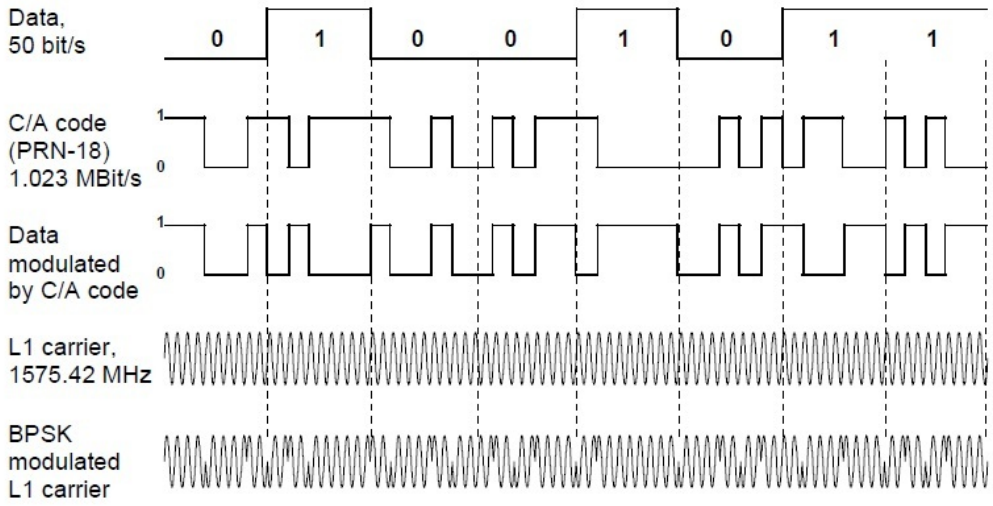

Figura 1.5: Si veda [24, Figure 42]. Struttura dati di un segnale GPS

darsi di dividere tutto per il numero di bit osservati.
Va infine detto che, come risultato dell'Effetto Doppler, poiché i satelliti ed i ricevitori sono in moto relativo tra loro, i segnali trasmessi possono essere spostati fino a ± 5000 Hz dal punto di ricezione. Quindi, la determinazione del tempo di trasmissione del segnale ed il recupero dei dati non richiede solo la correlazione con tutti i possibili codici in tutti i possibili sfasamenti, ma anche l'identificazione della corretta frequenza portante di fase.

Capitolo 2

Campi finiti

Ricordiamo la definizione di campo.

Definizione 2.1 (Campo)**.**
Un *campo* F è un insieme in cui sono definite due operazioni $(+, \cdot)$, ed in cui due elementi speciali distinti, denotati con 0 e $1 \in F$ sono identificati in modo tale da soddisfare le seguenti proprietà:

1. *commutatività*: $\forall\ a, b \in F \quad a + b = b + a$ e $a \cdot b = b \cdot a$.
2. *associatività*: $\forall\ a, b, c \in F \quad (a+b)+c = a+(b+c)$ e $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
3. *distributività*: $\forall\ a, b, c \in F \quad (a + b) \cdot c = (a \cdot c) + (b \cdot c)$.
4. *identità additiva e moltiplicativa*: $\forall\ a \in F \quad a + 0 = a$ e $a \cdot 1 = a$.
5. *esistenza dell'inverso additivo e moltiplicativo*:
   $\forall\ a \in F, \exists\ b \in F$ tale che $a + b = 0$.
   $\forall\ a \in F \setminus \{0\}, \exists\ b \in F$ tale che $a \cdot b = 1$.

Quindi un campo non è altro che un anello commutativo in cui ogni elemento non nullo ha un inverso. Inoltre, un campo F si dice *campo finito* se il numero di elementi in F è finito.

Esempio 2.1 (Il campo $\mathbb{Z}/(p)$)**.**
Sia n un intero positivo. Diciamo che due interi a e b sono equivalenti mod n,

in simboli $a \equiv b \pmod{n}$ se $n \mid a - b$. Le classi di equivalenza in cui $\mathbb{Z}$ viene ripartito dalla relazione di congruenza mod n sono gli insiemi

$$\begin{aligned} [0]_n &= \{\ldots, -2n, -n, 0, n, 2n, \ldots\}, \\ [1]_n &= \{\ldots, -2n+1, -n+1, 1, n+1, 2n+1, \ldots\}, \\ &\vdots \\ [n-1]_n &= \{\ldots, -3n-1, -2n-1, n-1, 2n-1, 3n-1, \ldots\}. \end{aligned}$$

Consideriamo l'anello finito $\mathbb{Z}/(n)$ delle classi di equivalenza di interi modulo n, dove $[a]_n = a + (n)$ e (n) è l'ideale principale generato da n.
Se $n = p$, con p primo, $\mathbb{Z}/(p)$ è un campo finito. Questo segue dal fatto che $\mathbb{Z}$ è un dominio ad ideali principali. Di conseguenza, (p) è un ideale massimale ed un ideale primo di $\mathbb{Z}$. (Per la dimostrazione si rimanda il lettore a [12, p.17].)

2.1 Il campo $\mathbb{F}_p$

Definizione 2.2 (Campo di Galois di ordine p)**.**
Sia p un primo e sia $\mathbb{F}_p := \{0, 1, \ldots, p-1\}$.
Sia $\gamma\colon \mathbb{Z}/(p) \to \mathbb{F}_p$ la mappa definita da $\gamma([a]_p) = a$ per $a = 0, 1, \ldots, p-1$.
Allora $\mathbb{F}_p$, dotato della struttura di campo indotta da γ, è un campo finito, detto il *campo di Galois di ordine p*.

La mappa γ è un isomorfismo, quindi i due campi sono isomorfi: in simboli $\mathbb{Z}/(p) \cong \mathbb{F}_p$. $\mathbb{F}_p$ contiene 0 l'elemento zero, 1 l'identità e la sua struttura è esattamente la struttura di $\mathbb{Z}/(p)$. Perciò il calcolo con gli elementi di $\mathbb{F}_p$ equivale all'aritmetica degli interi modulo p. Poiché $\mathbb{Z}$ è un dominio euclideo, possiamo dividere ogni intero a per p ottenendo un quoziente q ed un resto r tale che $0 \leq r < p$. Chiaramente, due elementi qualsiasi della stessa classe di equivalenza hanno stesso resto r, così prendiamo r come rappresentante canonico delle classi di equivalenza.

2.1.1 Aritmetica di $\mathbb{F}_p$

Definizione 2.3 (Somma, prodotto, opposto)**.**
Siano α e β due classi di congruenza modulo p. Scegliamo due interi a e b tali che $\alpha = [a]_p$ e $\beta = [b]_p$. Allora la *somma* e il *prodotto* di α e β sono definiti come:

$$\alpha + \beta := [a + b]_p.$$

$$\alpha \cdot \beta := [a \cdot b]_p.$$

Definiamo anche l'*opposto* di $[a]_p$ come

$$-\alpha := [-a]_p.$$

La somma ed il prodotto di classi di congruenza sono ben definite, ossia se scegliamo degli a e b diversi, le classi che abbiamo definito non cambiano. Per dividere basta sapere come invertire: se possiamo calcolare $1/b$, allora possiamo usare il prodotto per calcolare $a/b = a \cdot 1/b$. Ci sono quattro metodi per calcolare l'inverso.

Primo metodo di inversione in $\mathbb{F}_p$

Una possibilità è quella di cercare l'inverso. Poiché $\mathbb{F}_p$ è finito, è possibile elencare tutti gli elementi del campo e vedere qual è quello giusto.

Esempio 2.2. Vogliamo calcolare $[2]_5^{-1}$.

$$[2]_5 \cdot [0]_5 = [0]_5, \;\; [2]_5 \cdot [1]_5 = [2]_5, \;\; [2]_5 \cdot [2]_5 = [4]_5, \;\; [2]_5 \cdot [3]_5 = [1]_5 \; !$$

Così $[2]_5^{-1} = [3]_5$.

Il metodo di ricerca dell'inverso è però accettabile solo per p piccolo.

Secondo metodo di inversione in $\mathbb{F}_p$

La dimostrazione dell'esistenza di inversi mod p è una dimostrazione costruttiva che ci dà il secondo metodo di inversione.
Ricordando che l'operazione mod p è un omomorfismo da $\mathbb{Z}$ a $\mathbb{F}_p$, definiamo l'operazione opposta.

Definizione 2.4 (Sollevamento)**.**
Siano $[a]_p \in \mathbb{F}_p$ e $a \in \mathbb{Z}$. L'operazione che manda $[a]_p \longmapsto a$, si dice *sollevamento.*

Questa non è una funzione inversa, ovviamente, poiché l'operazione mod p non è iniettiva. Sia $\alpha = [a]_p$ un elemento di $\mathbb{F}_p$. Solleviamo α a $\mathbb{Z}$ ottenendo a. Se $a = 0$, cioè se a è un multiplo di p, allora non ci aspettiamo alcun inverso. Ma se a non è un multiplo di p, allora a è relativamente primo a p. Questo significa che il massimo comun divisore tra a e p è 1, in simboli: $\mathrm{mcd}(a, p) = 1$. Usando l'*algoritmo euclideo*, possiamo sempre scrivere il risultato dell'algoritmo come combinazione lineare di a e di p, cioè

$$1 = as + pt$$

per qualche intero s e t. Ma pt è un multiplo di p, quindi è $0 \bmod p$, così

$$1 \equiv as \pmod{p}.$$

Questo significa che la classe d'equivalenza contenente s è l'inverso della classe d'equivalenza contenente a. Sopra abbiamo ''sollevato'' a $\mathbb{Z}$ per fare il massimo comun divisore; ora torniamo al mod p in $\mathbb{F}_p$ e diciamo che $[s]_p$ è l'inverso di α.
Questo metodo è utile per provare l'esistenza degli inversi in $\mathbb{F}_p$. Vediamo quindi come si calcola il massimo comun divisore tra due interi positivi.

L'algoritmo euclideo

Siano a e b due numeri interi positivi. Supponiamo $b > a$ e dividiamo b per a:

$$b = q_1 a + r_1, \quad 0 \leq r_1 < a.$$

Se $r_1 = 0$, allora $a|b$ e $\mathrm{mcd}(a, b) = a$.
Supponiamo $r_1 \neq 0$. Ogni intero positivo d che divide sia a che b, dividerà anche $r_1 = b - q_1 a$. Viceversa, se d divide a ed r_1, dividerà pure b.
La conclusione è che $M = \mathrm{mcd}(a, b) = \mathrm{mcd}(a, r_1)$.

Dividiamo ora a per r_1: $a = q_2 r_1 + r_2, \quad 0 \leq r_2 < r_1$.
Se $r_2 = 0$, allora $M = r_1$, altrimenti $M = \text{mcd}(r_1, r_2)$.
Iterando il procedimento, se $r_i = 0$ ci fermiamo e se $r_i \neq 0$ otteniamo r_{i+1} come

$$r_{i-1} = q_{i+1} r_i + r_{i+1}, \quad 0 \leq r_{i+1} < r_i.$$

Siccome la successione dei resti r_i è strettamente decrescente e tutti i resti sono positivi o nulli, arriveremo al punto di doverci fermare, cioè esisterà un n tale che $r_n \neq 0$, mentre $r_{n+1} = 0$, ovvero $r_n | r_{n-1}$. Allora $M = r_n$.
In altri termini, il massimo comun divisore di a e di b è l'ultimo resto diverso da 0 di questa catena di divisioni successive.
Questo algoritmo, che ha considerevole importanza pratica e teorica, permette di calcolare il massimo comun divisore di a e di b senza scomporre né a né b in fattori primi. Questo è estremamente utile, in quanto la fattorizzazione in primi di un intero grande è molto difficile da calcolare.
Possiamo anche utilizzare questa successione di divisioni per calcolare gli interi s e t tali che $as + bt = M$. Si procede a ritroso partendo dall'ultima uguaglianza con resto non nullo:

$$r_n = r_{n-2} - q_n r_{n-1}.$$

Si sostituisce il valore di r_{n-1} ricavato dall'uguaglianza $r_{n-3} = q_{n-1} r_{n-2} + r_{n-1}$ e si ottiene

$$r_n = r_{n-2} - q_n (r_{n-3} - q_{n-1} r_{n-2}) = (1 + q_n q_{n-1}) r_{n-2} - q_n r_{n-3}.$$

Ora, sostituendo in questa equazione il valore di r_{n-2} ricavato da $r_{n-4} = q_{n-2} r_{n-3} + r_{n-2}$, esprimeremo r_n come combinazione lineare di r_{n-3} ed r_{n-4} e cos'ì via. Si arriverà ad esprimere r_n come combinazione lineare di a e di b.

Osservazione 1. Talvolta è possibile scrivere $b = q'a - r'$, dove $0 < r' < a$ ed è più piccolo del resto (positivo) della divisione di b per a. Infatti possiamo scrivere $q'a - r' = (q + 1)a - (b - r)$, e se $r > b/2$ allora $0 < b - r < r$. Si avrà ancora $M = \text{mcd}(a, r')$ e quindi nell'algoritmo euclideo è possibile ammettere dei resti negativi.

Esempio 2.3. Con $a = 50$ e $b = p = 97$, applicando l'algoritmo euclideo otteniamo:

$$\begin{aligned} 97 &= 50 + 47 && r_1 = 47 \\ 50 &= 47 + 3 && r_2 = 3 \\ 47 &= 3 \cdot 15 + 2 && r_3 = 2 \\ 3 &= 2 + 1 && r_4 = 1 \\ 2 &= 1 \cdot 2 + 0 && r_5 = 0 \end{aligned}$$

Effettuiamo ora lo stesso calcolo, avendo cura di prendere il resto negativo se questo è più piccolo in valore assoluto, ossia quando il resto positivo è maggiore di metà del divisore. Per esempio, nella prima divisione $97 = 50+47$, 47 è maggiore di metà di 50 e perciò converrà aumentare il quoziente di 1 e ottenere come resto $47 - 50 = -3$.

$$\begin{aligned} 97 &= 50 \cdot 2 - 3 && r_1 = -3 \\ 50 &= 3 \cdot 17 - 1 && r_2 = -1 \\ 3 &= 1 \cdot 3 + 0 && r_3 = 0 \end{aligned}$$

Col secondo metodo ci sono volute solo tre divisioni invece di cinque. Troviamo degli interi s e t con $50s + 97t = 1$ partendo dalla seconda divisione e procedendo all'indietro.

$$\begin{aligned} 1 &= 3 \cdot 17 - 50 \\ 3 &= 50 \cdot 2 - 97 \quad \text{sostituiamo nella precedente ottenendo} \\ 1 &= (50 \cdot 2 - 97) \cdot 17 - 50 \\ 1 &= 50 \cdot 33 - 97 \cdot 17. \end{aligned}$$

Sollevando ad $\mathbb{F}_{97}$, poiché $97 \cdot 17 \equiv 0 \mod 97$, possiamo concludere che $[50]_{97}^{-1} = [33]_{97}$.

Terzo metodo di inversione in $\mathbb{F}_p$

Il terzo metodo usa un po' di teoria dei gruppi. Ricordiamo che in ogni campo F, gli elementi non nulli formano un gruppo con la moltiplicazione detto *gruppo moltiplicativo* del campo, indicato con $F^\times$. Se F ha p elementi, allora ci sono $p - 1$ elementi non nulli in $F^\times$. Ricordiamo il teorema di

Lagrange: l'ordine di un sottogruppo divide l'ordine del gruppo. Inoltre, dato un elemento a di un gruppo finito, possiamo formare il sottogruppo ciclico generato dall'elemento, semplicemente prendendo potenze di a fino ad ottenere l'identità id. Ricordiamo anche che l'ordine di un elemento (un intero positivo k tale che $a^k = id$) è uguale all'ordine del sottogruppo ciclico generato da a.
Così, dato G gruppo finito, con $|G|$ ordine di G; se l'ordine di a è k, per il teorema di Lagrange k divide $|G|$ ed abbiamo che $\forall a \in G,\ a^{|G|} = id$.
Per il caso $G = \mathbb{F}_p^\times$, dove $|G| = p-1$ e l'identità è 1, mostreremo al Corollario 2.2.4, che $a^{p-1} = 1$ per ogni elemento non nullo in $\mathbb{F}_p$. (Si rimanda il lettore al [22, Teorema di Fermat 1.7.5]). Questo ha due conseguenze dirette: vediamo la prima. Possiamo scrivere $a^{p-2} = a^{-1}$ per ogni elemento non nullo, che ci fornisce il terzo algoritmo per l'inversione: l'algoritmo $p-2$. La seconda: moltiplicando $a^{p-1} = 1$ per a otteniamo $a^p = a$, relazione soddisfatta da tutti gli elementi di $\mathbb{F}_p$ (incluso lo zero), come vedremo nel Corollario 2.2.4.
Così, dato l'algoritmo $p-2$, possiamo facilmente invertire in $\mathbb{F}_p$ per p piccolo.

Esempio 2.4. Come prima, vogliamo calcolare $[2]_5^{-1}$.
Poiché $5-2 = 3$, possiamo invertire ogni elemento semplicemente elevandolo al cubo. $2^3 = 8$ che è 3 mod 5. Così $[2]_5^{-1} = [3]_5$.
Possiamo verificarlo scrivendo $[2]_5 \cdot [3]_5 = [1]_5$.

Quarto metodo di inversione in $\mathbb{F}_p$

Vedremo più avanti un altro teorema sui campi finiti (Teorema 2.2.14) in cui si afferma che il gruppo moltiplicativo di un campo finito è ciclico. Già Gauss provò l'esistenza di *elementi primitivi* mod p, cioè l'esistenza di almeno un elemento di $\mathbb{F}_p^\times$ tale che tutti gli altri elementi siano potenze di esso. Questo significa che esiste un elemento g (*generatore*) tale che tutti gli elementi in $\mathbb{F}_p^\times$ sono potenze di g. Poiché $a^{p-1} = 1$, questi esponenti in realtà vengono presi mod $p-1$. Se $a = g^k$, allora $a^{-1} = g^{-k}$, o equivalentemente $a^{-1} = g^{p-1-k}$.
In questo metodo ci vogliono dei calcoli per trovare un generatore g. Ne

cerchiamo uno. Così facendo viene fuori che col 2 o con il 3 funziona il più delle volte, altrimenti si prova con il 5 ecc.

Esempio 2.5. Con $p = 5$ e $p = 7$ le potenze di 2 sono

k	1	2	3	4
2^k	2	4	3	1

k	1	2	3
2^k	2	4	1

Con $p = 5$ funziona per tutto $\mathbb{F}_5^\times$, così 2 è primitivo mod 5. Con $p = 7$ non funziona per tutto $\mathbb{F}_7^\times$. Ma, le potenze di 3 sono

k	1	2	3	4	5	6
3^k	3	2	6	4	5	1

così 3 è primitivo mod 7. Analogamente si può vedere che 2 è primitivo mod 11.

Useremo il mod 11 come esempio di inversione log, poiché 11 è grande abbastanza per essere un po' più interessante.

Definizione 2.5 (Logaritmo discreto)**.**
Poiché $\mathbb{F}_p^\times$ è ciclico, possiamo scrivere ciascun a diverso da zero come $a = g^k$. $\forall\ a \in \mathbb{F}_p^\times$ definiamo il *logaritmo discreto* in base g come

$$\log_g \colon \mathbb{F}_p^\times \longrightarrow \mathbb{Z}/(p-1), \quad \text{con } \log_g(a) = k.$$

D'ora in poi, per semplicità, parleremo soltanto di logaritmi, sottointendendo l'aggettivo "discreti".
Ora possiamo trascrivere elementi di un campo e loro logaritmi, ottenendo la tabella dei logaritmi. Inoltre, ordinando i valori della tabella dei logaritmi, otteniamo la tabella degli antilogaritmi.

$\log_2(a) = k$	0	1	8	2	4	9	7	3	6	5
$a = 2^k$	1	2	3	4	5	6	7	8	9	10

$\log_2(a) = k$	0	1	2	3	4	5	6	7	8	9
$a = 2^k$	1	2	4	8	5	10	9	7	3	6

Poiché $a^{-1} = g^{-k} = g^{p-1-k}$, possiamo invertire semplicemente usando la tabella degli antilogaritmi. Ad esempio $8 = 2^3$ (tre caselle da destra nella prima tabella), quindi $[1]_{11}/[8]_{11} = [2^{10-3}]_{11} = [2^7]_{11} = [7]_{11}$ (tre caselle da destra nella seconda tabella).
Possiamo moltiplicare usando le tabelle dei logaritmi e degli antilogaritmi con la nota regola

$$\log_g(ab) \equiv \big(\log_g(a) + \log_g(b)\big) \mod (p-1), \quad \text{ottenendo}$$

$$ab = g^{\log_g(ab)} = g^{\log_g(a)+\log_g(b)}.$$

Analogamente, si può fare la divisione in un solo passaggio, usando

$$\log_g(a/b) \equiv \big(\log_g(a) - \log_g(b)\big) \mod (p-1), \quad \text{ottenendo}$$

$$a/b = g^{\log_g(a/b)} = g^{\log_g(a)-\log_g(b)}.$$

2.2 Struttura di campi finiti

2.2.1 Caratteristica ed ordine di campi finiti

Definizione 2.6 (Caratteristica)**.**
La *caratteristica* di un campo F è definita come il più piccolo intero $m > 0$ tale che

$$m \cdot 1 = \overbrace{1 + 1 + \cdots + 1}^{m \text{ volte}} = 0,$$

se un tale intero esiste, è 0 altrimenti.

