NTNU
Det skapende universitet

Kodeteori i et historisk perspektiv

Brage Sæth

Master i realfag
Innlevert: juli 2016
Hovedveileder: Sverre Olaf Smalø, MATH

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Kodeteori i et historisk perspektiv

Brage Sæth

Forord

Først og fremst ønsker jeg å takke min veileder, Sverre Olaf Smalø, for gode råd og forklaringer. Uten han hadde ikke denne oppgaven blitt ferdig.

Jeg ønsker også å takke datteren min, Hanna, for å ha lyst opp dagene mine med sitt gode humør og kunstverk til å ha på pulten. Takk til Øistein Søvik og Astrid Nerhus Dale som også har vært med på å gjøre livet på matteland hyggeligere de siste månedene av denne perioden.

Takk til alle venner og familiemedlemmer som har støttet meg gjennom studietiden.

Sammendrag

For tiden er det ikke stor framgang i utviklingen av kodeteori. Hver nye kode som blir utviklet er eventuelt bare marginalt bedre enn de gamle. Denne oppgaven gir en oversikt for kodeteoriens utvikling over tre kapitler. I det første kapitlet introduseres problemet som førte til utviklingen av kodeteori. Det andre kapittelet ser på den matematiske bakgrunnen som legger grunnlaget for en del av kodeteorien. Det siste kapitellet ser på utviklingen av ulike kodetyper, eksempler på disse og litt mer matematisk begrunnelse.

Summary

Nowadays there is no substantial progress in the development of coding theory. Each new code being developed, is only marginally better than the old ones. This paper gives an overview of the development of coding theory over three chapters. The first chapter introduces the problem that led to the development of coding theory. The second chapter looks at the mathematical background that forms the basis for part of the code theory. The last chapter takes a look at the development of different types of code, examples of these and a bit more mathematical reasoning.

Innhold

Kapittel 1

Historisk bakgrunn

I kodeteoriens barndom var hovedfokuset å kunne sende beskjeder fra et sted til et annet uten at uønskede personer hadde mulighet til å lese eller endre beskjeden. I det 5 århundret f.Kr utviklet spartanerene det de kalte en *skytale* [Se [Sin00]], som var en stokk med bestemt diameter som man surret en strimle lær rundt og skrev på læret langs stokken. Dermed kan en bare lese beskjeden dersom en selv hadde en stokk av samme tykkelse som den opprinnelige skytalen.

En beskjed som:

'angrip fienden fra nordøst ved daggry' skrevet på en gitt skytale:

a	n	g	r	i	p	f	i
e	n	d	e	n	f	r	a
n	o	r	d	ø	s	t	v
e	d	d	a	g	g	r	y

gir 'aenennodgdrdredainøgpfsgfrtriavy' på lærstrimlen.

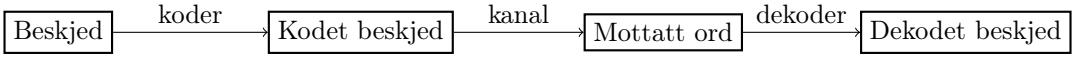
Etterhvert fikk denne grenen av matematikk navnet kryptografi. Som vi skal se handler ikke kodeteori om hemmelighold, men om å bevare informasjon.

1.1 Bevare informasjon

Claude Shannon, som er regnet blant grunnleggerene av informasjonsteori, hvor kodeteori er anvendt, gjorde seg denne bemererkningen;

> Det fundamentale problemet med kommunikasjon er det å kunne på et sted reprodusere en beskjed, enten den eksakte eller en tilnærming av beskjeden som ble sendt fra et annet sted.
>
> [Sha48, s. 379]

Kodeteori handler om å bevare informasjon. Hvis en skal sende en beskjed, kan man benytte en kode på den før den sendes, slik at selv om den utsettes for støy i kanalen den sendes over, så er mottaker i stand til å dekode beskjeden korrekt.