Quindi se $1 \in F$ ha ordine infinito la caratteristica è zero, altrimenti è uguale all'ordine di 1.

Teorema 2.2.1. *La caratteristica di un campo finito è sempre un primo.*

Dimostrazione. Innanzitutto, un campo finito ha sempre caratteristica positiva, perché l'ordine di un elemento di un gruppo finito è sempre finito. Supponiamo che la caratteristica di F sia un intero positivo p, mostriamo che p è primo. Sicuramente $p \neq 1$, perché altrimenti in F si avrebbe $1 = 0$ ed F sarebbe l'anello nullo. Se fosse $p = mn$ per m ed n interi positivi, allora si avrebbe $(m1)(n1) = (mn)1 = p1 = 0$ e quindi $m1 = 0$ oppure $n1 = 0$, perché F è un dominio. Ma p è l'ordine di 1 in F e questo implicherebbe che $p|m$ oppure $p|n$. Concludiamo che $m = p$ oppure $n = p$ e quindi p è primo. □

Vogliamo far vedere che **l'ordine di un campo finito è una potenza di un primo**. A questo scopo, seguono definizioni e risultati preliminari.
Sia F un campo. Un sottoinsieme K di F, che è un campo rispetto alle operazioni di F, si dice *sottocampo* di F. In questo contesto, F si dice *estensione* di K. Se $K \neq F$, diciamo che K è un *sottocampo proprio* di F.
Un campo che non contiene sottocampi propri si dice *campo primo*. L'intersezione di un qualsiasi insieme non vuoto di sottocampi di un dato campo F è ancora un sottocampo di F. Se facciamo l'intersezione di tutti i sottocampi di F, otteniamo il *sottocampo primo* di F.
Se K è un'estensione di F, allora K è uno spazio vettoriale su F rispetto alle operazioni di campo di F. Il *grado* di K su F è la dimensione di K come spazio vettoriale su F. Denoteremo con $[K\colon F]$ il grado di K su F.

Lemma 2.2.2. *Sia F un campo finito con q elementi e sia $F \subset K$, dove K è anch'esso un campo finito. Allora K ha q^n elementi con $n = [K\colon F]$.*

Dimostrazione. Essendo K finito ed essendo uno spazio vettoriale su F, è uno spazio vettoriale di dimensione finita su F. Sia $n = [K\colon F]$; allora K ha una base di n elementi su F. Sia $(v_1, \ldots, v_n)$ tale base, ogni elemento di K ha una rappresentazione unica nella forma $\alpha_1 v_1 + \cdots + \alpha_n v_n$ con $\alpha_1, \ldots, \alpha_n$

in F. Pertanto il numero degli elementi di K è uguale al numero degli $\alpha_1 v_1 + \cdots + \alpha_n v_n$ al variare di $\alpha_1, \ldots, \alpha_n$ in F; siccome ogni coefficiente può avere q valori, K deve chiaramente avere q^n elementi. □

Teorema 2.2.3. *Supponiamo che F sia un campo finito di caratteristica p, allora F contiene esattamente p^n elementi per qualche intero positivo n.*

Dimostrazione. Supponiamo che $\mathbb{F}_p$ sia il sottocampo primo di F. Applicando il Lemma 2.2.2, con $\mathbb{F}_p = F$ ed $F = K$, notiamo che ci sono soltanto p possibili scelte per ogni coordinata α_i, quindi il numero totale di elementi in F è

$$\overbrace{p \cdot p \cdots p}^{n \text{ volte}} = p^n.$$

□

Abbiamo visto nel Teorema 2.2.1 che la caratteristica di un campo finito è sempre un primo: possiamo quindi concludere che l'ordine di un campo finito è una potenza di un primo.

2.2.2 Esistenza ed unicità

Corollario 2.2.4. *Se un campo finito F ha q elementi, ogni elemento $a \in F$ verifica $a^q = a$.*

Dimostrazione. Se $a = 0$ l'asserto del corollario è banalmente vero. D'altro canto, gli elementi non nulli di F formano, rispetto alla moltiplicazione, un gruppo G con $q - 1$ elementi (il gruppo moltiplicativo di F). Dunque, per il Teorema di Fermat astratto, $a^{q-1} = 1$ per ogni $a \in G$. Moltiplicando per a questa relazione, otteniamo che $a^q = a$. □

Quindi gli elementi di F sono soluzioni dell'equazione $x^q - x = 0$. D'altra parte, poiché tali soluzioni sono al più q e F ha esattamente q elementi, ne segue che gli elementi di F sono tutte e sole le radici del polinomio $x^q - x$.

Lemma 2.2.5. *Se F è un campo finito con q elementi, il polinomio $x^q - x$ si fattorizza, in $F[x]$, nella forma*

$$x^q - x = \prod_{\lambda \in F} (x - \lambda).$$

Dimostrazione. Sappiamo che il polinomio $x^q - x$ ha al più q radici in F. Ma per il corollario precedente, sono già note q radici, in quanto sono tutti gli elementi di F. Inoltre, se $a \in F$ è radice di $p(x) \in F[x]$, allora $(x - a)|p(x)$. Si può dunque concludere che

$$x^q - x = \prod_{\lambda \in F} (x - \lambda).$$

□

Lemma 2.2.6. *Per ogni numero primo p ed ogni intero positivo n esiste un campo avente p^n elementi.*

Dimostrazione. Per $q = p^n$ consideriamo il polinomio $x^q - x$ in $\mathbb{F}_p[x]$, l'anello dei polinomi in x su $\mathbb{F}_p$. Sia K il campo di spezzamento di questo polinomio su $\mathbb{F}_p$. Questo polinomio ha q radici distinte in K poiché la sua derivata $qx^{q-1} - 1 = -1$ in $\mathbb{F}_p[x]$ e quindi non può avere alcuna radice in comune con $x^q - x$. Sia $S := \{a \in K : a^q - a = 0\}$. Allora S è un sottocampo di K poiché: (i) S contiene 0 e 1; (ii) $a, b \in S$ implica che $(a - b)^q = a^q - b^q = a - b$ e quindi $a - b \in S$; (iii) per $a, b \in S$ e $b \neq 0$ abbiamo $(ab^{-1})^q = a^q b^{-q} = ab^{-1}$ e quindi $ab^{-1} \in S$. Ma, d'altra parte $x^q - x$ si deve spezzare in K poiché K contiene tutte le sue radici. Quindi $K = S$ e poiché S ha q elementi, K è un campo finito con q elementi. □

Corollario 2.2.7. *Un campo F con q elementi è il campo di spezzamento del polinomio $x^q - x$.*

Dimostrazione. Per il lemma precedente, $x^q - x$ si spezza certamente in F. Però non si può spezzare in un campo più piccolo, perché tale campo dovrebbe contenere tutte le radici del polinomio e quindi avere almeno q elementi. Dunque il campo di spezzamento di $x^q - x$ è proprio F. □

Poiché due campi di spezzamento di un dato polinomio su un dato campo sono sempre isomorfi, otteniamo il seguente teorema.

Teorema 2.2.8. *Due campi finiti aventi lo stesso numero di elementi sono isomorfi.*

Combinando il Teorema 2.2.8 ed il Lemma 2.2.6 abbiamo il seguente risultato.

Teorema 2.2.9 (Unicità)**.**
Per ogni numero primo p ed ogni intero positivo n esiste un unico campo avente p^n elementi.

2.2.3 Sottocampi e gruppo moltiplicativo

Vediamo dei risultati che ci serviranno per dimostrare il criterio del sottocampo.

Lemma 2.2.10. *Siano F un campo ed $F[x]$ l'anello polinomiale nella variabile x a coefficienti in F; siano m ed n degli interi positivi. Allora $x^m - 1$ divide $x^n - 1$ in $F[x]$ se e solo se m divide n.*

Dimostrazione. Poniamo $n = ms + r$, con $s, r \in \mathbb{Z}$, $0 \leq r < m$. Allora

$$x^n - 1 = x^{ms+r} - 1 = x^r(x^{ms} - 1) + (x^r - 1).$$

Si ha

$$x^{ms} - 1 = (x^m - 1)(x^{m(s-1)} + x^{m(s-2)} + \cdots + x^m + 1),$$

e quindi $x^m - 1$ divide $x^{ms} - 1$. Perciò $x^r - 1$ è il resto della divisione di $x^n - 1$ per $x^m - 1$ e $x^r - 1 = 0$ se e solo se $r = 0$. □

Lemma 2.2.11. *Siano m, n, q interi positivi, con $q > 1$. Allora $q^m - 1$ divide $q^n - 1$ se e solo se m divide n.*

Dimostrazione. Ponendo $n = ms + r$, $0 \leq r < m$, si ha

$$q^n - 1 = q^{ms+r} - 1 = q^r(q^{ms} - 1) + (q^r - 1),$$

inoltre $(q^m - 1)|(q^{ms} - 1)$ per il Lemma 2.2.10. Siccome $q > 1$ e $m > r$, si ha $q^m - 1 > q^r - 1 \geq 0$, per cui $q^r - 1$ è il resto della divisione di $q^n - 1$ per $q^m - 1$ e $q^r - 1 = 0$ se e solo se $r = 0$. □

Teorema 2.2.12 (Criterio del sottocampo)**.**
Sia $\mathbb{F}_q$ il campo finito con $q = p^n$ elementi. Allora ogni sottocampo di $\mathbb{F}_q$ ha ordine p^m, con $m|n$. Viceversa, se $m|n$, allora esiste esattamente un sottocampo di $\mathbb{F}_q$ con p^m elementi.

Dimostrazione. È chiaro che un sottocampo K di $\mathbb{F}_q$ ha ordine p^m per qualche intero positivo $m \leq n$. Il Lemma 2.2.2 mostra che $q = p^n$ deve essere una potenza di p^m e quindi m è necessariamente un divisore di n.
Viceversa, se $m|n$, allora $p^m - 1|p^n - 1$ per il Lemma 2.2.11 e quindi $x^{p^m-1} - 1$ divide $x^{p^n-1} - 1$ in $\mathbb{F}_p[x]$ ([22, Lemma 5.14.9]). Di conseguenza, $x^{p^m} - x$ divide $x^{p^n} - x = x^q - x$ in $\mathbb{F}_p[x]$. Così, ogni radice di $x^{p^m} - x$ è una radice di $x^q - x$ e quindi appartiene a $\mathbb{F}_q$. Segue che $\mathbb{F}_q$ deve contenere come sottocampo un campo di spezzamento di $x^{p^m} - x$ su $\mathbb{F}_p$, che sappiamo avere ordine p^m. Se ci fossero due sottocampi distinti di ordine p^m in $\mathbb{F}_q$, insieme conterrebbero più di p^m radici di $x^{p^m} - x$ in $\mathbb{F}_q$, un'ovvia contraddizione. □

Introduciamo ora la funzione e la formula di inversione di Möbius: queste ci serviranno per mostrare che **il gruppo moltiplicativo di un campo finito è ciclico** e per trovare il numero di polinomi monici irriducibili in $\mathbb{F}_q[x]$ di grado n. Per le dimostrazioni si rimanda il lettore al [12, 3.2] o in alternativa al [22, 5.14].

Definizione 2.7 (Funzione di Möbius)**.**
La *funzione di Möbius* μ è la funzione su $\mathbb{N}$ definita da

$$\mu(n) = \begin{cases} 1 & \text{se } n = 1, \\ (-1)^k & \text{se } n \text{ è il prodotto di } k \text{ primi distinti}, \\ 0 & \text{se } n \text{ è divisibile per il quadrato di un primo.} \end{cases}$$

Formula di inversione di Möbius

Sia $f\colon \mathbb{Z}^+ \to \mathbb{Z}$ una funzione. Sia $g\colon \mathbb{Z}^+ \to \mathbb{Z}$ la funzione definita da

$$g(n) = \sum_{d|n} f(d) \tag{2.1}$$

dove la sommatoria è presa su tutti i divisori positivi di n. Allora

$$f(n) = \sum_{d|n} \mu(n/d)g(d) = \sum_{d|n} \mu(d)g(n/d). \tag{2.2}$$

Vediamo una delle applicazioni della formula di Möbius: ricaviamo un'espressione per la funzione φ di Eulero. Ricordiamo che se $n > 1$ è un intero, allora $\varphi(n)$ è il numero di interi k con $1 \leq k \leq n$ relativamente primi ad n.
Se $d = \mathrm{mcd}(k, n)$ ed $e = k/d$, allora $1 \leq e \leq n/d$. Inoltre $\mathrm{mcd}(e, n/d) = 1$: se l è un intero positivo con $l|e$ ed $l|(n/d)$, allora $dl|de = k$, $dl|n$ e quindi $dl \leq d$, ossia $l = 1$. Viceversa, se $d|n$, e un intero tale che $1 \leq e \leq n/d$, $\mathrm{mcd}(e, n/d) = 1$ e allora $1 \leq de \leq n$, $\mathrm{mcd}(de, n) = d$. Perciò se $d|n$, il numero di interi k con $1 \leq k \leq n$ e $\mathrm{mcd}(k, n) = d$ è uguale al numero di interi e con $1 \leq e \leq n/d$ e $\mathrm{mcd}(e, n/d) = 1$, ossia $\varphi(n/d)$. Ne deduciamo che

$$n = \sum_{d|n} \varphi(n/d) = \sum_{d|n} \varphi(d).$$

Applicando la formula di inversione di Möbius 2.2 otteniamo la formula seguente.

Proposizione 2.2.13. *Per ogni $n > 1$ si ha*

$$\varphi(n) = \sum_{d|n} \mu(d)n/d.$$

Teorema 2.2.14. *Sia F un campo e sia G un sottogruppo finito del gruppo abeliano $F^\times$. Allora G è ciclico.*

Ricaviamo subito un corollario.

Corollario 2.2.15. *Se F è un campo finito, il gruppo $F^\times$ è ciclico.*

Dimostrazione. (Teorema 2.2.14). Sia n il numero di elementi di G. Possiamo assumere $n > 1$. Il polinomio $x^n - 1$ ha esattamente n radici e

$$x^n - 1 = \sum_{a \in G} (x - a).$$

Prendiamo ora un divisore positivo d di n. Per il Lemma 2.2.10, $x^d - 1 | x^n - 1$ e perciò le radici di $x^d - 1$ sono tutte contenute in G e sono precisamente gli elementi di G il cui ordine divide d. Inoltre $x^d - 1$ sarà prodotto di fattori distinti del tipo $x - a$ con $a \in G$. Ne segue che $x^d - 1$ ha d radici distinte e quindi che G contiene esattamente d elementi il cui ordine divide d.
Per ogni intero positivo d chiamiamo $f(d)$ il numero di elementi di G di ordine d, e $g(d)$ il numero di elementi di G il cui ordine divide d. Abbiamo appena visto che $g(d) = d$ se $d|n$. Ovviamente $g(d) = \sum_{k|d} f(k)$. Per la formula di inversione di Möbius e per la proposizione 2.2.13, si ha

$$f(n) = \sum_{d|n} \mu(d) g(n/d) = \sum_{d|n} \mu(d) n/d = \varphi(n)$$

dove φ è la φ di Eulero. Ma $\varphi(n) > 0$ e G contiene un elemento di ordine n. Ne segue che G è ciclico. □

2.3 Il campo $\mathbb{F}_{p^n}$

Abbiamo visto nel paragrafo 2.2.1 che tutti i campi finiti di un dato ordine sono isomorfi tra loro. Questo significa che, dati p ed n, se possiamo costruire anche solo un campo finito di ordine p^n, in un certo senso li avremo costruiti tutti. Non ci resta che costruire tale campo.
Per fare ciò, iniziamo dando uno sguardo ad $\mathbb{F}_p[x]$ l'anello dei polinomi con coefficienti nel campo $\mathbb{F}_p$.

2.3.1 Aritmetica di $\mathbb{F}_p[x]$

$\mathbb{F}_p[x]$ è un dominio euclideo e, come per l'anello degli interi, esiste una divisione con resto.

Teorema 2.3.1 (Algoritmo di divisione)**.**
Siano $g \neq 0$ un polinomio in $\mathbb{F}_p[x]$ e $\deg(g)$ il grado di g. Allora per ogni $f \in \mathbb{F}_p[x]$ esistono dei polinomi $q, r \in \mathbb{F}_p[x]$ tali che

$$f = qg + r, \quad \textit{dove } \deg(r) < \deg(g).$$

Esempio 2.6. Siano $f, g \in \mathbb{F}_5[x]$ con

$$f(x) = 2x^5 + x^4 + 4x + 3, \qquad g(x) = 3x^2 + 1.$$

Calcoliamo i polinomi $q, r \in \mathbb{F}_5[x]$ con $f = qg + r$ usando la divisione in colonna:

Si procede in questo modo:

$$\begin{array}{r|l}
 & 4x^3 + 2x^2 + 2x + 1 \\
\hline
3x^2 + 1 & 2x^5 + x^4 \qquad\qquad + 4x + 3 \\
 & -2x^5 \qquad -4x^3 \\
\hline
 & x^4 + x^3 \\
 & -x^4 \qquad\quad -2x^2 \\
\hline
 & x^3 + 3x^2 + 4x \\
 & -x^3 \qquad\quad -2x \\
\hline
 & 3x^2 + 2x + 3 \\
 & -3x^2 \qquad -1 \\
\hline
 & 2x + 2
\end{array}$$

- Dividiamo $2x^5$ per $3x^2$ usando l'aritmetica di $\mathbb{F}_5$: otteniamo $4x^3$.
- Moltiplichiamo $4x^3$ per $3x^2 + 1$ e cambiamo segno.
- Incolonniamo sotto il dividendo e sommiamo avendo cura di aggiungere 5 quando si ottengono dei coefficienti negativi.
- Dividiamo x^4 per $3x^2$ e procediamo come nel punto di partenza.

Quindi, per ottenere le cifre del quoziente, dividiamo il coefficiente direttore del resto in corso per il coefficiente direttore del divisore (necessariamente non nullo). Ci fermiamo quando il grado del resto in corso è inferiore al grado del divisore. Il risultato trovato è:

$$q(x) = 4x^3 + 2x^2 + 2x + 1, \quad r(x) = 2x + 2, \qquad \deg(r) < \deg(g).$$

Per semplicità, possiamo scrivere un polinomio in $\mathbb{F}_p[x]$ omettendo le x, gli esponenti ed i segni $+$. Ossia, possiamo usare la cosiddetta *notazione compatta* in alternativa a quella completa. Ad esempio: $x^3 + x^2 + 1$ in notazione compatta diventa 1101.

Esempio 2.7. Siano dati due polinomi in $\mathbb{F}_5[x]$:

$$\begin{aligned} f(x) = &\quad x^3 + 4x + 4 &&= 1044 \\ g(x) = &\quad 2x^2 + 3x + 1 &&= 231. \end{aligned}$$

Vediamo somma, differenza e prodotto di f e g in notazione compatta:

$f + g$:	$f - g$:	$f * g$:
1044	1044	1044
$+$ 231	$-$ 231	$*$ 231
1220	1313	1044
		3022
		2033
		234014

La divisione di f per g diventa:

	33 =	quoziente
	+ - - - - - - - -	
231	| 1044	
	| −1430	1^a cifra quoz. $= 3$ $(1/2 = 3 \bmod 5;\ 231 * 3 = 143)$
	+ - - - - - - - -	
	| 114	
	|− 143	2^a cifra quoz. $= 3$ $(1/2 = 3 \bmod 5;\ 231 * 3 = 143)$
	+ - - - - - - - -	
	| 21 =	resto

Il quoziente è 33 ed il resto è 21. Usando la notazione completa, questo significa che $x^3 + 4x + 4$ è uguale a $2x + 1 \bmod 2x^2 + 3x + 1$.

Il fatto che $\mathbb{F}_p[x]$ ammetta un algoritmo di divisione, implica che ogni ideale di $\mathbb{F}_p[x]$ è principale. Definiamo un polinomio monico come un polinomio che ha il coefficiente del monomio di grado massimo (coefficiente direttore) uguale a 1.

Teorema 2.3.2. *$\mathbb{F}_p[x]$ è un dominio ad ideali principali. Infatti, per ogni ideale $J \neq (0)$ di $\mathbb{F}_p[x]$ esiste un polinomio monico $g \in \mathbb{F}_p[x]$ univocamente determinato, con $J = (g)$.*

Dimostrazione. Supponiamo $J \neq (0)$. Sia $h(x)$ un polinomio non nullo di grado minimo contenuto in J, sia b il coefficiente direttore di $h(x)$ e poniamo $g(x) = b^{-1}h(x)$. Allora $g \in J$ e g è monico. Sia $f \in J$, l'algoritmo di divisione fornisce $q, r \in \mathbb{F}_p[x]$ con $f = qg + r$ e $\deg(r) < \deg(g) = \deg(h)$. Poiché J è un ideale, otteniamo $f - qg = r \in J$ e per definizione di h deve essere $r = 0$. Quindi, f è un multiplo di g e quindi $J = (g)$. Se $g_1 \in \mathbb{F}_p[x]$ è un altro polinomio monico con $J = (g_1)$, allora $g = c_1 g_1$ e $g_1 = c_2 g$ con $c_1, c_2 \in \mathbb{F}_p[x]$. Questo implica che $g = c_1 c_2 g_1$, quindi $c_1 c_2 = 1$ e c_1, c_2 sono polinomi costanti. Poiché sia g che g_1 sono monici, segue che $g = g_1$ e l'unicità di g è provata. □

Teorema 2.3.3. *Siano $f_1, \ldots, f_n$ polinomi in $\mathbb{F}_p[x]$ non tutti nulli. Allora esiste un polinomio monico $d \in \mathbb{F}_p[x]$ univocamente determinato tale che:*

(i) *d divide ogni f_j, $1 \leq j \leq n$;*

(ii) *ogni polinomio $c \in \mathbb{F}_p[x]$ che divide ognuno degli f_j, $1 \leq j \leq n$, divide d. Inoltre, d può essere espresso nella forma*

$$d = b_1 f_1 + \cdots + b_n f_n, \quad \text{con } b_1, \ldots, b_n \in \mathbb{F}_p[x]. \tag{2.3}$$

Dimostrazione. Sia J costituito da tutti polinomi della forma $c_1 f_1 + \cdots + c_n f_n$ con $c_1, \ldots, c_n \in \mathbb{F}_p[x]$: è facile vedere che è un ideale di $\mathbb{F}_p[x]$. Poiché gli f_j non sono tutti nulli, abbiamo $J \neq (0)$ ed il Teorema 2.3.2 implica che $J = (d)$ per qualche polinomio monico $d \in \mathbb{F}_p[x]$. La proprietà (i) e la rappresentazione (2.3) seguono immediatamente dalla costruzione di d. La

proprietà (ii) segue dalla (2.3). Se d_1 è un altro polinomio monico in $\mathbb{F}_p[x]$ soddisfacente (i) e (ii), allora queste proprietà implicano che d e d_1 si dividono l'un l'altro e quindi $(d) = (d_1)$. Applicando l'unicità fornita dal Teorema 2.3.2 si ottiene $d = d_1$. □

Il polinomio monico d si dice *massimo comun divisore* di $f_1, \ldots, f_n$, in simboli $d = \text{mcd}(f_1, \ldots, f_n)$. d può essere calcolato con l'algoritmo euclideo, che abbiamo visto nell'aritmetica di $\mathbb{F}_p$ paragrafo 2.1.1. Vediamo un esempio in $\mathbb{F}_3[x]$.

Esempio 2.8. L'algoritmo euclideo applicato a

$$f(x) = 2x^6 + x^3 + x^2 + 2, \qquad g(x) = x^4 + x^2 + 2x, \qquad f, g \in \mathbb{F}_3[x]$$

fornisce:

$$2x^6 + x^3 + x^2 + 2 = (2x^2 + 1)(x^4 + x^2 + 2x) + (x + 2)$$

$$x^4 + x^2 + 2x = (x^3 + x^2 + 2x + 1)(x + 2) + 1$$

$$x + 2 = (x + 2)1 + 0.$$

Quindi $\text{mcd}(f, g) = 1$ e f e g si dicono relativamente primi.

2.3.2 Costruzione di $\mathbb{F}_{p^n}$ attraverso un polinomio irriducibile di $\mathbb{F}_p[x]$

Definizione 2.8 (Polinomio irriducibile)**.**
Un polinomio $f \in F[x]$ si dice *irriducibile* su F se f ha grado positivo e $f = bc$, con $b, c \in F[x]$, implica che b o c sono polinomi costanti.

Sia $r(x)$ un polinomio monico irriducibile (con coefficiente direttore uguale a 1) con $\deg(r) \geq 1$ e consideriamo $\langle r(x) \rangle$ l'ideale da esso generato.
Così come (p) ideale di $\mathbb{Z}$ è l'insieme di tutti i multipli di p in $\mathbb{Z}$, $\langle r(x) \rangle$ è l'insieme di tutti i multipli di $r(x)$ in $\mathbb{F}_p[x]$. Consideriamo l'anello quoziente

$$\mathbb{F}_p[x]/\langle r(x) \rangle = \{[g]_r = g + \langle r(x) \rangle \mid g \in \mathbb{F}_p[x]\},$$

dove $[g]_r = [h]_r$ se $g \equiv h \bmod r$ e cioè se $g - h$ è divisibile per r. Questo equivale a richiedere che g ed h abbiano stesso resto r_1 nella divisione per r. Ogni classe contiene un unico rappresentante $r_1 \in \mathbb{F}_p[x]$ con $\deg(r_1) < \deg(r)$. Il passaggio da g ad r_1 si dice *riduzione mod* r. Quindi, gli elementi di $\mathbb{F}_p[x]/\langle r(x)\rangle$ possono essere scritti esplicitamente come $r_1 + \langle r(x)\rangle$.
Se $n = \deg(r)$, allora il numero di elementi di $\mathbb{F}_p[x]/\langle r(x)\rangle$ è uguale al numero di polinomi in $\mathbb{F}_p[x]$ di grado $< n$. Ci sono p possibilità per ogni coefficiente ed n coefficienti in totale, quindi in questo insieme ci sono p^n elementi e questo combacia con le considerazioni fatte in precedenza sugli spazi vettoriali.
Inoltre, poiché $r(x)$ è irriducibile ed i polinomi irriducibili su un campo F sono esattamente gli elementi primi di $F[x]$, segue che $\mathbb{F}_p[x]/\langle r(x)\rangle$ è un campo ([12, Teorema 1.61]). Allora scriveremo

$$\mathbb{F}_{p^n} = \mathbb{F}_p[x]/\langle r(x)\rangle .$$

Vedremo di seguito che esiste almeno un polinomio monico irriducibile per ogni p ed n e se ne esiste più di uno, come già detto, tutti questi campi sono isomorfi tra loro.
Il passo successivo sarà trovare degli $r(x)$ irriducibili e a quel punto avremo tutto ciò di cui abbiamo bisogno per prendere i resti nella costruzione di $\mathbb{F}_p[x]/\langle r(x)\rangle$. Sia ora $u = x + \langle r(x)\rangle$, dove x è un elemento di $\mathbb{F}_p[x]$: questi sono tutti i polinomi che appartengono alla stessa classe d'equivalenza di x. Allora scriveremo anche

$$\mathbb{F}_p[x]/\langle r(x)\rangle \cong \mathbb{F}_p(u)$$

ossia aggiungiamo al campo $\mathbb{F}_p$ una radice u di $r(x)$, dove u è un elemento del campo $\mathbb{F}_{p^n}$.

2.3.3 Aritmetica di $\mathbb{F}_{p^n}$

Addizione, sottrazione e moltiplicazione in $\mathbb{F}_{p^n}$

Addizione e sottrazione di elementi di $\mathbb{F}_{p^n}$ sono facili: basta farle componente per componente come in $\mathbb{F}_p[x]$.

Per la moltiplicazione ci sono due metodi. Il primo lo chiamiamo *metodo dell'anello-riduzione* (AR), il secondo lo chiamiamo *metodo della riduzione lungo il cammino* (RC).

1. **Metodo AR**
   Basta combinare sollevamento, moltiplicazione in un anello e riduzione mod $r(x)$ come descritti in precedenza (per calcolare $3 \cdot 4$ in $\mathbb{F}_p$, sollevavamo 3 e 4 da $\mathbb{F}_p$ a $\mathbb{Z}$, prendevamo $3 \cdot 4 = 12$ nell'anello $\mathbb{Z}$, poi dividevamo per 5 ottenendo resto 2 di nuovo in $\mathbb{F}_p$). Vediamo come procedere con un esempio.

   Esempio 2.9.
   Calcoliamo $(u^3+1) \cdot (u^2+u+1)$ con $p = 2$, $r(x) = x^4 + x + 1$.
   Per prima cosa solleviamo da $\mathbb{F}_{2^4}$ a $\mathbb{F}_2[x]$ ottenendo $(x^3+1)\cdot(x^2+x+1)$, poi moltiplichiamo in $\mathbb{F}_2[x]$. La moltiplicazione ci restituisce il prodotto non ridotto, quindi dividiamo per $r(x)$ e prendiamo il resto, scartando il quoziente:

$$\begin{array}{r} 1001 \\ * \quad 0111 \\ \hline 1001 \\ 10010 \\ 100100 \\ \hline 111111 \end{array} \quad \longrightarrow \quad \begin{array}{rl|r} & & 11 \\ & + - - - - - & \\ 10011 & \quad 111111 & \\ & \quad - 100110 & \\ & + - - - - - & \\ & \quad 11001 & \\ & \quad - 10011 & \\ & + - - - - - & \\ & \quad 1010 & \end{array} \quad \begin{array}{l} = \text{quoziente} \\ \\ \\ \\ \\ \\ \\ = \text{resto} \end{array}$$

   Quindi $1001 \cdot 0111 = 1010$, che in notazione completa corrisponde a $(u^3+1) \cdot (u^2+u+1) = u^3 + u$.

2. **Metodo RC**
   Il metodo RC è molto simile al metodo AR, ma è più comodo perché semplifica i calcoli. Speghiamo il metodo RC attraverso l'esempio 2.9.

Abbiamo $u^4 + u + 1 = 0$, da cui $u^4 = -u - 1$, che in $\mathbb{F}_{2^4}$ diventa $u^4 = u + 1$.

$$
\begin{array}{rr}
 & 1001 \\
* & 0111 \\
\hline
 & 1001 \\
 & 10010 \\
 & 100100 \\
\hline
 & 111111 = 101100
\end{array}
\quad \longrightarrow \quad
\begin{array}{rl|rl}
 & & & 10 \quad = \text{quoziente} \\
\hline
10011 & & + & \\
 & 101100 & & \\
 & - 100110 & & \\
\hline
 & 001010 & & = \text{resto}
\end{array}
$$

Nel metodo AR, poiché i gradi in input sono inferiori o uguali ad $n-1$, il grado del prodotto non ridotto può essere grande fino a $2n-2$.
Nel metodo RC, quando ci imbattiamo in un termine di grado n, lo riduciamo sostituendo ad esempio u^4 con $u+1$.

I metodi di inversione in $\mathbb{F}_{p^n}$

Come in $\mathbb{F}_p$ il primo metodo di inversione consiste nel cercare un reciproco andando per tentativi, ma è conveniente solo se p^n è piccolo; il secondo metodo sfrutta l'algoritmo euclideo ed il terzo metodo si basa sul fatto che ogni elemento a di $\mathbb{F}_{p^n}$ verifica $a^{p^n-2} = a^{-1}$ (è la regola di inversione $p^n - 2$). Nel quarto metodo di inversione si calcola un elemento primitivo, poi si usano i logaritmi.

Esempio 2.10. Sia $r(x) = x^4 + x + 1$ con $p = 2$, $n = 4$.
Usando $g = u$, prendiamo le potenze di u applicando il metodo RC ed ordiniamo gli elementi del campo per ottenere la tabella dei logaritmi. Infine ordiniamo in base al valore dei logaritmi per ottenere la tabella degli antilogaritmi.

$\log_g(a) = k$	$a = g^k$
15	0001
1	0010
4	0011
2	0100
8	0101
5	0110
10	0111
3	1000
14	1001
9	1010
7	1011
6	1100
13	1101
11	1110
12	1111

$\log_g(a) = k$	$a = g^k$
1	0010
2	0100
3	1000
4	0011
5	0110
6	1100
7	1011
8	0101
9	1010
10	0111
11	1110
12	1111
13	1101
14	1001
15	0001

Tabella 2.1: Tabella dei logaritmi e tabella degli antilogaritmi

Ora è facile invertire semplicemente usando le tabelle. Ad esempio, usando la prima tabella $0011 = 0010^4$, quindi $1/0011 = 0010^{15-4} = 0010^{11} = 1110$ usando la seconda tabella.

2.3.4 Potenze di u ed LFSR (Linear feedback shift register)

Prendendo le potenze di u facciamo "brillare" il metodo della riduzione lungo il cammino (RC) per la moltiplicazione (si veda paragrafo 2.3.3, metodo 2). Possiamo procedere in questo modo:

- Iniziare con u.
- Far scorrere verso sinistra di una posizione tutti i coefficienti della notazione compatta. Questo equivale a moltiplicare per u.

- Se c'è un coefficiente diverso da zero nella posizione u^n, usare $r(u)$ per eliminarlo.

Esempio 2.11. Prendiamo $r(x) = x^4 + x + 1$ (10011 in notazione compatta) con $p = 2, n = 4$, $u^4 + u + 1 = 0$. Possiamo risolvere per u^4 ottenendo $u^4 = u + 1$. Così, ogni volta che vediamo 10000, possiamo eliminarlo sostituendolo con 0011.

k	u^k	
1		0010
2		0100
3		1000
4	10000 =	0011
5		0110
6		1100
7	11000 =	1011
8	10110 =	0101
9		1010
10	10100 =	0111
11		1110
12	11100 =	1111
13	11110 =	1101
14	11010 =	1001
15	10010 =	0001

Quando $p = 2$, indipendentemente da n, l'addizione corrisponde all'operazione XOR (exclusive or). Questo "scorrimento a sinistra ed incorporamento" può essere efficientemente implementato in un circuito elettronico, che viene chiamato in modo abbastanza appropriato, *registro a scorrimento a retroazione lineare* o LFSR, linear feedback shift register (si veda il capitolo 3). Questi circuiti sono importanti nella crittografia e nella teoria dei codici.

Con i dati iniziali impostati a 1, chiaramente il numero di iterazioni prima che la sequenza si ripeta sarà massimo quando u è un elemento primitivo, cioè quando $r(x)$ è primitivo, poiché uno scorrimento a sinistra equivale a moltiplicare per u. Da qui il detto: "i polinomi primitivi danno luogo al massimo periodo dei registri a scorrimento a retroazione lineare" che vedremo in modo approfondito nel capitolo 3.

2.3.5 Fattorizzazione in $\mathbb{F}_{p^n}$

Siano p un primo ed $n = [\mathbb{F}_{p^n} : \mathbb{F}_p]$. $x^{p^n} - x$ si dice *polinomio universale*: questo è riducibile in $\mathbb{F}_p$ e si fattorizza in monici irriducibili di grado m per ogni m che divide n.

Esempio 2.12. Abbiamo già elencato tutti i monici irriducibili per $p = 2$ ed $n = 1, \ldots, 4$. I divisori di 4 sono ovviamente $1, 2, 4$. È facile vedere che $x^{16} - x$ ha i seguenti sei fattori in $\mathbb{F}_2[x]$:

$$\begin{aligned} m = 1: &\quad x, && x+1 \\ m = 2: &\quad x^2 + x + 1 \\ m = 4: &\quad x^4 + x + 1, && x^4 + x^3 + 1, \quad x^4 + x^3 + x^2 + x + 1. \end{aligned}$$

Vediamo cosa comporta il fatto che $\mathbb{F}_{p^n}$ sia un campo di spezzamento di un polinomio con coefficienti in $\mathbb{F}_p$, ossia che $\mathbb{F}_{p^n}$ sia un'estensione di campo di $\mathbb{F}_p$. È proprio qui che serve aver preso in precedenza, x ed u distinti: vediamo perché, prima con un'analogia e poi con un esempio.
Analogia: il polinomio $x^2 + 1$ è irriducibile su $\mathbb{R}$, ma si spezza su $\mathbb{C}$ come prodotto di due fattori lineari: $x^2 + 1 = (x + i)(x - i)$. $\mathbb{C}$ è un'estensione di $\mathbb{R}$ ottenuta aggiungendo i radice del polinomio $x^2 + 1$ (infatti $i^2 = -1$), ossia $\mathbb{C} = \mathbb{R}(i) = \mathbb{R}[x]/\langle x^2 + 1 \rangle$. Analogamente per i campi finiti.

Esempio 2.13. Abbiamo visto che $x^2 + x + 1$ è irriducibile su $\mathbb{F}_2$. Quindi, aggiungendo ad $\mathbb{F}_2$ un elemento u tale che $u^2+u+1 = 0$, si ottiene l'estensione di campo $\mathbb{F}_{2^2} = \mathbb{F}_2(u) = \mathbb{F}_2[x]/\langle x^2 + x + 1 \rangle$ campo di spezzamento di $x^2 + x + 1$. Verifichiamo che $(x + 10)(x + 11)$ (usando la notazione compatta), o

$(x+u)(x+u+1)$ (usando la notazione completa), è una fattorizzazione in $\mathbb{F}_{2^2}$.

$$\begin{aligned}(x+u)(x+u+1) &= x^2+ux+x+ux+u^2+u\\ &= x^2+x+u^2+u\\ &= x^2+x+u+1+u\\ &= x^2+x+1.\end{aligned}$$

La tabella delle radici

La fattorizzazione di $x^{p^n}-x$ in irriducibili su $\mathbb{F}_p$ è unica, ma i dettagli della fattorizzazione in termini lineari su $\mathbb{F}_{p^n}$ chiaramente dipenderanno dalla scelta di $r(x)$ per definire l'aritmetica di $\mathbb{F}_{p^n}$.
Abbiamo visto che per $p=2$, $n=4$ ci sono tre possibilità per $r(x)$. Siano

$$\begin{aligned}r_1(x) &= x^4+x+1 &&= 10011\\ r_2(x) &= x^4+x^3+1 &&= 11001\\ r_3(x) &= x^4+x^3+x^2+x+1 &&= 11111\end{aligned}$$

i tre monici irriducibili per $\mathbb{F}_{2^4}$.
La fattorizzazione di $x^{16}-x$ in termini lineari su $\mathbb{F}_{2^4}$ usando r_1, r_2, r_3 per definire l'aritmetica del campo finito, avviene come segue. Poiché stiamo fattorizzando in termini lineari su un campo di spezzamento, tutti i fattori sono della forma $(x-a)$ per qualche a. Per guadagnare spazio scriveremo soltanto le radici (le a).

Si rimanda il lettore all'appendice A, per la descrizione dei metodi per la costruzione di una simile tabella. Nota: Nella notazione compatta 00010 significa nella prima colonna $x \in \mathbb{F}_2[x]$ e nelle colonne successive $u \in \mathbb{F}_{2^4}$.
Inoltre, nella prima colonna, dove vengono riportati i monici irriducibili di grado m che divide 4, sono stati marcati gli r_i, poiché questi sono gli stessi tre polinomi che definiscono l'aritmetica di $\mathbb{F}_{2^4}$. In altri termini, stiamo fattorizzando $r_i \bmod r_j$.

Irrid. su $\mathbb{F}_2$	Radici mod $r_1 = 10011$	Radici mod $r_2 = 11001$	Radici mod $r_3 = 11111$
00010	0000	0000	0000
00011	0001	0001	0001
00111	0110 0111	1011 1010	1101 1100
r_1	0010 0100 0011 0101	0111 1100 0110 1101	0111 1010 0110 1011
r_2	1011 1001 1101 1110	0010 0100 1001 1110	0011 0101 1110 1001
r_3	1000 1100 1111 1010	1000 1111 0011 0101	1111 1000 0010 0100

Tabella 2.2: Tabella delle radici per $x^{16} - x$ su $\mathbb{F}_{2^4}$

Criterio del sottocampo attraverso i logaritmi

Tornando ai campi finiti, se $\mathbb{F}_{p^n}^{\times}$ ha generatore g, allora $\mathbb{F}_{p^m}^{\times}$ ha generatore $g^{(p^n-1)/(p^m-1)}$.

Nella tabella delle radici 2.2 possiamo vedere i sottocampi. Consideriamo gli elementi 0 e 1 insieme alle radici di $r_v = x^2 + x + 1$ (riga 3 della tabella 2.2). Operando mod $x^4 + x + 1$, cioè mod r_1 (colonna 2 della tabella 2.2), questi quattro elementi sono 0, 1, $u^2 + u$, $u^2 + u + 1$. Usando addizione, sottrazione, moltiplicazione e divisione possiamo vedere che questi quattro elementi soddisfano tutti gli assiomi di un campo e quindi deve essere un campo: lo chiamiamo prima copia di $\mathbb{F}_{2^2}$.

Facendo riferimento alla tabella degli antilogaritmi 2.1, notiamo che $1, u^2 + u, u^2+u+1$ hanno logaritmi $5, 10, 15$ rispetto al generatore u; ossia gli elementi della prima copia di $\mathbb{F}_{2^2}$ dentro $\mathbb{F}_{2^4}$ sono 0, u^5, u^{10} e $u^{15} = 1$. Questo è il *criterio del sottocampo attraverso i logaritmi*: elementi di un sottocampo possono essere determinati dai loro logaritmi. In particolare un elemento di $\mathbb{F}_{p^n}$ è anche un elemento di $\mathbb{F}_{p^m}$ dove $m|n$ se e solo se è 0 o il suo logaritmo è un multiplo di $(p^n - 1)/(p^m - 1)$. Infatti, un elemento di $\mathbb{F}_{2^4}$ è anche un elemento della prima copia di $\mathbb{F}_{2^2}$ se e solo se è 0 o il suo logaritmo è un multiplo di $5 = (2^4 - 1)/(2^2 - 1)$.