Figur 1.1: En beskjed sendt over en kanal

Det finnes flere former for kanaler avhengig av medium. En telefonledning kan utsettes for vibrasjoner som gir opphav til støy, mens det trådløse alternativet er utsatt for kosmisk stråling. Et krav til denne koden er at sannsynligheten for at et symbol kommer korrekt frem til mottaker ikke må være akkurat 50%, hvis ikke er det umumlig å sende noe som helst over kanalen. Det er ønskelig at denne sannsynligheten skal ligge nærmest mulig 100%.

En annen grunn til utviklingen av kodeteori, er for bruk i datamaskiner. Datamaskiner fra *Bells Telephone Lab* skulle ofte arbeide uten tilsyn over lengre tid, på natten og i helgene [Se Ham50]. Derfor var det nødvendig at datamaskinene selv var i stand til å oppdage og eventuelt rette opp feil. Slike feil kan blant annet skyldes alfapartikler. Dette var et økende problem etterhvert som datamaskiner skulle utføre mer og mer beregninger. Men dette betyr at det er det samme problemet for både lagring og sending av data, så søker en felles løsning.

Datamaskiner benytter seg av det binære tallsystemet, forkortet bits. Ofte benytter en seg av benevningen byte for åtte bits. Dersom en datamaskin har åtte harddisker kan den skrive sju bits til sju av hardisskene, mens på den åttende harddisken noterer seg summen av de øvrige sju hardiskene. Dersom man hver dag undersøker om harddiskene er skadet, vet en om en fortsatt har informasjonen en har lagret eller ikke. Vi antar her at vi er istand til å oppdage hvilken harddisk som er skadet. Siden den åttende harddisken har informasjon om alle harddiskene, er en istand til å gjennskape den harddisken som går i stykker. Sannsynligheten for at to harddisker skal gå istykker er cirka kvadratet av sannsynligheten for at en skal gå i stykker, derfor må en vurdere hva sannsynligheten for dette skjer og om den er tilfredsstillende. Dersom en ikke er istand til å se hvilken harddisk som er skadet må en finne på noe smartere, hvilket er formålet med kodeteori.

1.2 Variabel lengde koder

Rundt 1840 ved utviklingen av telegrafen, ble Morsealfabetet utviklet av Samuel Morse og assistenten hans [Se [Mor09]]. Dette er en kode som gjør om bokstaver og siffer til sammensetninger av $\cdot$, $-$ og 'mellomrom' av ulik lengde. Denne koden var nødvendig ettersom telegrafen ikke kunne sende annet enn sammensetninger av $\cdot$, $-$ og 'mellomrom'.

$$\begin{cases} \cdot & \text{lengde } 1 \\ - & \text{lengde } 3 \\ \text{'mellomrom' mellom } \cdot \textit{ og} - & \text{lengde } 1 \\ \text{'mellomrom' mellom bokstaver} & \text{lengde } 3 \\ \text{'mellomrom' mellom ord} & \text{lengde } 7 \end{cases}$$

Ved utvikling av denne koden foretok de en grov frekvensanalyse på hyppigheten av bokstaver i det engelske språket, og satt det opp slik at bokstaver som forekom hyppig, fikk et kort kodeord, mens de sjeldne hadde lengre kodeord. Eksempel på dette er E som ble kodet '$\cdot$', mens X ble kodet '$- \cdot\cdot -$', så X tar 11 ganger så lang tid som E å sende.