Ora, abbiamo visto che per $p = 2$ ed $n = 2$, l'unico monico irriducibile è

$r_v = x^2 + x + 1$. Questo porta ad una seconda copia di $\mathbb{F}_{2^2}$ con elementi 0, 1, v, $v+1$ dove abbiamo usato v anziché u per sottolineare che queste sono classi d'equivalenza differenti: sono classi mod r_v.
Poiché queste due copie di $\mathbb{F}_{2^2}$ sono campi ed hanno lo stesso ordine, devono essere isomorfi. L'isomorfismo tra i due campi si chiama automorfismo di Frobenius, che vedremo al paragrafo 2.4.4.

Algoritmo di Berlekamp

Vediamo ora un algoritmo per la fattorizzazione di polinomi su un campo finito ideato da Elwyn Berlekamp nel 1967 (si veda [3, §6] o in alternativa [12, §4]; per informazioni aggiuntive si veda [7, p.154]).
Ogni polinomio f in $\mathbb{F}_p[x]$ di grado positivo ha una fattorizzazione canonica in $\mathbb{F}_p[x]$. Il nostro scopo è quello di esprimere un polinomio monico $f \in \mathbb{F}_p[x]$ di grado positivo nella forma

$$f = f_1^{e_1} \cdots f_k^{e_k},$$

dove $f_1, \ldots, f_k$ sono polinomi monici irriducibili in $\mathbb{F}_p[x]$ e gli $e_1, \ldots, e_k$ sono interi positivi.
Per prima cosa semplifichiamo il nostro obiettivo nella fattorizzazione di un polinomio che non ha fattori multipli, ossia una fattorizzazione in cui tutti gli $e_1, \ldots, e_k$ sono uguali a 1 (o equivalentemente consideriamo un polinomio che non ha radici multiple). A questo scopo, calcoliamo

$$d(x) = \mathrm{mcd}(f(x), f'(x)),$$

il massimo comun divisore tra $f(x)$ e la sua derivata, applicando l'algoritmo euclideo come visto nell'esempio 2.8.
Se $d(x) = 1$, allora sappiamo che $f(x)$ non ha fattori multipli per il [12, Teorema 1.68].
Se $d(x) = f(x)$, dobbiamo avere $f'(x) = 0$.
Quindi, $f(x) = g(x)^p$, dove $g(x)$ è un polinomio opportuno di $\mathbb{F}_p[x]$ e p è la caratteristica di $\mathbb{F}_p$. Se necessario, il procedimento di riduzione può essere

continuato applicando il metodo a $g(x)$.
Se $d(x) \neq 1$ e $d(x) \neq f(x)$, allora $d(x)$ è un fattore non banale di $f(x)$ e $f(x)/d(x)$ non ha fattori multipli per l'ipotesi fatta su f.
La fattorizzazione di $f(x)$ è ottenuta fattorizzando $d(x)$ e $f(x)/d(x)$ separatamente. Nel caso in cui $d(x)$ avesse ancora fattori multipli, dovremmo continuare con ulteriori applicazioni del procedimento di riduzione.
La fattorizzazione canonica di questi polinomi conduce direttamente alla fattorizzazione canonica del polinomio originale. Quindi, restringiamo la nostra attenzione ai polinomi senza fattori multipli. Il teorema seguente è cruciale.

Teorema 2.3.4. *Se $f \in \mathbb{F}_p[x]$ è monico e $h \in \mathbb{F}_p[x]$ è tale che $h^p \equiv h \bmod f$, allora*

$$f(x) = \prod_{c \in \mathbb{F}_p} \mathrm{mcd}(f(x), h(x) - c). \tag{2.4}$$

Dimostrazione. Ogni $\mathrm{mcd}(f(x), h(x) - c)$ divide $f(x)$. Poiché i polinomi $h(x) - c$, $c \in \mathbb{F}_p$ sono a due a due relativamente primi, allora sono i massimi comun divisori con $f(x)$ e quindi il prodotto di questi divide $f(x)$. D'altra parte, $f(x)$ divide

$$h(x)^p - h(x) = \prod_{c \in \mathbb{F}_p} (h(x) - c),$$

e quindi $f(x)$ divide il membro destro della (2.4). Quindi i due membri della (2.4) sono polinomi monici che si dividono l'un l'altro e quindi devono essere uguali. □

Sia h come nel Teorema 2.3.4. Se h fornisce una fattorizzazione non banale di f, diciamo che h è un polinomio *f-riducente*. Ogni h con $h^p \equiv h \bmod f$ e $0 < \deg(h) < \deg(f)$ è ovviamente f-riducente. Per ottenere degli algoritmi di fattorizzazione sulla base del Teorema 2.3.4, dobbiamo trovare dei metodi di costruzione di polinomi f-riducenti. Poiché la fattorizzazione dipende dal calcolo di p massimi comun divisori, un'applicazione diretta di questa formula sarà a basso costo computazionale per campi finiti con p piccolo.
Il primo metodo per la costruzione di polinomi f-riducenti fa uso del Teorema

cinese del resto per polinomi (si veda ad esempio il [22]). Sia f un polinomio senza fattori multipli. Sia $(c_1, \ldots, c_k)$ una k-upla arbitraria di elementi di $\mathbb{F}_p$. Il Teorema cinese del resto implica che esiste un unico $h \in \mathbb{F}_p[x]$ con $h(x) = c_i \bmod f_i(x)$ per $1 \leq i \leq k$ e $\deg(h) < \deg(f)$. Il polinomio $h(x)$ soddisfa la condizione

$$h(x)^p \equiv c_i^p = c_i \equiv h(x) \bmod f_i(x) \quad \text{per} 1 \leq i \leq k,$$

e quindi

$$h^p \equiv h \bmod f, \quad \deg(h) < \deg(f). \tag{2.5}$$

D'altra parte, se h è una soluzione della (2.5), allora l'identità

$$h(x)^p - h(x) = \prod_{c \in \mathbb{F}_p} (h(x) - c)$$

implica che ogni fattore irriducibile di f divide uno dei polinomi $h(x) - c$. Quindi, tutte le soluzioni della (2.5) soddisfano $h(x) \equiv c_i \bmod f_i(x)$, $1 \leq i \leq k$, per qualche k-upla $c_1, \ldots, c_k$ di elementi di $\mathbb{F}_p$. Di conseguenza, ci sono esattamente p^k soluzioni della (2.5).
Troviamo queste soluzioni riducendo la (2.5) in un sistema di equazioni lineari. Con $n = \deg(f)$ costruiamo la matrice quadrata $B = (b_{ij})$, $0 \leq i, j \leq n-1$, calcolando le potenze $x^{ip} \bmod f(x)$. Nello specifico, sia

$$x^{ip} = \sum_{j=0}^{n-1} b_{ij} x^j \bmod f(x) \quad \text{per } 0 \leq i \leq n-1.$$

Allora $h(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \in \mathbb{F}_p[x]$ è una soluzione di (2.5) se e solo se

$$(a_0, a_1, \ldots, a_{n-1})B = (a_0, a_1, \ldots, a_{n-1}). \tag{2.6}$$

Questo segue dal fatto che (2.6) vale se e solo se

$$h(x) = \sum_{j=0}^{n-1} a_j x^j$$

$$= \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} a_j b_{ij} x^j$$

$$\equiv \sum_{i=0}^{n-1} a_i x^{ip} = h(x)^p \bmod f(x).$$

Il sistema (2.6) può essere scritto nella forma equivalente

$$(a_0, a_1, \ldots, a_{n-1})(B - I) = (0, \ldots, 0), \tag{2.7}$$

dove I è la matrice identità $n \times n$ su $\mathbb{F}_p$. Il sistema (2.7) ha p^k soluzioni. Quindi, la dimensione dello spazio nullo della matrice $B - I$ è k, il numero di fattori monici irriducibili distinti di f ed il rango di $B - I$ è $n - k$.
Poiché il polinomio costante $h_1(x) = 1$ è sempre una soluzione di (2.5), il vettore $(1, 0, \ldots, 0)$ è sempre una soluzione di (2.7) (lo si può verificare direttamente). Esisteranno dei polinomi $h_2(x), \ldots, h_k(x)$ di grado $\leq n - 1$ tali che $h_1(x), h_2(x), \ldots, h_k(x)$ formano una base per lo spazio nullo di $B - I$. I polinomi $h_2(x), \ldots, h_k(x)$ hanno grado positivo e quindi sono f-riducenti.
Abbiamo $r = n - k$: noto il rango r, sappiamo anche che **il numero di fattori monici irriducibili distinti di f è dato da** $n - r$. Sulla base di questa informazione possiamo decidere quando interrompere la procedura di fattorizzazione. Per determinare r è consigliabile usare soltanto operazioni colonna, poiché lasciano lo spazio nullo invariante; una volta determinato r, abbiamo $k = n - r$.
Se $k = 1$, sappiamo che f è irriducibile su $\mathbb{F}_p$ e ci fermiamo. In questo caso, le uniche soluzioni di (2.5) sono i polinomi costanti e lo spazio nullo di $B - I$ contiene solo i vettori della forma $(c, 0, \ldots, 0)$ con $c \in \mathbb{F}_p$.
Se $k \geq 2$, prendiamo $h_2(x)$ la base polinomiale f-riducente e calcoliamo $\mathrm{mcd}(f(x), h_2(x) - c)$ per ogni $c \in \mathbb{F}_p$.
Il risultato sarà una fattorizzazione non banale di $f(x)$ fornita dalla (2.4). Se l'uso di $h_2(x)$ non spezza $f(x)$ in k fattori, calcoliamo $\mathrm{mcd}(g(x), h_3(x) - c)$ per ogni $c \in \mathbb{F}_p$ e tutti i fattori non banali $g(x)$ trovati finora. Si procede in questo modo finché non si ottengono k fattori di $f(x)$.

Esempio 2.14. Usando l'algoritmo di Berlekamp, fattorizziamo il seguente polinomio:

$$f(x) = x^8 + x^6 + x^4 + x^3 + 1 \quad \text{su } \mathbb{F}_2.$$

Poiché $\text{mcd}(f(x), f'(x)) = 1$, $f(x)$ non ha fattori multipli. Dobbiamo calcolare x^{2i} mod $f(x)$ con $0 \leq i \leq 7$. Questo fornisce le seguenti congruenze mod $f(x)$:

$$\begin{array}{lccccccc}
x^0 \equiv 1 & & & & & & & \\
x^2 \equiv & & x^2 & & & & & \\
x^4 \equiv & & & & x^4 & & & \\
x^6 \equiv & & & & & & x^6 & \\
x^8 \equiv 1 & & & +x^3 & +x^4 & & +x^6 & \\
x^{10} \equiv 1 & & +x^2 & +x^3 & +x^4 & +x^5 & & \\
x^{12} \equiv & & x^2 & & +x^4 & +x^5 & +x^6 & +x^7 \\
x^{14} \equiv 1 & +x & & +x^3 & +x^4 & +x^5 & &
\end{array}$$

Quindi, la matrice B 8×8 è data da

$$B = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0
\end{pmatrix}$$

e $B - I$ è data da

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

La matrice $B-I$ ha rango 6 e i due vettori $(1,0,0,0,0,0,0,0)$ e $(0,1,1,0,0,1,1,1)$ formano una base dello spazio nullo di $B-I$. I corrispondenti polinomi sono $h_1(x) = 1$ e $h_2(x) = x + x^2 + x^5 + x^6 + x^7$. Calcoliamo $\mathrm{mcd}(f(x), h_2(x) - c)$ per $c \in \mathbb{F}_2$ con l'algoritmo euclideo ed otteniamo $\mathrm{mcd}(f(x), h_2(x)) = x^6 + x^5 + x^4 + x + 1$, $\mathrm{mcd}(f(x), h_2(x) - 1) = x^2 + x + 1$. La fattorizzazione canonica desiserata è quindi

$$f(x) = (x^6 + x^5 + x^4 + x + 1)(x^2 + x + 1).$$

2.4 Polinomi irriducibili

Vogliamo fornire una formula per il numero di polinomi irriducibili su un campo finito. Sia μ la funzione di Möbius vista nel paragrafo 2.2.3.

2.4.1 Formula per il numero di polinomi irriducibili

Teorema 2.4.1. *Sia F un campo finito con q elementi e sia n un intero positivo. Allora il numero di polinomi monici irriducibili di grado n in $F[x]$ è*

$$\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}.$$

La sommatoria è presa su tutti i divisori positivi d di n.

Dimostrazione. Il teorema segue dalla formula di inversione di Möbius e dalla scomposizione in fattori irriducibili del polinomio $x^{q^n} - x$ vista in 2.2.5. Sia $f(n)$ il numero di polinomi monici irriducibili di grado n in $F[x]$. Nella scomposizione del polinomio $x^{q^n} - x \in F[x]$ si hanno $f(d)$ fattori irriducibili di grado d per ciascun divisore d di n, per cui, prendendo i gradi,

$$q^n = \sum_{d|n} df(d).$$

Dal Teorema di inversione di Möbius si ricava allora

$$nf(n) = \sum_{d|n} \mu(d) q^{n/d},$$

che è la tesi. □

Se n è un primo, allora i soli divisori positivi di n sono 1 ed n, per cui il numero di polinomi monici irriducibili di grado n in $F[x]$ è

$$\frac{1}{n}(\mu(1)q^n + \mu(n)q^{n/n}) = \frac{(q^n - q)}{n}.$$

Se $n = p^r$, con p primo, allora i soli divisori positivi di n sono $1, p, p^2, \ldots, p^r$ e perciò il numero di polinomi monici irriducibili di grado n in $F[x]$ è

$$\frac{1}{n}(\mu(1)q^n + \mu(p)q^{n/p}) = \frac{(q^{p^r} - q^{p^{r-1}})}{p^r}.$$

Se $n = rs$, dove r ed s sono primi distinti, allora il numero di polinomi irriducibili di grado n in $F[x]$ è

$$\frac{1}{n}(\mu(1)q^n + \mu(r)q^{n/r} + \mu(s)q^{n/s} + \mu(rs)q^{n/rs}) = \frac{(q^{rs} - q^s - q^r + 1)}{n}.$$

Esistenza di polinomi irriducibili di grado arbitrario

Segue dalla formula per il numero di polinomi irriducibili su un campo finito.

Corollario 2.4.2. *Sia F un campo finito. Per ciascun intero positivo n, esiste almeno un polinomio irriducibile di grado n in $F[x]$.*

Dimostrazione. Poniamo

$$\partial(n) = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}.$$

Basta dimostrare che $\partial(n) \neq 0, \forall n \geq 1$. Ma $n\partial(n)$ è una somma di potenze di q, tutte distinte, con coefficienti 1 o -1. Se r è la più piccola potenza di q che appare in questa somma, possiamo scrivere $n\partial(n)$ come la somma di $\pm q^r$ con altri termini del tipo $\pm q^t, t > r$. Raccogliendo q^r, risulta che $\partial(n)$ è un prodotto del tipo $q^r(\pm 1 + m)$, dove m è divisibile per q. Ovviamente $\pm 1 + m \neq 0$, perché altrimenti ± 1 sarebbe divisibile per q. Ne segue che $\partial(n) \neq 0$. □

Prima di dimostrare il Teorema 2.4.1, diamo un'altra applicazione della formula di Möbius: ricaviamo un'espressione per la funzione φ di Eulero. Ricordiamo che se $n > 1$ è un intero, allora $\varphi(n)$ è il numero di interi k con $1 \leq k \leq n$, che sono relativamente primi ad n.
Se $d = \text{mcd}(k, n)$ e poniamo $e = k/d$, allora $1 \leq e \leq n/d$.
Inoltre $\text{mcd}(e, n/d) = 1$: se l è un intero positivo con $l|e$, $l|(n/d)$, allora $dl|de = k$, $dl|n$ e quindi $dl \leq d$, ossia $l = 1$.
Viceversa, se d è un divisore di n ed e è un intero tale che $1 \leq e \leq n/d$ e $\text{mcd}(e, n/d) = 1$, allora $1 \leq de \leq n$, $\text{mcd}(de, n) = d$.
Perciò se d è un divisore di n, il numero di interi k con $1 \leq k \leq n$, $\text{mcd}(k, n) = d$ è uguale al numero di interi e con $1 \leq e \leq n/d$, $\text{mcd}(e, n/d) = 1$, ossia $\varphi(n/d)$. Ne deduciamo che

$$n = \sum_{d|n} \varphi\left(\frac{n}{d}\right) = \sum_{d|n} \varphi(d).$$

Applicando la formula di inversione di Möbius 2.2 otteniamo la formula seguente.

Proposizione 2.4.3. *Per ogni $n > 1$ si ha*

$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.$$

Vediamo un altro lemma preliminare alla dimostrazione del Teorema 2.4.1.

Lemma 2.4.4. *Se F è un campo, $q > 1$ un intero, m ed n interi positivi, allora $x^{q^m} - x$ divide $x^{q^n} - x$ in $F[x]$ se e solo se m divide n.*

Dimostrazione. Abbiamo $x^{q^m} - x = x(x^{q^m-1} - 1)$ e $x^{q^n} - x = x(x^{q^n-1} - 1)$, perciò $x^{q^m} - x$ divide $x^{q^n} - x$ se e solo se $x^{q^m-1} - 1$ divide $x^{q^n-1} - 1$. I lemmi 2.2.10 e 2.2.11 concludono la dimostrazione. □

2.4.2 Metodi semplici per la determinazione di irriducibili in $\mathbb{F}_p[x]$

Vediamo quindi dei metodi semplici per provare l'irriducibilità.

1. **Metodo della radice.**
   Se un polinomio ha grado inferiore o uguale a 3 e si fattorizza in termini di grado più piccolo, allora almeno uno dei fattori deve essere lineare. Sappiamo che i fattori lineari corrispondono alle radici: se a è una radice di $r(x)$, allora $r(x)$ ha $x - a$ come fattore e viceversa. È facile verificare la presenza di radici in $\mathbb{F}_p$, poiché $\mathbb{F}_p$ è finito: basta provare tutti gli elementi. Questo ci fornisce un metodo per provare l'irriducibilità per $n \leq 3$: calcoliamo $r(0), r(1), \ldots, r(p-1)$ e se questi sono tutti diversi da zero, $r(x)$ è irriducibile.
   Il metodo della radice può essere usato anche per polinomi di grado più alto di 3 per poter stabilire velocemente che un polinomio non è irriducibile, cioè se il polinomio ha una radice. Quando $p = 2$, questo è particolarmente facile: $r(0)$ non è altro che il termine costante, così se il termine costante è 0, $r(x)$ è riducibile (eccetto $r(x) = x$). Inoltre, $r(1)$ non è altro che la somma dei coefficienti mod 2, quindi se il numero dei coefficienti non nulli è pari, $r(x)$ è riducibile (eccetto $r(x) = x + 1$).

2. **Metodo della divisione**
   Il secondo metodo semplice per verificare l'irriducibilità, dopo aver controllato i fattori lineari tramite il metodo delle radici è il metodo della

divisione. Poiché in $\mathbb{F}_p[x]$ esiste un numero finito di polinomi di grado inferiore a quello di $r(x)$, basta provarli tutti. Ovviamente non prendiamo in considerazione i fattori che sappiamo già essere riducibili. Inoltre, proviamo solo i fattori con grado fino a $\lfloor n/2 \rfloor$, perché se ci fossero fattori di grado più alto avremmo già trovato il corrispondente fattore di grado più basso.

Ora applichiamo i metodi visti per elencare tutti i polinomi irriducibili di grado inferiore o uguale a 4 con $p = 2$.
Per $n = 1$, abbiamo x ed $x + 1$ (10 e 11 in notazione compatta) che sono già lineari e quindi irriducibili.
Per $n = 2$, abbiamo 100, 101, 110 e 111. Due di questi hanno il termine costante nullo e quindi hanno 0 come radice; 101 ha un numero pari di 1 e così ha 1 come radice. Rimane 111 ($x^2 + x + 1$) che è quadratico e privo di radici, quindi irriducibile.
Per $n = 3$, ci sono otto possibilità

$$1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111$$

Tutti quelli che terminano con lo 0 hanno 0 come radice; 1001, 1010, 1100 e 1111 hanno 1 come radice. Restano 1011 e 1101 ($x^3 + x + 1$ e $x^3 + x^2 + 1$) che sono irriducibili poiché quadratici e privi di radici.
Per $n = 4$, ci sono sedici possibilità

$$10000, 10001, 10010, 10011, 10100, 10101, 10110, 10111,$$

$$11000, 11001, 11010, 11011, 11100, 11101, 11110, 11111$$

Il metodo della radice riduce la lista alla seguente

$$10011, 10101, 11001, 11111$$

Poiché il metodo della radice esclude tutti i fattori lineari, rimangono solo i fattori quadratici e possiamo limitarci ai quadratici irriducibili. Ce n'è uno solo e precisamente 111. La divisione sintetica come sopra mostra che

$10101 = 111^2$; gli altri tre hanno resto diverso da zero e quindi devono essere irriducibili. Riassumiamo i risultati ottenuti nella seguente tabella.

n	Notazione compatta	Notazione completa
1	10	x
	11	$x+1$
2	111	x^2+x+1
3	1011	x^3+x+1
	1101	x^3+x^2+1
4	10011	x^4+x+1
	11001	x^4+x^3+1
	11111	$x^4+x^3+x^2+x+1$

Tabella 2.3: Polinomi monici irriducibili con $p = 2$, $n = 1, \ldots, 4$

Si possono trovare delle utili tabelle, in cui vengono elencati i polinomi irriducibili con $p = 2, 3, 5, 7$ fino ad un certo n, in [12, TABLE C].