1.3 Kodeteori i Norge

Kodeteori er knyttet til alle personer her i Norge i form av personnummer [se [Sel64]]. På tidlig 1960-tallet valgte Statistisk Sentralbyrå (SSB) at folketallet skulle registreres med egne nummer for å lettere bearbeide folkeregisteret som da skulle overføres fra papirformat til magnetbånd. Hver person skulle ha et unikt nummer og det ble nødvendig å ha med minst ni siffer. De seks første siffrene angir fødselsdatoen til personen, hvor de to første angir fødselsdag, måned og år. De tre neste siffrene angir individualnummer for hver fødselsdato, hvor det det siste av disse er oddetall for menn og partall for kvinner. Dette personnummeret blir da på formen:

$$d_1 d_2 m_1 m_2 å_1 å_2 n_1 n_2 n_3$$

Sannsynligheten for feil er stor med kun 9 siffer, da det er mulighet for feil med at de som behandler informasjonen stempler feil i hullkortene, eller at de som oppgir opplysningene sier feil. Derfor eksisterer det to kontrollsiffer k_1 og k_2 som er i stand til å finne alle enslige og nesten alle doble feil: Disse feilene finner en ved å undersøke om

$$\begin{pmatrix} d_1 & d_2 & m_1 & m_2 & å_1 & å_2 & n_1 & n_2 & n_3 & k_1 & k_2 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 7 & 4 \\ 6 & 3 \\ 1 & 2 \\ 8 & 7 \\ 9 & 6 \\ 4 & 5 \\ 5 & 4 \\ 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} = (0,0) \pmod{11}$$

Dersom det skjer en feil i fjerde og tiende siffer slik at m_2 blir endret til m_2' hvor $m_2' = m_2 + x_1$ og k_1 blir endret til k_1' hvor $k_1' = k_1 + x_2$ kan feilen gå uoppdaget hvis:

$$\left(\begin{pmatrix} d_1 & d_2 & m_1 & m_2 & å_1 & å_2 & n_1 & n_2 & n_3 & k_1 & k_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & x_1 & 0 & 0 & 0 & 0 & 0 & x_2 & 0 \end{pmatrix} \right) \begin{pmatrix} 3 & 5 \\ 7 & 4 \\ 6 & 3 \\ 1 & 2 \\ 8 & 7 \\ 9 & 6 \\ 4 & 5 \\ 5 & 4 \\ 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \end{pmatrix} + \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \pmod{11}$$

Så høyresiden må bli null modulo 11 i begge ligningene, hvilket her er tilfelle så lenge $x_1 = -x_2 \pmod{11}$

Ola Nordmann ble født 17 mai 1946 med individualnummer 225. Dette gir kontrollnummer som tilfredstiller:

$$\begin{pmatrix} 1 & 7 & 0 & 5 & 4 & 6 & 2 & 2 & 5 & k_1 & k_2 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 7 & 4 \\ 6 & 3 \\ 1 & 2 \\ 8 & 7 \\ 9 & 6 \\ 4 & 5 \\ 5 & 4 \\ 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 6+k_1 & 8+2k_1+k_2 \end{pmatrix} \pmod{11} = \begin{pmatrix} 0 & 0 \end{pmatrix} \pmod{11}$$

Altså må $k_1 = 5$ og $k_2 = 4$.

En del eldre personer sier ofte fødselsår som nittenseksogførti, noe som kan resultere i at Ola skrev ned personnummeret sitt feil som; 17056422554

$$\begin{pmatrix} 1 & 7 & 0 & 5 & 6 & 4 & 2 & 2 & 5 & 5 & 4 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 7 & 4 \\ 6 & 3 \\ 1 & 2 \\ 8 & 7 \\ 9 & 6 \\ 4 & 5 \\ 5 & 4 \\ 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 2 \end{pmatrix} \pmod{11}$$

Så her oppdager en at det har skjedd en lesefeil.

Kapittel 2

Matematisk bakgrunn

2.1 Kongruens

I 1801 ga Gauss ut boken *Disquistiones Arthmeticae* (lat: undersøkelse av aritmetikk) hvor kongruensteori først dukket opp [Se [Bur11]]. Kongruens ble innført med symbolet $\equiv$ og beskriver sammenhengen mellom tre heltall.

Definisjon 2.1.1. *For et gitt positivt heltall n, så sier en at to heltall a og b er kongruent modulo n:*

$$a \equiv b \pmod{n}$$

hvis $n \mid a - b$.