2.4.3 Altri metodi per la determinazione di irriducibili in $\mathbb{F}_p[x]$

Finora abbiamo visto che:

- Per polinomi di grado inferiore o uguale a tre, usiamo il test della radice, poiché se esiste una fattorizzazione non banale, almeno un fattore deve essere lineare.

- Per polinomi di grado superiore o uguale a quattro, prima usiamo il test della radice per vedere se ci sono fattori lineari, poi usiamo la prova della divisione per fattori di grado da 2 fino a $\lfloor n/2 \rfloor$.

Vediamo altre possibilità.

Algoritmo del polinomio universale

Esiste un algoritmo di irriducibilità, molto più efficiente della prova della divisione. Abbiamo già definito il polinomio universale $x^{p^n} - x$ in 2.3.5 ed abbiamo notato che questo si fattorizza nel prodotto di monici irriducibili di grado d per ogni d che divide n.
Ricordiamo che p è un primo ed $n = [\mathbb{F}_{p^n} : \mathbb{F}_p]$ e che possiamo scrivere

$$x^{p^n} - x = r_1 \cdots r_s, \quad r_i \text{ monici irriducibili, } \deg(r_i) = d \ \ \forall\, d|n, \ i = 1, \ldots, s.$$

Siano $r(x) \in \mathbb{F}_p[x]$ ed $m = \deg(r)$. Dividiamo $r(x)$ per il suo coefficiente direttore, rendendolo monico (se già non lo fosse): quindi anche ogni fattore di $r(x)$ sarà monico.
Se $r(x)$ ha un fattore irriducibile r_1 con $\deg(r_1) = i$ per $i < m$, allora r_1 divide $x^{p^i} - x$. Viceversa, se $r(x)$ non ha fattori in comune con $x^{p^i} - x$ per $i < m$, allora $r(x)$ è irriducibile.
Ad esempio: supponiamo che $r(x)$ sia riducibile di grado 8, ossia $r(x) = r_1 r_2$ con r_1, r_2 irriducibili, $\deg(r_1) = 3$ e $\deg(r_2) = 5$. Allora r_1 deve essere uno dei fattori di $x^{p^3} - x$ ed r_2 deve essere uno dei fattori di $x^{p^5} - x$.
Procediamo come segue: per ogni i, da 1 fino ad $m-1$ incluso, troviamo $x^{p^i} - x \bmod r(x)$. Ovvero, dividiamo $x^{p^i} - x$ per $r(x)$ ottenendo $x^{p^i} - x = q(x)r(x) + t(x)$ e calcoliamo $\mathrm{mcd}(r(x), t(x))$. Se c'è un fattore comune, sarà ritenuto mod $r(x)$, poiché la riduzione mod $r(x)$ mantiene il grado ragionevole. Se $\mathrm{mcd}(r(x), t(x)) = 1 \ \forall i < m$, $r(x)$ è irriducibile.
Di seguito, l'algoritmo nello specifico:

- p primo, $r(x)$ in $\mathbb{F}_p[x]$
- Posto $m := \deg(r)$
- Posto $U(x) := x$
- Per i da 1 ad $m - 1$:
  - Posto $U(x) := (U(x))^{p^i} \bmod r(x)$.
  - Posto $g(x) := \mathrm{mcd}(r(x), U(x) - x)$
  - Se $g(x) \neq 1$ allora restituisci RIDUCIBILE

- Restituisci IRRIDUCIBILE

Test della derivata per fattori multipli

Si può verificare velocemente per fattori che si ripetono calcolando il massimo comun divisore di r e della sua derivata formale. Questo non è un test di irriducibilità, ma può essere un modo veloce per svelare alcuni fattori.

Algoritmo di tutti gli irriducibili

Un'altra tecnica è quella che, dato un irriducibile (trovato ad esempio con uno dei metodi precedenti), produce tutti i rimanenti irriducibili di un dato grado.

- Dato $r(x)$ di grado n in $\mathbb{F}_p[x]$, si scrivono tutti i p^n elementi del campo finito $\mathbb{F}_p[x]/\langle r(x)\rangle$.
- Per ogni elemento α si prendono ripetute potenze p-esime fino ad ottenere una ripetizione (per facilitare i calcoli si può costruire la tabella dei logaritmi). Questo fornirà le orbite di Frobenius di ogni elemento.
- Per ogni orbita di Frobenius, si prendono tutti i distinti $\rho^i(\alpha)$, si scrive il prodotto $\prod_i(x-\rho^i(\alpha))$ e lo si moltiplica usando l'aritmetica definita da $r(x)$. Questo fornirà i polinomi minimi per ogni elemento.
- Si otterrà un risultato simile a quello della tabella della radice. Si scartano tutti i polinomi minimi di grado inferiore ad n. I restanti polinomi minimi sono tutti monici irriducibili di grado n.

Esempio 2.15. Consideriamo x^3+x+1 di grado 3 in $\mathbb{F}_2[x]$.

- Abbiamo otto elementi nel campo: 000, 001, 010, 011, 100, 101, 110, 111.
- $000^2 = 000$
- $001^2 = 001$

- $010^2 = 100, \quad 100^2 = 110, \quad 110^2 = 010.$
- $011^2 = 101, \quad 101^2 = 111, \quad 111^2 = 011.$
- $(x - 000) = x$
- $(x - 001) = x + 1$
- $(x - 010)(x - 100)(x - 110) = x^3 + x + 1$
- $(x - 011)(x - 101)(x - 111) = x^3 + x^2 + 1$
- I polinomi di grado 3 sono $x^3 + x + 1$ e $x^3 + x^2 + 1$.

2.4.4 Radici di polinomi irriducibili

Sia K un sottocampo di F e $\theta \in F$. Se θ soddisfa un'equazione polinomiale non banale con coefficienti in K, cioè se $a_n\theta^n + \cdots + a_1\theta + a_0 = 0$ con $a_i \in K$ non tutti nulli, allora θ si dice *algebrico* su K.

Definizione 2.9 (Polinomio minimo)**.**
Se $\theta \in F$ è algebrico su K, allora il polinomio monico $g \in K[x]$ che genera l'ideale $J = \{f \in K[x] \colon f(\theta) = 0\}$ di $K[x]$ si dice *polinomio minimo* di θ su K. Con *grado* di θ su K intendiamo il grado di g.

Teorema 2.4.5. *Se $\theta \in F$ è algebrico su K, allora il suo polinomio minimo g su K gode delle seguenti proprietà:*

(i) *g è irriducibile in $K[x]$.*

(ii) *Per $f \in K[x]$ abbiamo $f(\theta) = 0$ se e solo se g divide f.*

(iii) *g è il polinomio monico in $K[x]$ di grado minimo che ha θ come radice.*

Teorema 2.4.6. *Sia $\theta \in F$ algebrico di grado n su K e sia g il polinomio minimo di θ su K, quindi* $\deg(g) = n$*. Allora:*

(i) *$K(\theta)$ è isomorfo a $K[x]/(g)$.*

(ii) $[K(\theta) : K] = n$ *e* $(1, \theta, \ldots, \theta^{n-1})$ *è una base di* $K(\theta)$ *su* K.

(iii) *Ogni* $\alpha \in K(\theta)$ *è algebrico su* K *ed il suo grado su* K *è un divisore di* n.

Per le dimostrazioni dei teoremi precedenti si rimanda il lettore al [12, 4.1].

Definizione 2.10 (Coniugati)**.**
Sia $\mathbb{F}_{p^n}$ un'estensione di $\mathbb{F}_p$ e sia $\alpha \in \mathbb{F}_{p^n}$. Allora, gli elementi $\alpha, \alpha^p, \ldots, \alpha^{p^{n-1}}$ si dicono i *coniugati* di α rispetto ad $\mathbb{F}_p$.

I coniugati di $\alpha \in \mathbb{F}_{p^n}$ rispetto ad $\mathbb{F}_p$ sono distinti se e solo se il polinomio minimo di α su $\mathbb{F}_p$ ha grado n. Altrimenti, il grado d di questo polinomio minimo è un divisore proprio di n, e allora i coniugati di α rispetto ad $\mathbb{F}_p$ sono gli elementi distinti $\alpha, \alpha^p, \ldots, \alpha^{p^{d-1}}$, ognuno ripetuto n/d volte.

Teorema 2.4.7. *I coniugati di* $\alpha \in \mathbb{F}_{p^n}^{\times}$ *rispetto ad ogni sottocampo di* $\mathbb{F}_{p^n}$ *hanno lo stesso ordine nel gruppo* $\mathbb{F}_{p^n}^{\times}$.

Dimostrazione. Poiché $\mathbb{F}_{p^n}^{\times}$ è un gruppo ciclico per il Teorema 2.2.14, il risultato segue dai seguenti due fatti. 1) In un gruppo ciclico finito di ordine m generato da un elemento a, l'elemento a^k genera un sottogruppo di ordine $m/\mathrm{mcd}(k, m)$ (per la dimostrazione si veda [12, p.7]). 2) Ogni potenza della caratteristica di $\mathbb{F}_{p^n}$, ossia ogni potenza di p, è relativamente prima all'ordine $p^n - 1$ di $\mathbb{F}_{p^n}^{\times}$. □

Corollario 2.4.8. *Se* α *è un elemento primitivo di* $\mathbb{F}_{p^n}^{\times}$*, allora lo sono anche tutti i suoi coniugati rispetto a ciascun sottocampo di* $\mathbb{F}_{p^n}^{\times}$.

Esempio 2.16. Sia $\alpha \in \mathbb{F}_{2^4}$ una radice del polinomio $x^4 + x + 1 \in \mathbb{F}_2[x]$.
I coniugati di α rispetto ad $\mathbb{F}_2$ sono $\alpha, \alpha^2, \alpha^4 = \alpha + 1$ e $\alpha^8 = \alpha^2 + 1$, ognuno dei quali è un elemento primitivo di $\mathbb{F}_{2^4}$.
I coniugati di α rispetto ad $\mathbb{F}_{2^2}$ sono α ed $\alpha^4 = \alpha + 1$.

Automorfismo di Frobenius

C'è una stretta relazione tra gli elementi coniugati e certi automorfismi di un campo finito. Sia $\mathbb{F}_{p^n}$ un'estensione di $\mathbb{F}_p$. Con un automorfismo σ di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ intendiamo un automorfismo di $\mathbb{F}_{p^n}$ che fissa gli elementi di $\mathbb{F}_p$. Quindi, nel dettaglio richiediamo che σ sia una mappa iniettiva da $\mathbb{F}_{p^n}$ in se stesso con $\sigma(\alpha+\beta) = \sigma(\alpha)+\sigma(\beta)$ e $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$ $\forall \alpha, \beta \in \mathbb{F}_{p^n}$ e $\sigma(a) = a$ $\forall a \in \mathbb{F}_p$.

Teorema 2.4.9. *Gli automorfismi distinti di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ sono esattamente le mappe $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$, definite da $\sigma_j(\alpha) = \alpha^{p^j}$ per $\alpha \in \mathbb{F}_{p^n}$ e $0 \leq j \leq n-1$.*

Per la dimostrazione del Teorema 2.4.9 sfruttiamo il seguente risultato.

Teorema 2.4.10. *Se f è un polinomio irriducibile in $\mathbb{F}_p[x]$ di grado n, allora f ha una radice α in $\mathbb{F}_{p^n}$. Inoltre tutte le radici di f sono semplici e sono date dagli n elementi distinti $\alpha, \alpha^p, \ldots, \alpha^{p^{n-1}}$ di $\mathbb{F}_{p^n}$.*

Per la dimostrazione del Teorema 2.4.10 si rimanda il lettore a [12, p.52].

Dimostrazione. (Teorema 2.4.9).
Per ogni σ_j ed ogni $\alpha, \beta \in \mathbb{F}_{p^n}$ ovviamente abbiamo

$$\sigma_j(\alpha\beta) = (\alpha\beta)^p = \alpha^p\beta^p = \sigma_j(\alpha)\sigma_j(\beta),$$
$$\sigma_j(\alpha+\beta) = (\alpha+\beta)^p = \alpha^p + \beta^p = \sigma_j(\alpha) + \sigma_j(\beta)$$

e quindi σ_j è un endomorfismo di $\mathbb{F}_{p^n}$. Inoltre, $\sigma_j(\alpha) = 0$ se e solo se $\alpha = 0$ e quindi σ_j è iniettivo. Poiché $\mathbb{F}_{p^n}$ è un insieme finito, segue che è un automorfismo di $\mathbb{F}_{p^n}$. Abbiamo $\sigma_j(a) = a$, poiché per ogni $a \in \mathbb{F}_p$ $a^p = a$ e quindi ogni σ_j è un automorfismo di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$. Le mappe $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$ sono distinte poiché assumono valori distinti per un elemento primitivo di $\mathbb{F}_{p^n}$.
Ora supponiamo che σ sia un automorfismo arbitrario di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$. Sia β un elemento primitivo di $\mathbb{F}_{p^n}$ e sia $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{F}_p[x]$ il suo polinomio minimo su $\mathbb{F}_p$. Allora

$$\begin{aligned} 0 &= \sigma(\beta^n + a_{n-1}\beta^{n-1} + \cdots + a_0) \\ &= \sigma(\beta)^n + a_{n-1}\sigma(\beta)^{n-1} + \cdots + a_0, \end{aligned}$$

quindi $\sigma(\beta)$ è una radice di f in $\mathbb{F}_{p^n}$. Segue dal Teorema 2.4.10 che $\sigma(\beta) = \beta^{q^j}$ per qualche j, $0 \leq j \leq n-1$. Poiché σ è un omomorfismo, otteniamo allora $\sigma(\alpha) = \alpha^{p^j}$ per ogni $\alpha \in \mathbb{F}_{p^n}$. □

Sulla base del Teorema 2.4.9, è evidente che i coniugati di $\alpha \in \mathbb{F}_{p^n}$ rispetto ad $\mathbb{F}_p$ sono ottenuti applicando tutti gli automorfismi di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ all'elemento α. Gli automorfismi di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ formano un gruppo con l'operazione della composizione di mappe. Il Teorema 2.4.9 mostra che questo gruppo di automorfismi di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ è un gruppo ciclico di ordine n generato da σ_1. L'automorfismo σ_1 di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ si dice *automorfismo di Frobenius* di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$. Il gruppo di automorfismi di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$ si dice *gruppo di Galois* di $\mathbb{F}_{p^n}$ su $\mathbb{F}_p$.

Gruppo di Galois e orbite di Frobenius

Vediamo l'azione di σ_1 su un elemento α di $\mathbb{F}_{p^n}$:

$$\begin{aligned}
&\sigma_1(\alpha) = \alpha^p;\\
&\sigma_1^2(\alpha) = \sigma(\sigma(\alpha)) = (\alpha^p)^p = \alpha^{p^2};\\
&\vdots\\
&\sigma_1^{n-1}(\alpha) = \alpha^{p^{n-1}}.
\end{aligned}$$

Ma allora $\sigma_1^n(\alpha) = \alpha^{p^n} = \alpha$ quindi $\sigma_1^n = \sigma_1^0 = id$, la mappa identità. Questo ci permette di affermare che il gruppo di Galois è effettivamente ciclico di ordine n, generato dalla mappa σ_1 p-esima potenza. Ovvero:

$$\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle \sigma_1 \rangle \,;\ |\sigma_1| = n.$$

Dato a in $\mathbb{F}_{p^n}$, $\sigma_1^i(a)$ sono le *orbite* di a sotto l'azione dell'automorfismo di gruppo. Ovviamente c'è una relazione d'equivalenza sul campo per tutti quegli elementi che sono potenze p^i di un altro. Queste classi d'equivalenza formano una partizione del campo.
Abbiamo anche visto che $a^{p^n} = a$ per tutti gli elementi di $\mathbb{F}_{p^n}$ e quindi $a^{p^d} = a$ per tutti gli elementi di $\mathbb{F}_{p^d}$. Questo significa che quando prenderemo in

considerazione la corrispondenza di Galois, con $\mathbb{F}_{p^d}$ visto come sottocampo di $\mathbb{F}_{p^n}$, avremo σ_1^d identità sul campo $\mathbb{F}_{p^d}$ e

$$\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_{p^d}) = \langle \sigma_1^d \rangle \,;\ |\sigma_1^d| = n/d$$

che come gruppo avrà ordine n/d.

La corrispondenza di Galois

Abbiamo visto che il gruppo di Galois di $\mathbb{F}_{p^n}$ è ciclico di ordine n, generato dall'automorfismo di Frobenius σ_1. Abbiamo anche visto qualcosa sui sottocampi. Vogliamo ora saperne di più riguardo alla corrispondenza di Galois per campi finiti.
Possiamo così riassumere le discussioni sui sottocampi e sui gruppi di Galois:

$$\begin{array}{ll} \mathbb{F}_{2^4} = \mathbb{F}_2(\gamma),\ \gamma^4+\gamma+1=0=\mathbb{F}_2(\alpha)(\beta) & \langle \sigma_1^4 = id \rangle \\ \Big| & \Big| \\ \mathbb{F}_{2^2} = \mathbb{F}_2(\alpha),\ \alpha^2+\alpha+1=0 & \langle \sigma_1^2 \rangle \\ \Big| & \Big| \\ \mathbb{F}_2 & \langle \sigma_1 \rangle \end{array}$$

Ricapitolando:

- La mappa p-esima potenza, σ_1, è l'identità su $\mathbb{F}_2$. Le quattro mappe $\sigma_1,\ \sigma_1^2,\ \sigma_1^4,\ \sigma_1^{16} = id$, sono automorfismi distinti su $\mathbb{F}_{2^4}$ e formano il gruppo di Galois di $\mathbb{F}_{2^4}$ su $\mathbb{F}_2$.

- La mappa p-esima potenza, σ_1, è l'identità su $\mathbb{F}_2$. Le due mappe $\sigma_1, \sigma_1^2 = id$ sono automorfismi distinti su $\mathbb{F}_{2^2}$ e formano il gruppo di Galois di $\mathbb{F}_{2^2}$ su $\mathbb{F}_2$.

- La mappa p^2-esima potenza, σ_1^2, è l'identità su $\mathbb{F}_{2^2}$. Le due mappe $\sigma_1^2, \sigma_1^4 = id$ sono automorfismi distinti su $\mathbb{F}_{2^2}$ e formano il gruppo di Galois di $\mathbb{F}_{2^4}$ su $\mathbb{F}_{2^2}$.

2.4.5 Traccia di un elemento e polinomio caratteristico

Sia $F = \mathbb{F}_{p^n}$ estensione del campo $K = \mathbb{F}_p$ visto come spazio vettoriale su K. Allora F ha dimensione n su K e se $(\alpha_1, \ldots, \alpha_n)$ è una base di F su K, ogni elemento $\alpha \in F$ può essere univocamente rappresentato nella forma

$$\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n, \quad \text{con } c_j \in K \quad \text{per1} \leq j \leq n.$$

Introduciamo ora un'importante mappa lineare da F a K.

Definizione 2.11 (Traccia)**.**
Per $\alpha \in F = \mathbb{F}_{p^n}$ e $K = \mathbb{F}_p$, la *traccia* $T(\alpha)$ di α su K è definita da

$$T(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{n-1}}.$$

In altri termini, la traccia di α su K è la somma dei coniugati di α rispetto a K. Un'altra definizione di traccia può essere ottenuta come segue.

Definizione 2.12 (Polinomio caratteristico)**.**
Sia $f \in K[x]$ il polinomio minimo di α su K, il cui grado d è un divisore di n. Allora

$$g(x) = f(x)^{n/d} \in K[x]$$

si dice *polinomio caratteristico* di α su K.

Per il Teorema 2.4.10, le radici di f in F sono date da $\alpha, \alpha^p, \ldots, \alpha^{p^{d-1}}$, allora l'osservazione successiva alla definizione 2.10 implica che le radici di g in F sono precisamente i coniugati di α rispetto a K. Quindi

$$\begin{aligned} g(x) &= x^n + a_{n-1}x^{n-1} + \cdots + a_0 \\ &= (x-\alpha)(x-\alpha^p)\cdots(x-\alpha^{p^{n-1}}), \end{aligned}$$

e confrontando i coefficienti si dimostra che

$$T(\alpha) = -a_{n-1}.$$

In particolare, $T(\alpha)$ è sempre un elemento di K.

Teorema 2.4.11. *Sia $K = \mathbb{F}_p$ e $F = \mathbb{F}_{p^n}$. Allora la funzione traccia soddisfa le seguenti proprietà:*

(i) $\forall \alpha, \beta \in F, \quad T(\alpha + \beta) = T(\alpha) + T(\beta)$;

(ii) $\forall c \in K, \alpha \in F \quad T(c\alpha) = cT(\alpha)$;

(iii) *T è una trasformazione lineare suriettiva da F a K, dove sia F che K sono intesi come spazi vettoriali su K;*

(iv) $\forall a \in K, \quad T(a) = na$;

(v) $\forall \alpha \in F, \quad T(\alpha^p) = T(\alpha)$.

Dimostrazione. (i) Usiamo il fatto che in un campo di caratteristica p vale: $(\alpha + \beta)^{p^n} = \alpha^{p^n} + \beta^{p^n}$ (si veda [12, Teorema 1.46]). Per $\alpha, \beta \in F$, si ha

$$\begin{aligned} T(\alpha + \beta) &= \alpha + \beta + (\alpha + \beta)^p + \cdots + (\alpha + \beta)^{p^{n-1}} \\ &= \alpha + \beta + \alpha^p + \beta^p + \cdots + \alpha^{p^{n-1}} + \beta^{p^{n-1}} \\ &= T(\alpha) + T(\beta). \end{aligned}$$

(ii) Per $c \in K$ abbiamo $c^{p^j} = c \ \forall \ j \geq 0$ per il Corollario 2.2.4. Quindi, otteniamo per $\alpha \in F$

$$\begin{aligned} T(c\alpha) &= c\alpha + c^p\alpha^p + \cdots + c^{p^{n-1}}\alpha^{p^{n-1}} \\ &= c\alpha + c\alpha^p + \cdots + c\alpha^{p^{n-1}} \\ &= cT(\alpha). \end{aligned}$$

(iii) Le proprietà (i) e (ii) insieme al fatto che $T(\alpha) \in K$ per ogni $\alpha \in F$ mostrano che la funzione traccia è una trasformazione lineare da F a K. Per provare che è una mappa suriettiva, basta mostrare l'esistenza di un $\alpha \in F$ con $T(\alpha) \neq 0$. Ora, $T(\alpha) = 0$ se e solo se α è una radice del polinomio $x^{p^{n-1}} + \cdots + x^p + x \in K[x]$ in F. Ma questo polinomio può avere al più p^{n-1} radici in F ed F ha p^n elementi.
(iv) Segue immediatamente dalla definizione della funzione traccia e dal Corollario 2.2.4.
(v) Per $\alpha \in F$ abbiamo $\alpha^{p^n} = \alpha$ per il Corollario 2.2.4 e quindi $T(\alpha^p) = \alpha^p + \alpha^{p^2} + \cdots + \alpha^{p^n} = T(\alpha)$. □

2.5 Primitività in $\mathbb{F}_{p^n}$

Definizione 2.13 (Polinomio primitivo)**.**
Un polinomio $r(x) \in \mathbb{F}_p[x]$ di grado $n \geq 1$ si dice polinomio *primitivo* su $\mathbb{F}_p$ se è il polinomio minimo su $\mathbb{F}_p$ di un elemento primitivo di $\mathbb{F}_{p^n}$.

Quindi, un polinomio primitivo su $\mathbb{F}_p$ di grado n può essere descritto come un polinomio monico che è irriducibile su $\mathbb{F}_p$ ed ha una radice $\alpha \in \mathbb{F}_{p^n}$ che genera il gruppo moltiplicativo di $\mathbb{F}_{p^n}$.

2.5.1 La formula per il numero di polinomi primitivi

La tabella delle radici 2.2 rende ovvie le derivazioni delle formule per il numero di monici irriducibili e di primitivi.
Vediamo la formula per il numero di polinomi primitivi. Il gruppo moltiplicativo di $\mathbb{F}_{p^n}$ è ciclico con ordine $p^n - 1$. Sappiamo che questo gruppo ciclico è generato da degli elementi i cui logaritmi sono relativamente primi a $p^n - 1$. Il numero di tali logaritmi è uguale a $\varphi(p^n - 1)$ dove φ è la funzione di Eulero. Così, ci sono $\varphi(p^n - 1)$ generatori del gruppo moltiplicativo di $\mathbb{F}_{p^n}$. Poiché generano tutto $\mathbb{F}_{p^n}^{\times}$, non possono essere elementi di un sottocampo proprio, quindi i loro polinomi minimi hanno tutti grado n. Inoltre, se una delle radici di un polinomio è primitiva, lo sono tutte. Quindi i $\varphi(p^n - 1)$ generatori sono valutati per i loro

$$\frac{1}{n}\varphi(p^n - 1)$$

polinomi minimi, ognuno dei quali è primitivo. Anche questa funzione assume valori positivi, garantendoci l'esistenza di almeno un monico irriducibile primitivo per ogni primo p e per ogni intero positivo n.

2.5.2 Come trovare i polinomi primitivi in $\mathbb{F}_p[x]$

La chiave è che $\mathbb{F}_p^{\times}$ è ciclico di ordine $p - 1$; dato a in $\mathbb{F}_p$, basta verificare che $a^d \neq 1$ per tutti i divisori propri d di $p - 1$.

Qui si applica lo stesso concetto. Dato $r(x)$, prima si controlla se $r(x)$ è irriducibile (come visto in precedenza). Successivamente, $r(x)$ sarà primitivo se u è un elemento primitivo mod r. Sia $n = \deg(r)$: lavoriamo nel campo finito $\mathbb{F}_p[x]/\langle r(x)\rangle$. Ricordiamo che questo campo ha un gruppo moltiplicativo di ordine $p^n - 1$. Si trovano tutti i divisori propri d di $p^n - 1$. Se $u^d \neq 1$ per ognuno di essi, allora u è un elemento primitivo mod r ed r è un polinomio primitivo. Altrimenti, r non è primitivo.
Nota: Il grosso del lavoro richiesto dipende fortemente dalla fattorizzazione di $p^n - 1$. Notiamo anche che per $p = 2$ ed n opportuno si ottengono i primi di Mersenne, ossia dei primi della forma $2^n - 1$. In alcuni casi, tutti i monici irriducibili di grado n sono primitivi per il Teorema di Lagrange. Infatti per il Teorema di Lagrange tutti gli elementi del campo, ad eccezione dello 0 e dell'1, hanno ordine $p^n - 1$.

Capitolo 3

Registri a scorrimento a retroazione lineare

3.1 Sequenze ricorrenti lineari

3.1.1 Feedback shift register e periodicità

Sia k un intero positivo. Siano $a, a_0, \ldots, a_{k-1}$ elementi di un campo finito $\mathbb{F}_p$.

Definizione 3.1 (Sequenza ricorrente lineare affine e lineare)**.**
Una sequenza $s_0, s_1, \ldots$ di elementi di $\mathbb{F}_p$ che soddisfano la relazione

$$s_{n+k} = a_{k-1}s_{n+k-1} + a_{k-2}s_{n+k-2} + \cdots + a_0 s_n + a \quad \text{per } n = 0, 1, \ldots \tag{3.1}$$

si dice *sequenza ricorrente lineare affine* in $\mathbb{F}_p$ di ordine k.
I termini $s_0, s_1, \ldots, s_{k-1}$, che determinano la sequenza univocamente si dicono *valori iniziali*.
Una relazione della forma (3.1) si dice *relazione a ricorrenza lineare affine* di ordine k.
Se $a = 0$, parliamo di *relazione a ricorrenza lineare*:

$$s_{n+k} = a_{k-1}s_{n+k-1} + a_{k-2}s_{n+k-2} + \cdots + a_0 s_n \quad \text{per } n = 0, 1, \ldots; \tag{3.2}$$

analogamente quando ci si riferisce alla sequenza.

La generazione di sequenze ricorrenti lineari affini può essere implementata su un *feedback shift register* (registro a scorrimento a retroazione). Si tratta di un particolare circuito elettronico che tratta l'informazione sotto forma di elementi di $\mathbb{F}_p$ opportunamente rappresentati. Il registro utilizza quattro dispositivi. Il primo è un *sommatore*, che ha due input ed un output, che è somma in $\mathbb{F}_p$ dei due input. Il secondo è un *moltiplicatore costante*, che ha un input e fornisce come output il prodotto dell'input per una costante di $\mathbb{F}_p$. Il terzo è un *sommatore costante* che somma all'input una costante di $\mathbb{F}_p$. Il quarto tipo di dispositivo è un *elemento di ritardo* ("flip-flop"), che ha un input ed un output. Il flip-flop è regolato da un contatore sincrono esterno, in modo tale che il suo input, in un particolare istante, appare come il suo output dopo un'unità di tempo. La rappresentazione delle componenti negli schemi dei circuiti è mostrata in figura 3.1. Un registro a scorrimento a

Figura 3.1: Si veda [12, Figure 8.1]. Componenti del registro a scorrimento a retroazione. (a) Sommatore. (b) Moltiplicatore costante per moltiplicare per a. (c) Sommatore costante per sommare a. (d) Elemento di ritardo.

retroazione è costruito interconnettendo un numero finito di sommatori, moltiplicatori costanti, sommatori costanti ed elementi di ritardo su un circuito chiuso, in modo tale che due output non sono mai connessi insieme. In realtà, per generare sequenze ricorrenti lineari affini, basta connettere le componenti in un modo piuttosto particolare. Un registro a scorrimento che genera una sequenza ricorrente lineare affine soddisfacente la (3.1), è mostrata in figura 3.2. Sin dall'inizio, ogni elemento di ritardo D_j, $j = 0, 1, \ldots, k-1$, contiene il valore iniziale s_j. Se pensiamo alle operazioni aritmetiche ed al trasferimento lungo i collegamenti da eseguire istantaneamente, allora dopo un'unità di tempo ogni D_j conterrà s_{j+1}. Procedendo in questo modo, vediamo che l'output del registro a scorrimento è la stringa di elementi $s_0, s_1, s_2, \ldots$

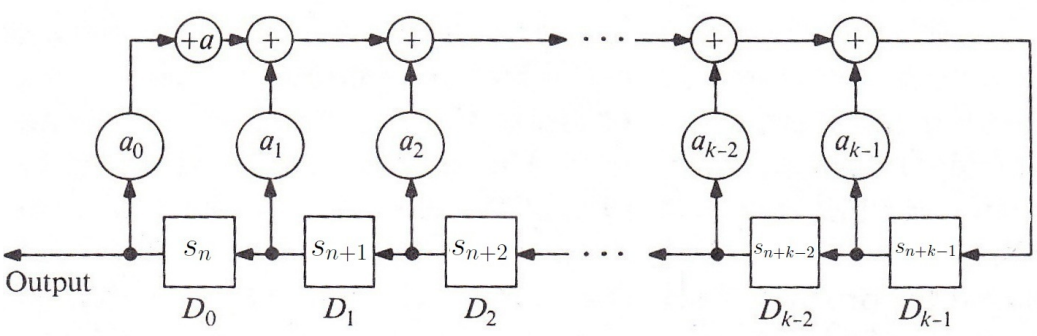

Figura 3.2: Esempio di registro a scorrimento

ricevuta negli intervalli di un'unità di tempo. Nella maggior parte delle applicazioni, la sequenza ricorrente lineare affine desiderata è lineare ed in questo caso il sommatore costante non è necessario.

Esempio 3.1. Consideriamo la relazione a ricorrenza lineare

$$s_{n+7} = s_{n+4} + s_{n+3} + s_{n+2} + s_n \quad n = 0, 1, \ldots \text{ in } \mathbb{F}_2.$$

Un registro a scorrimento corrispondente a questa relazione a ricorrenza lineare è mostrato in figura 3.3. Poiché la moltiplicazione per una costante in $\mathbb{F}_2$ conserva o annulla gli elementi, l'effetto di un moltiplicatore costante può essere simulato da un collegamento con fili o da una disconnessione. Quindi, un registro a scorrimento per la generazione di sequenze binarie a ricorrenza lineare, richiede solo elementi di ritardo, sommatori e collegamenti con fili.

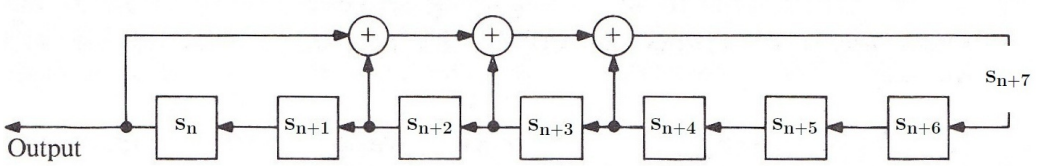

Figura 3.3: Esempio di registro a scorrimento nel caso $p = 2$ e $k = 7$

Se n è un intero non negativo, allora dopo n unità di tempo l'elemento di ritardo D_j conterrà s_{n+j}. Quindi è naturale chiamare il vettore riga

$$\mathbf{s_n} = (s_n, s_{n+1}, \ldots, s_{n+k-1})$$

vettore dell' n-esimo stato della sequenza a ricorrenza lineare (o del registro a scorrimento). Il vettore $\mathbf{s_0} = (s_0, s_1, \ldots, s_{k-1})$ si dice *vettore dello stato iniziale*.
Prima di studiare la periodicità delle sequenze a ricorrenza lineare nei campi finiti, introduciamo un po' di terminologia sulle *sequenze periodiche alla fine*.

Definizione 3.2 (Sequenza periodica alla fine)**.**
Sia S un insieme arbitrario non vuoto e $s_0, s_1, \ldots$ una sequenza di elementi di S. Se esistono degli interi $r > 0$ e $n_0 \geq 0$ tali che $s_{n+r} = s_n$ per ogni $n \geq n_0$, allora la sequenza si dice *periodica alla fine* ed r si dice *periodo* della sequenza. Il più piccolo dei periodi possibili di una sequenza periodica alla fine si dice *periodo minimo* della sequenza.

Definizione 3.3 (Sequenza periodica)**.**
Una sequenza periodica alla fine $s_0, s_1, \ldots$ con periodo minimo r si dice *periodica* se $s_{n+r} = s_n$ vale per ogni $n = 0, 1, \ldots$.

Se $s_0, s_1, \ldots$ è periodica alla fine con periodo minimo r, allora il più piccolo intero non negativo n_0 tale che $s_{n+r} = s_n$ per ogni $n \geq n_0$ si dice *preperiodo*. La sequenza è periodica se il preperiodo è 0.

Teorema 3.1.1. *Siano $\mathbb{F}_q$ un campo finito e k un intero positivo. Allora ogni sequenza ricorrente lineare affine di ordine k è periodica alla fine con periodo minimo r tale che $r \leq q^k$ e $r \leq q^k - 1$ se la sequenza è lineare.*

Dimostrazione. Notiamo che ci sono esattamente q^k k-uple distinte di elementi di $\mathbb{F}_q$. Quindi, considerando i vettori dello stato $\mathbf{s_m}$, $0 \leq m \leq q^k$, di una data sequenza ricorrente lineare affine di ordine k in $\mathbb{F}_q$, segue che $\mathbf{s_j} = \mathbf{s_i}$ per qualche i e j con $0 \leq i < j \leq q^k$. Usando la relazione a ricorrenza lineare affine e l'induzione, arriviamo a $\mathbf{s_{n+j-i}} = \mathbf{s_n}$ per ogni $n \geq i$, dimostrando che la sequenza stessa è periodica alla fine con periodo minimo $r \leq j - i \leq q^k$. Nel caso in cui la sequenza fosse lineare e nessun vettore dello stato fosse il vettore nullo, si potrebbe trattare lo stesso argomento ma con $q^k - 1$ al posto di q^k, per ottenere $r \leq q^k - 1$. Se invece, uno dei vettori dello stato di una

sequenza lineare fosse il vettore nullo, allora tutti i vettori dello stato successivi sarebbero vettori nulli e quindi la sequenza avrebbe periodo minimo $r = 1 \leq q^k - 1$. □

Un'importante **condizione sufficiente per la periodicità** di una sequenza ricorrente lineare affine è fornita dal risultato seguente.

Teorema 3.1.2. *Se $s_0, s_1, \ldots$ è una sequenza ricorrente lineare affine in un campo finito, soddisfacente la relazione a ricorrenza lineare affine* (3.1) *e se il coefficiente a_0 in* (3.1) *è non nullo, allora la sequenza $s_0, s_1, \ldots$ è periodica.*

Dimostrazione. In base al Teorema 3.1.1, la sequenza data è periodica alla fine. Se r è il suo periodo minimo ed n_0 il suo preperiodo, allora $s_{n+r} = s_n$ per ogni $n \geq n_0$. Supponiamo $n_0 \geq 1$. Da (3.1) con $n = n_0 + r - 1$ e per $a_0 \neq 0$, otteniamo

$$\begin{aligned} s_{n_0-1+r} &= a_0^{-1}(s_{n_0+k-1+r} - a_{k-1}s_{n_0+k-2+r} - \cdots - a_1 s_{n_0+r} - a) \\ &= a_0^{-1}(s_{n_0+k-1} - a_{k-1}s_{n_0+k-2} - \cdots - a_1 s_{n_0} - a). \end{aligned}$$

Usando la (3.1) con $n = n_0 - 1$, troviamo la stessa espressione per s_{n_0-1} e quindi $s_{n_0-1+r} = s_{n_0-1}$. Questa è una contraddizione per la definizione di preperiodo. □

Sia $s_0, s_1, \ldots$ una sequenza ricorrente lineare di ordine k in $\mathbb{F}_q$ soddisfacente la relazione (3.2), dove $a_j \in \mathbb{F}_q$ per $0 \leq j \leq k-1$. A questa sequenza associamo la matrice A $k \times k$ su $\mathbb{F}_q$ definita da

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{k-1} \end{pmatrix}. \tag{3.3}$$

Notiamo che la matrice A dipende solo dalla relazione a ricorrenza lineare soddisfatta dalla sequenza data.

Lemma 3.1.3. *Se $s_0, s_1, \ldots$ è una sequenza ricorrente lineare in $\mathbb{F}_q$ soddisfacente la* (3.2) *e A è la matrice in* (3.3) *ad essa associata, allora per i vettori dello stato della sequenza abbiamo*

$$\mathbf{s_n} = \mathbf{s_0} A^n \quad \textit{per } n = 0, 1, \ldots. \tag{3.4}$$

Dimostrazione. Poiché $\mathbf{s_n} = (s_n, s_{n+1}, \ldots, s_{n+k-1})$, si verifica facilmente che $\mathbf{s_{n+1}} = \mathbf{s_n} A$ per ogni $n \geq 0$, così (3.4) segue per induzione. □

Notiamo che l'insieme di tutte le matrici non singolari $k \times k$ su $\mathbb{F}_q$ forma un gruppo finito col prodotto di matrici, detto *gruppo lineare generale* $GL(k, \mathbb{F}_q)$.

3.1.2 Sequenze di lunghezza massima e polinomio caratteristico

Tra tutte le sequenze ricorrenti lineari in $\mathbb{F}_q$ che soddisfano una data relazione a ricorrenza di ordine k come in (3.2), possiamo individuarne una che fornisce il massimo valore per il periodo minimo in questa classe di sequenze. Vedremo nel corso della dimostrazione del Teorema 3.2.2 che si possono trovare più di una sequenza di questo tipo. Questa è la *sequenza risposta impulsiva* $d_0, d_1, \ldots$ univocamente determinata dai suoi valori iniziali $d_0 = \cdots = d_{k-2} = 0$, $d_{k-1} = 1$ ($d_0 = 1$ se $k = 1$) e la relazione a ricorrenza lineare

$$d_{n+k} = a_{k-1} d_{n+k-1} + a_{k-2} d_{n+k-2} + \cdots + a_0 d_n \quad \text{per } n = 0, 1, \ldots. \tag{3.5}$$

Esempio 3.2. Consideriamo la relazione a ricorrenza lineare

$$s_{n+5} = s_{n+1} + s_n, \quad n = 0, 1, \ldots, \text{ in } \mathbb{F}_2.$$

La sequenza risposta impulsiva $d_0, d_1, \ldots$ corrispondente ad essa è data dalla stringa di cifre binarie

$$0\,0\,0\,0\,1\,0\,0\,0\,1\,1\,0\,0\,1\,0\,1\,0\,1\,1\,1\,1\,1\,0\,0\,0\,0\,1\cdots$$

di periodo minimo 21. Un registro a scorrimento a retroazione che genera questa sequenza è mostrato in figura 3.4. Possiamo pensare questa sequenza come ottenuta partendo dallo stato in cui ogni elemento di ritardo è vuoto (cioè contiene 0) e poi inviando l'impulso 1 all'elemento di ritardo più a destra. Questo spiega il termine "sequenza risposta impulsiva". Vediamo

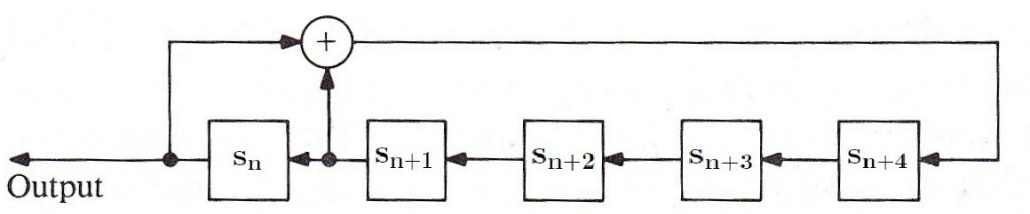

Figura 3.4: Registro a scorrimento che genera una sequenza di lunghezza massima

ora che il periodo minimo della sequenza è uguale all'ordine della matrice

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

in $GL(5, \mathbb{F}_2)$. Se il vettore dello stato iniziale di una sequenza ricorrente lineare in $\mathbb{F}_2$, soddisfacente la relazione data, è uguale ad uno dei 21 vettori di stato differenti che compaiono nella sequenza risposta impulsiva, allora il periodo minimo è ancora 21 (poiché tale sequenza non è altro che la sequenza risposta impulsiva traslata). Se prendiamo $(1, 1, 1, 0, 1)$ come vettore dello stato iniziale, otteniamo la stringa di cifre binarie $1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 1\cdots$ di periodo minimo 7 e lo stesso periodo minimo risulta da uno qualsiasi dei 7 vettori di stato differenti di questa sequenza nel ruolo di vettore dello stato iniziale.