Kongruens sees på som rest ved divisjon. For et heltall a, med kvotient q og rest r ved divisjon av n

$$a = qn + r, \text{ for } 0 \leq r < n$$

Så per definisjon av kongruens, så er $a \equiv r \pmod{n}$, hvor $r \in \{0, 1, \ldots, n-1\}$. Mengden av disse n verdiene danner en komplett mengde av restklasser modulo n. Mengden av alle heltall modulo n kan skrives $\mathbb{Z}_n$. Som regel med:

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$$

med operasjonene $+_n$ og $\cdot_n$ definert som

$$a +_n b = (a + b) \pmod{n} \text{ og } a \cdot_n b = (a \cdot b) \pmod{n}$$

2.2 Endelig kropper

En kropp er en mengde K med to opersjoner $+$ og $\cdot$ (addisjon og multiplikasjon) som tilfredsstiller følgende egenskaper [pr [Fra03]]:

- $(K, +)$ er en abelsk gruppe.
- $(K \setminus 0, \cdot)$ er en abelsk gruppe.
- Multiplikasjon er distributativ over addisjon.

$$\text{For } a, b, c \in K : a \cdot (b + c) = a \cdot b + a \cdot c$$

Mengden $\mathbb{Z}_n$ som innført i forrige avsnitt blir en endelig kropp hvis og bare hvis n er et primtall.

Setning 2.2.1. *Enhver endelig undergruppe av $(K \setminus (0), \cdot)$ er en syklisk gruppe.*

Bevis. La G være en endelig undergruppe av $(K \setminus (0), \cdot)$ og $a \in G$.
Da må G også inneholde alle elementer på formen a^n hvor $n \in \mathbb{Z}$. Ettersom G er endelig må en ende opp med en verdi i slik at $a^i \equiv 1$, ergo er G syklisk. □

Korollar 2.2.2. *Dersom K er endelig, så er $K \setminus (0)$ syklisk.*

Definisjon 2.2.3. *Vi gir en kropp K med q elementer notasjonen K_q, eventuelt $GF(q)$. Gir også notasjonen $K_q^* = K_q \setminus (0)$.*

Det eksisterer et primitivt element α i en kropp $\mathbb{K}_q^*$ slik at

$$K_q = \{0, \alpha^0 = 1, \alpha, \alpha^2, \ldots, \alpha^{q-2}\}$$

Den binære kroppen $M_2(\mathbb{Z}_2)$ med fire elementer $\{\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)\}$ har følgende addisjon og multiplikasjonstabell:

$+$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$

og

$\cdot$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$
$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$

Her er f.eks $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$ et primitivt element siden;

$\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^0 = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^1 = \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^2 = \left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$ og $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^3 = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^0$

For hvert primtall p og positivt heltall n finnes en kropp med p^n elementer og enhver endelig kropp K har p^n elementer for et primtalll p og positivt heltall n.

Teorem 2.2.4 (Kronecker's teorem [PSS94])**.** *La K være en kropp og $f(x)$ ett ikkekonstant polynom i $K[x]$. Da eksisterer det en kroppsutvidelse E av K og $\alpha \in E$ slik at $f(\alpha) = 0$.*

For gitt primtall p og positivt heltall n, vil røttene av polynomet $x^{p^n} - x$ i $\mathbb{Z}_p[x]$ danner nøyaktig elementene av K_{p^n}. Vi kan skrive $x^{p^n} - x = \prod_i^{p^n}(x - \alpha^i)$, så polynomet blir et produkt av lineære faktorer i K_{p^n}

Korollar 2.2.5. *For å konstruere en kropp med p^n elementer tar en et irredusibelt polynom $f(x)$ av grad n over K_p og tar $K_{p^n} = K_p[x]/f(x)$.*

2.3 Vektorrom

Et vektorrom V over en kropp K består av de to operasjonene addisjon og skalar multiplikasjon slik at for alle elementer $x, y \in V$ så eksisterer det ett unikt element $x + y \in$V og for hver $a \in K$ så eksisterer det et unikt element $ax \in V$.