Se prendiamo $(1, 1, 0, 1, 1)$ come vettore dello stato iniziale, otteniamo la stringa di cifre binarie $1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\cdots$ di periodo minimo 3 e lo stesso periodo minimo risulta da uno qualsiasi dei 3 vettori di stato differenti di

questa sequenza preso come vettore dello stato iniziale.
Il vettore di stato iniziale $(0, 0, 0, 0, 0)$ produce una sequenza di periodo minimo 1. Abbiamo così esaurito le 32 possibilità per i vettori dello stato iniziale.

Sia $s_0, s_1, \ldots$ una sequenza ricorrente lineare di ordine k in $\mathbb{F}_q$, soddisfacente la relazione (3.2). Il polinomio

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

si dice *polinomio caratteristico* della sequenza che, ovviamente, dipende solo dalla relazione (3.2). Se A è la matrice in (3.3), allora si vede facilmente che $f(x)$ è identico al polinomio caratteristico di A nel senso dell'algebra lineare, cioè $f(x) = \det(xI - A)$, con I matrice identità $k \times k$ su $\mathbb{F}_q$. D'altra parte, la matrice A può essere pensata come la matrice associata al polinomio monico $f(x)$.
Come prima applicazione del polinomio caratteristico, mostriamo come i termini di una sequenza ricorrente lineare possano essere rappresentati esplicitamente in un importante caso speciale.

Teorema 3.1.4. *Sia $s_0, s_1, \ldots$ una sequenza ricorrente lineare di ordine k in $\mathbb{F}_q$ con polinomio caratteristico $f(x)$. Se le radici $\alpha_1, \ldots, \alpha_k$ di $f(x)$ sono tutte distinte, allora*

$$s_n = \sum_{j=1}^{k} \beta_j \alpha_j^n \quad \textit{per } n = 0, 1, \ldots, \tag{3.6}$$

dove $\beta_1, \ldots, \beta_k$ sono elementi univocamente determinati dai valori iniziali della sequenza ed appartengono al campo di spezzamento di $f(x)$ su $\mathbb{F}_q$.

Dimostrazione. Le costanti $\beta_1, \ldots, \beta_k$ possono essere determinate dal sistema di equazioni lineari

$$\sum_{j=1}^{k} \alpha_j^n \beta_j = s_n, \quad n = 0, 1, \ldots, k-1.$$

Poiché il determinante di questo sistema è un determinante di Vandermonde, non nullo per la condizione sulle $\alpha_1, \ldots, \alpha_k$, gli elementi $\beta_1, \ldots, \beta_k$ sono univocamente determinati e appartengono al campo di spezzamento $\mathbb{F}_q(\alpha_1, \ldots, \alpha_k)$ di $f(x)$ su $\mathbb{F}_q$, come si vede dalla regola di Cramer. Per provare l'identità (3.6) per ogni $n \geq 0$, basta verificare se gli elementi del membro destro della (3.6), con questi specifici valori dei $\beta_1, \ldots, \beta_k$, soddisfano la relazione (3.2). Ma

$$\sum_{j=1}^{k} \beta_j \alpha_j^{n+k} - a_{k-1} \sum_{j=1}^{k} \beta_j \alpha_j^{n+k-1} - a_{k-2} \sum_{j=1}^{k} \beta_j \alpha_j^{n+k-2} - \cdots - a_0 \sum_{j=1}^{k} \beta_j \alpha_j^{n}$$

$$= \sum_{j=1}^{k} \beta_j f(\alpha_j) \alpha_j^n = 0,$$

per ogni $n \geq 0$, poiché le α_j sono le radici di $f(x)$ e la dimostrazione è completa. □

Osservazione 2. Una formula simile alla (3.6) è valida se la molteplicità di ogni radice di $f(x)$ è pari al più alla caratteristica p di $\mathbb{F}_q$. Nel dettaglio, siano $\alpha_1, \ldots, \alpha_m$ le radici distinte di $f(x)$ e supponiamo che ogni α_i $(i = 1, \ldots, m)$, abbia molteplicità $e_i \leq p$ e che $e_i = 1$ se $\alpha_i = 0$. Allora abbiamo

$$s_n = \sum_{i=1}^{k} P_i(n) \alpha_i^n \quad \text{per } n = 0, 1, \ldots,$$

dove ogni P_i, $i = 1, 2, \ldots, m$, è un polinomio di grado inferiore a e_i i cui coefficienti sono univocamente determinati dai valori iniziali della sequenza ed appartengono al campo di spezzamento di $f(x)$ su $\mathbb{F}_q$. L'intero n chiaramente è identificato come di consueto con un elemento di $\mathbb{F}_q$.

Nel caso in cui il polinomio carratteristico fosse irriducibile, gli elementi della sequenza ricorrente lineare possono essere rappresentati in termini di un'appropriata funzione traccia (si veda Definizione 2.11).
Vediamo ora un teorema che ci tornerà utile nella dimostrazione di un risultato importante nel caso di polinomio caratteristico irriducibile della sequenza.

Teorema 3.1.5. *Sia F un'estensione finita del campo finito K, entrambi considerati come spazi vettoriali su K. Allora le trasformazioni lineari da F a K sono esattamente le mappe L_β, $\beta \in F$, dove $L_\beta(\alpha) = T(\beta\alpha) := \beta\alpha + \beta^q\alpha^q + \cdots + \beta^{q^{k-1}}\alpha^{q^{k-1}}$ per ogni $\alpha \in F$. Inoltre $L_\beta \neq L_\gamma$ se β e γ sono elementi distinti di F.*

Per la dimostrazione si veda [12, Teorema 2.24].

Teorema 3.1.6. *Sia $s_0, s_1, \ldots$ una sequenza ricorrente lineare omogenea di ordine k in $K = \mathbb{F}_q$ il cui polinomio caratteristico $f(x)$ è irriducibile su K. Sia α una radice di $f(x)$ nell'estensione di campo $F = \mathbb{F}_{q^k}$. Allora esiste $\theta \in F$ univocamente determinato tale che*

$$s_n = T(\theta\alpha^n) \quad \textit{per } n = 0, 1, \ldots.$$

Dimostrazione. Poiché $(1, \alpha, \ldots, \alpha^{k-1})$ costituisce una base di F su K, possiamo definire una mappa lineare L univocamente determinata da F a K impostando $L(\alpha^n) = s_n$ per $n = 0, 1, \ldots, k-1$. Per il Teorema 3.1.5 esiste $\theta \in F$ univocamente determinato tale che $L(\gamma) = T(\theta\gamma)$ per ogni $\gamma \in F$. In particolare abbiamo

$$s_n = T(\theta\alpha^n) \quad \text{per } n = 0, 1, \ldots, k-1.$$

Rimane da dimostrare che gli elementi $T(\theta\alpha^n)$, $n = 0, 1, \ldots$, formano una sequenza ricorrente lineare omogenea con polinomio caratteristico $f(x)$. Ma se $f(x) = x^k - a_{k-1}x^{k-1} - \cdots - a_0 \in K[x]$, allora usando le proprietà della funzione traccia otteniamo

$$\begin{aligned} &T(\theta\alpha^{n+k}) - a_{k-1}T(\theta\alpha^{n+k-1}) - \cdots - a_0T(\theta\alpha^n) \\ &\quad = T(\theta\alpha^{n+k} - a_{k-1}\theta\alpha^{n+k-1} - \cdots - a_0\theta\alpha^n) \\ &\quad = T(\theta\alpha^n f(\alpha)) = 0 \end{aligned}$$

per ogni $n \geq 0$. □

Teorema 3.1.7. *Sia $s_0, s_1, \ldots$ una sequenza ricorrente lineare omogenea di ordine k in $\mathbb{F}_q$ che soddisfa la relazione (3.2) ed è periodica con periodo r.*

Sia $f(x)$ il polinomio caratteristico della sequenza. Allora vale l'identità

$$f(x)s(x) = (1 - x^r)h(x) \tag{3.7}$$

con $s(x) = s_0x^{r-1} + s_1x^{r-2} + \cdots + s_{r-2}x + s_{r-1} \in \mathbb{F}_q[x]$ *e*

$$h(x) = \sum_{j=0}^{k-1} \sum_{i=0}^{k-1-j} a_{i+j+1}s_ix^j \in \mathbb{F}_q[x], \tag{3.8}$$

dove poniamo $a_k = -1$.

Per la dimostrazione si veda [12, Teorema 8.25].

Osservazione 3. Notiamo che per $f(0) \neq 0$ il periodo minimo della sequenza risposta impulsiva può essere ottenuto dall'identità (3.7) nel modo seguente. Per la sequenza risposta impulsiva con polinomio caratteristico $f(x)$, il polinomio $h(x)$ in (3.8) è dato da $h(x) = -1$. Quindi, se r è il periodo minimo della sequenza risposta impulsiva, allora $f(x)$ divide $x^r - 1$ per la (3.7).

Le sequenze ricorrenti lineari i cui periodi minimi sono molto grandi sono di particolare importanza nelle applicazioni. Sappiamo dal Teorema 3.1.1 che per una sequenza ricorrente lineare omogenea di ordine k in $\mathbb{F}_q$ il periodo minimo può essere al più $q^k - 1$. Per poter generare tali sequenze per le quali il periodo minimo è uguale a $q^k - 1$, dobbiamo usare la nozione di polinomio primitivo (vedere Definizione 2.13).

Definizione 3.4 (Sequenza di periodo massimo)**.**
Una sequenza ricorrente lineare in $\mathbb{F}_q$ il cui polinomio caratteristico è un polinomio primitivo su $\mathbb{F}_q$ e che ha un vettore dello stato iniziale non nullo si dice *sequenza di periodo massimo* in $\mathbb{F}_q$.

Teorema 3.1.8. *Ogni sequenza di periodo massimo di ordine k in $\mathbb{F}_q$ è periodica ed il suo periodo minimo è uguale al più grande valore possibile per il periodo minimo di una qualsiasi sequenza ricorrente lineare di ordine k in $\mathbb{F}_q$, ossia $q^k - 1$.*

Esempio 3.3. Consideriamo la relazione a ricorrenza lineare

$$s_{n+7} = s_{n+4} + s_{n+3} + s_{n+2} + s_n \quad n = 0, 1, \ldots \text{ in } \mathbb{F}_2$$

che ha $f(x) = x^7 - x^4 - x^3 - x^2 - 1 \in \mathbb{F}_2[x]$ come polinomio caratteristico. Poiché $f(x)$ è un polinomio primitivo su $\mathbb{F}_2$, ogni sequenza con vettore dello stato iniziale non nullo derivante da questa relazione è una sequenza di periodo massimo in $\mathbb{F}_2$. Se scegliamo un particolare vettore dello stato iniziale non nullo, allora la sequenza risultante $s_0, s_1, \ldots$ ha periodo minimo $2^7 - 1 = 127$ in base al Teorema 3.1.8. Quindi, tutti i possibili vettori non nulli di $\mathbb{F}_2^7$ appaiono come vettori di stato in questa sequenza. Ogni altra sequenza di periodo massimo derivante dalla relazione data non è altro che una versione traslata della sequenza $s_0, s_1, \ldots$.

3.2 Registri a scorrimento lineare e segnale GPS

I registri a scorrimento lineare permettono la generazione di sequenze che hanno delle proprietà eccellenti, poiché consentono ad un ricevitore di sincronizzarsi con esse. Questi dispositivi sono di facile costruzione (abbiamo visto nel paragrafo 3.1, che si può riprodurre un registro a scorrimento lineare con poche componenti elettriche di base). Inoltre, questi registri generano segnali pseudo-casuali, ovvero generano segnali che appaiono per lo più casuali, anche se essi sono generati da degli algoritmi deterministici.
Costruiremo un registro a scorrimento lineare che genera un segnale periodico di periodo $2^k - 1$. Questo avrà la proprietà di essere estremamente poco correlato con tutte le traslazioni di se stesso e con altri segnali generati dallo stesso registro usando coefficienti differenti. La proprietà di avere un segnale che correla scarsamente le sue traslazioni ed altri segnali simili, permette ai ricevitori GPS di identificare facilmente i segnali di satelliti GPS individuali e di sincronizzarsi con essi. Il segnale prodotto da un registro a scorrimento lineare può essere immaginato come una sequenza di zeri e di uno. Il registro stesso può essere immaginato come un nastro di k caselle contenenti gli elementi $s_{n-1}, \ldots, s_{n-r}$, ognuna delle quali possiede un valore che può essere 0 o 1. Ad ogni casella viene associato un numero $a_i \in \{0, 1\}$. I k valori a_i

sono fissati e distinti per tutti i satelliti. Generiamo una sequenza casuale nel modo seguente:

- Scegliamo un insieme di condizioni iniziali $s_0, \ldots, s_{k-1} \in \{0, 1\}$, non tutti nulli.
- Dati $s_{n-k}, \ldots, s_{n-1}$, il registro calcola l'elemento successivo nella sequenza in questo modo

$$s_n \equiv s_{n-k}a_0 + s_{n-k+1}a_1 + \cdots + s_{n-1}a_{k-1} = \sum_{i=0}^{k-1} a_{n-k+i}a_i (\text{mod } 2).$$

- Trasliamo ogni elemento verso destra, eliminando s_{n-k}. Il valore calcolato s_n viene inserito nella casella più a sinistra.
- Iteriamo il procedimento precedente.

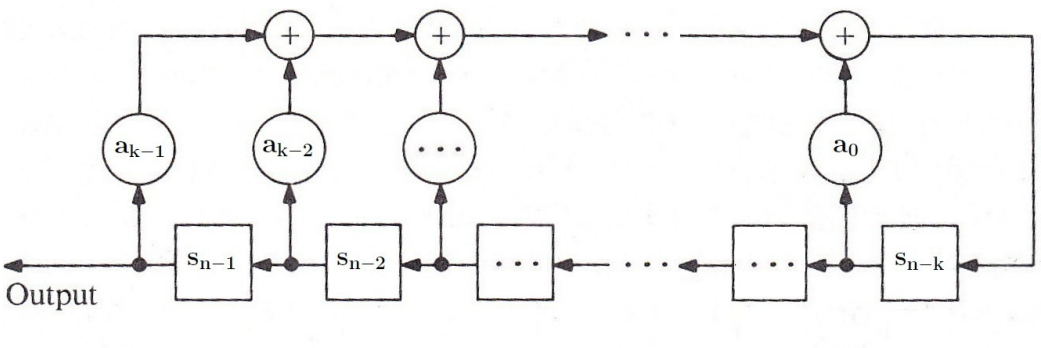

Figura 3.5: Registro a scorrimento lineare per il segnale GPS

Poiché la procedura sopra è perfettamente deterministica ed il numero di condizioni iniziali è finito, genereremo una sequenza che deve diventare periodica. Allo scopo di generare una sequenza con proprietà interessanti, dobbiamo soltanto scegliere i coefficienti $a_0, \ldots, a_{k-1} \in \{0, 1\}$ e le condizioni iniziali $s_0, \ldots, s_{k-1} \in \{0, 1\}$ in modo opportuno.
Non vedremo mai la sequenza completa, ma osserveremo soltanto una finestra di $M = 2^k - 1$ elementi consecutivi $\{s_n\}_{n=m}^{n=m+M-1}$, che chiamiamo $\mathbf{t}=$

$\{t_1, \ldots, t_M\}$. Vogliamo confrontarla con un'altra finestra $\mathbf{v} = \{v_1, \ldots, v_M\}$ della forma $\{s_n\}_{n=p}^{n=p+M-1}$. Per esempio: la sequenza $\mathbf{t}$ viene inviata dal satellite e la sequenza $\mathbf{v}$ è una traslazione ciclica della stessa sequenza generata dal ricevitore GPS. Per determinare la traslazione tra le due, il ricevitore trasla ripetutamente la sua sequenza di un'unità fin quando non diventa identica a $\mathbf{t}$.

Definizione 3.5 (Correlazione)**.**
La *correlazione* tra due sequenze $\mathbf{t}$ e $\mathbf{v}$ di lunghezza M è il numero di elementi i dove $t_i = v_i$ meno il numero di elementi i dove $t_i \neq v_i$. Denotiamo la correlazione con Cor($\mathbf{t}$,$\mathbf{v}$).

Osservazione 4. Se il registro consiste di k elementi, allora la correlazione tra ogni coppia di sequenze $\mathbf{t}$ e $\mathbf{v}$ deve soddisfare $-M \leq \text{Cor}(\mathbf{t},\mathbf{v}) \leq M$. Diciamo che le sequenze sono poco correlate se Cor($\mathbf{t}$,$\mathbf{v}$) è vicino a zero.

Proposizione 3.2.1. *La correlazione tra due sequenze è data da*

$$\text{Cor}(\mathbf{t},\mathbf{v}) = \sum_{i=1}^{M} (-1)^{t_i}(-1)^{v_i}.$$

Dimostrazione. Il numero Cor($\mathbf{t}$,$\mathbf{v}$) viene calcolato come segue: ogni volta che $t_i = v_i$ dobbiamo aggiungere 1. Allo stesso modo, ogni volta che $t_i \neq v_i$ dobbiamo sottrarre 1. Quindi, se $t_i = v_i$, allora o $(-1)^{t_i} = (-1)^{v_i} = 1$ o $(-1)^{t_i} = (-1)^{v_i} = -1$. In ogni caso vediamo che $(-1)^{t_i}(-1)^{v_i} = 1$. Analogamente, se $t_i \neq v_i$, esattamente uno di $(-1)^{t_i}$ e $(-1)^{v_i}$ è uguale a 1 e l'altro a -1. Quindi $(-1)^{t_i}(-1)^{v_i} = -1$. □

Il teorema seguente ci mostra che possiamo inizializzare un registro a scorrimento lineare in modo tale da fargli generare una sequenza poco correlata con ogni traslazione di se stessa.

Teorema 3.2.2. *Dato un registro a scorrimento lineare, esistono dei coefficienti $a_0, \ldots, a_{k-1} \in \{0,1\}$ e delle condizioni iniziali $s_0, \ldots, s_{k-1} \in \{0,1\}$, tali che la sequenza generata dal registro abbia un periodo di lunghezza $2^k - 1$. Consideriamo due finestre $\mathbf{t}$ e $\mathbf{v}$ di questa sequenza di lunghezza $M = 2^k - 1$,*

dove $\mathbf{t} = \{s_n\}_{n=m}^{n=m+M-1}$ *e* $\mathbf{v} = \{s_n\}_{n=p}^{n=p+M-1}$ *con* $p > m$. *Se* M *non divide* $p - m$, *allora*

$$\mathrm{Cor}(\mathbf{t},\mathbf{v}) = -1.$$

In altri termini, il numero di bit discordanti è sempre uno in più rispetto al numero di bit concordanti.

Vedremo tra poco la dimostrazione di questo teorema.

Esempio 3.4. Prendiamo $k = 4$, $(a_0, \ldots, a_3) = (1, 1, 0, 0)$ e $(s_0, \ldots, s_3) = (0, 0, 0, 1)$. Questi valori generano una sequenza di periodo $2^4 - 1 = 15$ che ripete il seguente blocco di simboli

$$0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad .$$

Se trasliamo la sequenza a sinistra di un simbolo, mandiamo il primo 0 alla fine, ottenendo

$$0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad .$$

Notiamo che i due blocchi di simboli differiscono alle posizioni 3, 4, 6, 8, 9, 10, 11 e 15. Quindi, differiscono in otto posizioni e combaciano in sette posizioni, fornendo una correlazione pari a -1.
Allo scopo di calcolare la correlazione con le altre 14 traslazioni della sequenza, scriviamo esplicitamente tutte le possibili traslazioni. Una verifica diretta mostra che due qualsiasi delle seguenti sequenze combaciano esattamente in

sette posizioni e differiscono nelle rimanenti otto.

0	0	0	1	0	0	1	1	0	1	0	1	1	1	1
0	0	1	0	0	1	1	0	1	0	1	1	1	1	0
0	1	0	0	1	1	0	1	0	1	1	1	1	0	0
1	0	0	1	1	0	1	0	1	1	1	1	0	0	0
0	0	1	1	0	1	0	1	1	1	1	0	0	0	1
0	1	1	0	1	0	1	1	1	1	0	0	0	1	0
1	1	0	1	0	1	1	1	1	0	0	0	1	0	0
1	0	1	0	1	1	1	1	0	0	0	1	0	0	1
0	1	0	1	1	1	1	0	0	0	1	0	0	1	1
1	0	1	1	1	1	0	0	0	1	0	0	1	1	0
0	1	1	1	1	0	0	0	1	0	0	1	1	0	1
1	1	1	1	0	0	0	1	0	0	1	1	0	1	0
1	1	1	0	0	0	1	0	0	1	1	0	1	0	1
1	1	0	0	0	1	0	0	1	1	0	1	0	1	1
1	0	0	0	1	0	0	1	1	0	1	0	1	1	1

Abbiamo visto nel capitolo precedente la struttura e la costruzione di campi finiti di ordine p^n (si rimanda il lettore ai paragrafi 2.2 e 2.3.2). Vediamo ora gli elementi di $\mathbb{F}_2^k$: sono le k-uple $(t_0, \ldots, t_{k-1})$ dove $t_i \in \{0, 1\}$. L'addizione di due k-uple siffatte è semplicemente l'addizione modulo 2, definita componente per componente. Per definire un'operazione moltiplicazione scegliamo un polinomio irriducibile

$$P(x) = x^k + p_{k-1}x^{k-1} + \cdots + p_1 x + p_0$$

sul campo $\mathbb{F}_2$. Interpretiamo ogni k-upla $(t_0, \ldots, t_{k-1})$ come un polinomio di grado inferiore o uguale a $k-1$. Per moltiplicare le due k-uple moltiplichiamo i due polinomi associati. Il prodotto è un polinomio di grado inferiore o uguale a $2(k-1)$, che è allora ridotto ad un polinomio in x di grado $k-1$ ottenuto prendendo il resto della divisione per P. Questo equivale ad applicare la regola $P(x) = 0$, cioè $x^k = p_{k-1}x^{k-1} + \cdots + p_1 x + p_0$ ed iterando. Interpretiamo allora i coefficienti del risultante polinomio di grado $k-1$

come le componenti di una k-upla. Abbiamo esplicitamente mostrato che gli elementi vettoriali di $\mathbb{F}_2^k$ possono essere interpretati come polinomi.
La funzione traccia (si veda Definizione 2.11) $T : \mathbb{F}_{2^k} \to \mathbb{F}_2$ è data da

$$T(t_{k-1}x^{k-1} + \cdots + t_1x + t_0) = t_{k-1}.$$

Proposizione 3.2.3. (i) *La funzione T è lineare e suriettiva.*

(ii) *Sia q il numero di elementi di $\mathbb{F}_{2^k}$. T vale 0 su esattamente $q/2$ elementi di $\mathbb{F}_{2^k}$ e vale 1 sui restanti $q/2$ elementi.*

Dimostrazione. (i) Si veda il Teorema 2.4.11, proprietà (i), (ii), (iii).
(ii) Abbiamo visto che $\mathbb{F}_{2^k} \cong \mathbb{F}_2^k$. Sia $q = 2^k$ il numero di elementi di $\mathbb{F}_2^k$. T è lineare suriettiva per la (i), quindi t_{k-1} vale 0 su $2^k/2$ elementi di $\mathbb{F}_2^k$ e vale 1 sui restanti $2^k/2$ elementi. □

Dimostrazione. (Teorema 3.2.2) Scegliamo un polinomio primitivo $P(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$ su $\mathbb{F}_2$, che ci permetterà di costruire il campo $\mathbb{F}_{2^k}$. Gli a_i del registro a scorrimento lineare sono i coefficienti del polinomio $P(x)$. Al fine di costruire delle buone condizioni iniziali scegliamo un qualsiasi polinomio non nullo $t = t_{k-1}x^{k-1} + \cdots + t_1x + t_0$ da $\mathbb{F}_{2^k}$. Definiamo le condizioni iniziali come

$$\begin{aligned} s_0 &= T(t) &&= t_{k-1}, \\ s_1 &= T(xt), \\ &\vdots \\ s_{k-1} &= T(x^{k-1}t). \end{aligned} \tag{3.9}$$

Consideriamo come viene calcolato il valore di s_1:

$$\begin{aligned} s_1 = T(tx) &= T(t_{k-1}x^k + t_{k-2}x^{k-1} + \cdots + t_0x) \\ &= T(t_{k-1}(t_{k-1}x^{k-1} + \cdots + a_1x + a_0) + t_{k-2}x^{k-1} + \cdots + t_0x) \\ &= T((t_{k-1}a_{k-1} + t_{k-2})x^{k-1} + \cdots \\ &= t_{k-1}a_{k-1} + t_{k-2}. \end{aligned}$$

Un calcolo simile permette la determinazione dei valori $s_2, \ldots, s_{k-1}$.
Prima di proseguire con la dimostrazione, consideriamo il seguente esempio.

Esempio 3.5. Nell'esempio 3.4 avevamo usato il polinomio $P(x) = x^4 + x + 1$. Come polinomio t avevamo preso semplicemente $t = 1$. Questo crea le seguenti condizioni iniziali:

$$\begin{aligned} s_0 &= T(1) = 0 \\ s_1 &= T(x) = 0 \\ s_2 &= T(x^2) = 0 \\ s_3 &= T(x^3) = 1. \end{aligned}$$

Proposizione 3.2.4. *Scegliamo i coefficienti $a_0, \ldots, a_{k-1}$ di un registro a scorrimento come quelli di un polinomio primitivo $P(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. Sia $t = t_{k-1}x^{k-1} + \cdots + t_1 x + t_0$. Scegliamo gli elementi iniziali $s_0, \ldots, s_{k-1}$ come nella* (3.9)*. Allora la sequenza $\{s_n\}_{n\geq 0}$ generata dal registro a scorrimento è data da $s_n = T(x^n t)$ e si ripete con periodo che divide $2^k - 1$.*

Dimostrazione. Usiamo il fatto che $P(x) = 0$, che è come dire $x^k = a_{k-1}x^{k-1} + \cdots + a_1 x + a_0$. Allora

$$\begin{aligned} T(x^k t) &= T((a_{k-1}x^{k-1} + \cdots + a_1 x + a_0)t) \\ &= a_{k-1}T(x_{k-1}t) + \cdots + a_1 T(xt) + a_0 T(t) \\ &= a_{k-1}a_{k-1} + \cdots + a_1 s_1 + a_0 s_0 \\ &= s_k. \end{aligned}$$

Procediamo per induzione. Supponiamo che gli elementi della sequenza soddisfino $s_i = T(x^i t)$ per $i \leq n - 1$. Allora

$$\begin{aligned} T(x^n t) = T(x^k x^{n-k} t) &= T((a_{k-1}x^{k-1} + \cdots + a_1 x + a_0)x^{n-k}t) \\ &= a_{k-1}T(x_{n-1}t) + \cdots + a_1 T(x^{n-k+1}t) + a_0 T(x^{n-k}t) \\ &= a_{k-1}s_{n-1} + \cdots + a_1 s_{n-k+1} + a_0 s_{n-k} \\ &= s_n. \end{aligned}$$

Quindi la moltiplicazione per x corrisponde esattamente al calcolo effettuato dal registro a scorrimento e $s_n = T(x^n t)$ per ogni n. Vediamo immediatamente che il periodo minimo può essere al più $2^k - 1$, poiché $x^{2^k-1} = 1$. □

Vogliamo vedere che il periodo minimo di questa sequenza è effettivamente $2^k - 1$: tanto per cominciare deve essere un divisore di $2^k - 1$. Infatti, come visto al Teorema 3.1.8, il periodo minimo è esattamente $2^k - 1$ quando P è primitivo. Non dimentichiamoci che il nostro scopo principale è quello di creare sequenze poco correlate con traslazioni di se stesse. Calcoliamo quindi la correlazione tra due finestre qualsiasi $\mathbf{t}$ e $\mathbf{v}$ di lunghezza $M = 2^k - 1$, con $\mathbf{t}= \{s_n\}_{n=m}^{n=m+M-1}$ e $\mathbf{v} = \{s_n\}_{n=p}^{n=p+M-1}$.

Proposizione 3.2.5. *Se* $\mathbf{t}= \{s_n\}_{n=m}^{n=m+M-1}$ *e* $\mathbf{v}= \{s_n\}_{n=p}^{n=p+M-1}$*, allora* $\mathrm{Cor}(\mathbf{t},\mathbf{v})= -1$ *se* M *non divide* $p - m$*.*

Dimostrazione. Possiamo supporre $m \leq p$. Allora

$$\begin{aligned}
\mathrm{Cor}(\mathbf{t}, \mathbf{v}) &= \textstyle\sum_{i=0}^{M-1}(-1)^{s_{m+i}}(-1)^{s_{p+i}} \\
&= \textstyle\sum_{i=0}^{M-1}(-1)^{T(x^{m+i}t)}(-1)^{T(x^{p+i}t)} \\
&= \textstyle\sum_{i=0}^{M-1}(-1)^{T(x^{m+i}t)+T(x^{p+i}t)} \\
&= \textstyle\sum_{i=0}^{M-1}(-1)^{T(x^{m+i}t+x^{p+i}t)} \\
&= \textstyle\sum_{i=0}^{M-1}(-1)^{T(tx^{i+m}(1+x^{p-m}))} \\
&= \textstyle\sum_{i=0}^{M-1}(-1)^{T(x^{i+m}\tau)},
\end{aligned}$$

dove $\tau = t(1+x^{p-m})$. Per la nostra scelta di P sappiamo che x è un elemento primitivo del nostro campo e quindi che $x^M = 1$ e $x^N \neq 1$ se $1 \leq N < M$. Deduciamo che $x^N = 1$ se e solo se M divide N.
Se M divide $p - m$ allora $x^{p-m} = 1$ e $\tau = t(1 + 1) = t \cdot 0 = 0$ ed in questo caso $\mathrm{Cor}(\mathbf{t},\mathbf{v})= M$.
Se M non divide $p - m$, allora il polinomio $(1 + x^{p-m})$ non è il polinomio nullo; quindi anche $\tau = t(1 + x^{p-m})$ è non nullo, poiché è il prodotto di due elementi non nulli.
Così τ è della forma x^r, dove $r \in \{0, \ldots, 2^k - 2\}$, questo implica che l'insieme $\{\tau x^{i+m}, 0 \leq i \leq M - 1\}$ è una permutazione degli elementi di $\mathbb{F}_{2^k}^{\times} = \{1, x, \ldots, x^{2^k-2}\}$. La funzione traccia T assume il valore 1 su metà degli elementi di $\mathbb{F}_{2^k}$ ed il valore 0 sui rimanenti elementi. Poiché assume il valore 0 sull'elemento zero, assume valore 1 su 2^{k-1} degli elementi di $\mathbb{F}_{2^k}^{\times}$ e valore 0 sui rimanenti $2^{k-1} - 1$. Quindi $\mathrm{Cor}(\mathbf{t},\mathbf{v})= -1$. □

Corollario 3.2.6. *Il periodo della sequenza pseudo-casuale generata dal registro a scorrimento lineare è esattamente* $M = 2^k - 1$.

Dimostrazione. Se il periodo fosse uguale a $K < M$, allora la sequenza coinciderebbe con la sua traslazione per K elementi, e le due sequenze avrebbero una correlazione uguale ad M. Questo è in contraddizione con la proposizione precedente. □

Abbiamo così concluso la dimostrazione del Teorema 3.2.2.
Se ora volessimo generare altre sequenze pseudo-casuali della stessa lunghezza, potremmo usare lo stesso principio ma cambiando il polinomio $P(x)$, poiché vogliamo una sequenza distinta per ogni satellite. La teoria di Galois vista al paragrafo 2.4.4, ci permette di calcolare (in alcuni casi) la correlazione di questa nuova sequenza con la prima e le sue traslazioni. Gli ingegneri, comunque, si accontentano di verificare questi valori della correlazione in tabelle precalcolate.

Appendice A

Come costruire una tabella delle radici del polinomio universale

Ci sono due metodi per costruire una tabella delle radici del polinomio universale (un esempio appare nel paragrafo 2.3.5):

(i) Dato il polinomio $x^{p^n} - x$ ed $r(x)$, un polinomio monico irriducibile di grado n in $\mathbb{F}_p[x]$, con cui definire l'aritmetica per un $\mathbb{F}_{p^n}$, usiamo l'algoritmo di Berlekamp (si veda 2.3.5) per fattorizzare $x^{p^n} - x$ in $\mathbb{F}_p[x]$. Allora, per ogni fattore monico irriducibile risultante, usiamo l'algoritmo di Berlekamp in $\mathbb{F}_{p^n}[x]$, oppure proviamo per tentativi su tutti gli elementi di $\mathbb{F}_{p^n}$ di trovare radici in $\mathbb{F}_{p^n}$ per ognuno dei fattori risultanti.

(ii) Usiamo la teoria di Galois, come descritta nel paragrafo 2.4.4:

- Scriviamo una tabella dei logaritmi per $\mathbb{F}_{p^n}$.
- Usiamo il criterio del sottocampo attraverso i logaritmi per trovare elementi di $\mathbb{F}_{p^m}$, per divisori crescenti m di n, cioè elementi il cui logaritmo è un multiplo di $(p^n - 1)/(p^m - 1)$.

- Per ogni elemento risultante a di $\mathbb{F}_{p^m}$, elenchiamo gli m elementi distinti dell'orbita di Frobenius $a, a^p, \ldots, a^{p^{m-1}}$. Le cancelliamo dalla tabella dei logaritmi, così le prenderemo in considerazione una sola volta. Quindi calcoliamo il prodotto

$$\prod_{i=1}^{m}(x - a^{p^i}).$$

  Questo è il polinomio minimo per tutti gli m elementi nell'orbita di Frobenius. Scriviamo questo polinomio minimo seguito dalle m radici.

- Una volta cancellati tutti gli elementi nella tabella dei logaritmi, ordiniamo le righe della tabella delle radici lessicalmente in base ai polinomi minimi nella colonna a sinistra.

Bibliografia

[1] Alfke P. *Efficient shift registers, LFSR counters, and long pseudo-random sequence generators.* Xilinx Corporation, 1996.
`http://www.xilinx.com/support/documentation/application_notes/ xapp052.pdf`

[2] Argento D. *Registri a scorrimento lineari.* 2004.
`http://www.dia.uniroma3.it/ dispense/rota/LFSR.pdf`

[3] Berlekamp E.R. *Algebraic coding theory.* McGraw-Hill Book Company, New York, 1968.

[4] Biagi L. *I fondamentali del GPS.* Geomatics Workbooks, Volume 8. Politecnico di Milano, 2009.
`http://geomatica.como.polimi.it/workbooks/n8/`

[5] Conte A., Picco Botta L. e Romagnoli D. *Algebra.*
Editrice Levrotto & Bella, Torino, 1986.

[6] Daemen J. and Rijmen V. *The Rijndael Block Cipher.*
`http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf`

[7] *Finite fields, coding theory, and advances in communications and computing.* Proceedings of the International Conference held at the University of Nevada, Las Vegas, Nevada, August 7-10, 1991.
Edited by Gary L. Mullen and Peter Jau-Shyong Shiue. Lecture Notes in Pure and Applied Mathematics, 141. Marcel Dekker, Inc., New York, 1993.

[8] Herstein I.N. *Algebra.* Editori Riuniti, 2003.

[9] Kerl J. *Computation in finite fields.*
Arizona State University and Lockheed Martin Corporation. April, 2004.
http://johnkerl.org/doc/ffcomp.pdf

[10] Klapper A. and Goresky M. *Feedback Shift Registers, 2-Adic Span, and Combiners with Memory.* Journal of Cryptology, 1997.
http://www.math.ias.edu/ goresky/pdf/2adic.jour.pdf

[11] Lang S. *Algebra.* Springer-Verlag, 2002.

[12] Lidl R. and Niederreiter H. *Finite fields*, in: Encyclopedia of mathematics and its applications, Volume 20. Cambridge University Press, 1983.

[13] Mullen G.L. and Mummert C. *Finite fields and applications.* Student mathematical library, Volume 41, 2008.

[14] Murphy T. *Course 373, Finite Fields.*
http://pet.ece.iisc.ernet.in/sathish/FiniteFields.pdf

[15] Neve A.C. *Generazione ed uso dei numeri pseudocasuali.*
http://www.fermilecce.it/index.php?option=com_docman&task=cat_view&gid=77&Itemid=92

[16] *Registro a scorrimento a retroazione lineare.*
http://it.wikipedia.org/wiki/Registro_a_scorrimento_a_retroazione_lineare

[17] Rossi G., Bottoni F. e Albanese G. *Generatori di Numeri Pseudorandom, "la generazione controllata della casualità".* Università degli Studi Roma Tre, Corso di Laurea Magistrale in Ingegneria delle Tecnologie della Comunicazione e dell'Informazione, Progetto di Crittografia, Prof.ssa F.Merola, a.a. 2009/2010.

[18] Rousseau C. and Saint-Aubin Y. *Mathematics and Technology.* Springer, 2010.

[19] Rousseau C. und Saint-Aubin Y. *Mathematik und Technologie.* Springer, 2012.

[20] Tsaban B. and Vishne U. *Efficient linear feedback shift registers with maximal period.* http://arxiv.org/pdf/cs/0304010.pdf

[21] Van der Waerden B.L. *Algebra I.* Springer-Verlag, Berlin-Göttingen-Heidelberg, 1955.

[22] Vistoli A. *Note di Algebra.* Centro Copie Bononia, Bologna. 1993-94 (Non pubblicato).

[23] Williams R. *A painless guide to CRC error detection algorithms.* http://www.cse.sc.edu/ jimdavis/Courses/2004-Fall %20CSCE%20491/crc-Ross.pdf

[24] Zogg J.M. *GPS - Essentials of Satellite Navigation.* u-blox Holding AG, 2009. http://www.u-blox.com/images/downloads/Product_Docs/GPS_ Compendium%28GPS-X-02007%29.pdf

[25] Zogg J.M. *GPS und GNSS: Grundlagen der Ortung und Navigation mit Satelliten.* u-blox Holding AG, Ottobre 2011. http://zogg-jm.ch/Dateien/Update_Zogg_Deutsche_Version_ Jan_09_Version_Z4x.pdf

Lista dei simboli

$\mathbb{N}$	Insieme dei numeri naturali (interi positivi)
$\mathbb{Z}$	Insieme degli interi
$\mathbb{R}$	Insieme dei numeri reali
$\mathbb{C}$	Insieme dei numeri complessi
A^n	Insieme di tutte le n-uple $(a_1, \ldots, a_n)$ con $a_i \in A$ per $1 \leq i \leq n$
$\lvert A \rvert$	Cardinalità (= numero degli elementi) dell'insieme finito A
$[a]_p$	Classe d'equivalenza di a mod p
$\log_a(z) = k$	Logaritmo discreto in base a di $z = a^k \in \mathbb{F}_p^\times,\ k \in \mathbb{Z}/(p-1)$
$\lfloor q \rfloor$	Il più grande intero $\leq q \in \mathbb{R}$
$\mathrm{mcd}(k_1, \ldots, k_n)$	Massimo comun divisore di $(k_1, \ldots, k_n)$
$a \equiv b \bmod n$	a congruente a b modulo n
$\varphi(n)$	Funzione di Eulero di n
$\mu(n)$	Funzione di Möbius di n
$\mathbb{Z}_n$	Gruppo degli interi modulo n
$GL(r, \mathbb{F}_q)$	Gruppo lineare generale di matrici $r \times r$ non singolari su $\mathbb{F}_q$
$\lvert G \rvert$	Ordine del gruppo finito G
$\langle r(x) \rangle$	Ideale generato dal polinomio $r(x)$ (se l'anello è sottointeso)
$(a) = a\mathbb{Z}$	Ideale principale generato da $a \in \mathbb{Z}$
$[a]_n, a + (n)$	Classe residuo dell'elemento a dell'anello, modulo l'ideale (n), è un elemento di $\mathbb{Z}/(n)$
$\mathbb{Z}/(n)$	Anello degli interi modulo n

$F[x]$	Anello polinomiale sul campo F
$\deg(f)$	Grado di f
f'	Derivata del polinomio f
$K(M)$	Estensione del campo K ottenuta aggiungendo M
$[L\colon K]$	Grado del campo L sul campo K
$\mathbb{F}_q$	Campo finito di ordine q
$\mathbb{F}_q^\times$	Gruppo moltiplicativo degli elementi non nulli di $\mathbb{F}_q$
$T(\alpha)$	Traccia di $\alpha \in F$ su K
$f(n)$	Numero di polinomi monici irriducibili in $\mathbb{F}_q[x]$ di grado n