- $(V, +)$ er en abelsk gruppe.

- Multiplikasjon er distributativ over addisjon.

$$\begin{aligned} \text{For alle } x, y \in V \text{ og } a, b \in K : a \cdot (x + y) &= a \cdot x + a \cdot y \\ \text{og } (a + b) \cdot x &= a \cdot x + b \cdot x \end{aligned}$$

Vi definerer K^n som mengden av $n-$tupler $(a_1, \ldots, a_n)$ med elementer a_i fra kroppen K. En vektor i K^n blir i de fleste grener av lineær algebra skrevet som en kollonnevektor, $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, men det er blitt mer konvensjonelt å skrive det som en radvektor $v = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ i kodeteori.

En $m \times n$ matrise M med koeffisienter i en kropp K, er en funksjon fra $\{1, \ldots, m\} \times \{1, \ldots, n\}$ til kroppen K, vanligvis representeres dette ved en rektangulær tabell på formen;

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}, \text{ der } A_{21} = A(2, 1)$$

Her anser en radene som vektorer i K^n og kolonnene som vektorer i K^m.

Mengden av alle $m \times n$-matriser med elementer A fra en kropp K, med notasjon $M_{m \times n}(K)$, danner et vektorrom over K, med $A, B \in M_{m \times n}(K)$ og $c \in K$, når:

$$\begin{aligned} (A + B)(i, j) &= A(i, j) + B(i, j) \\ (cA)(i, j) &= c \cdot A(i, j) \end{aligned} \quad \forall (i, j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$$

Ett polynom med koeffisienter fra kroppen K er et uttrykk

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

hvor $a_i \in K$. Graden av polynomet er definert som den største eksponenten hvor koeffisienten er ulik null.

En delmengde W av et vektorrom V over en kropp K er kalt et underrom av V hvis W er et vektorrom over K med addisjon og skalarmultiplikasjon definert fra V.

La V være et vektorrom og S være en ikketom delmengde av V. En vektor $v \in V$ er en lineær kombinasjona av vektorer av S hvis det eksisterer et endelig

antall vektorer $u_1, u_2, \ldots, u_n \in S$ og skalarer $a_1, a_2, \ldots, a_n \in K$ slik at

$$v = a_1u_1 + a_2u_2 + a_nu_n.$$

Vektoren v er da en lineær kombinasjon av $u_1, \ldots, u_n$ med koeffisienter $a_1, \ldots, a_n$. Dersom enhver $v \in V$ kan skrives som en lineærkombinasjon av elementene på S, sier vi at S genererer V.

En delmengde S av V er lineært avhengig hvis det eksissterer en endelig mengde distinkte vektorer $u_1, \ldots, u_n$ i S og skalarer $a_1, \ldots, a_n$, ikke alle null, slik at

$$a_1u_1 + \cdots + a_nu_n = 0$$

Da sier vi at vektorerene av S er lineært avhengige. Dersom delmengden S ikke er lineært avhengig er den lineært uavhengig.

En basis $\mathcal{B}$ for V er en lineær uavhengig delmengde av V, som genererer V. Delmengden $\mathcal{B}$ er en basis hvis og bare hvis det til hver vektor $v \in V$, eksisterer en n, $u_1, \ldots, u_n$ entydig i $\mathcal{B}$ og $a_1, \ldots, a_n$ entydig i K slik at;

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

Merk at det eksisterer basiser for vektorrom av uendelig størrelse ($n \to \infty$), men disse er ikke interessant for oss her.

Alle basiser i et gitt vektorrom har like mange elementer. Dette unike tallet er definert som dimensjonen til V, med notasjon $dim(V)$. Dersom basisen har uendelig antall elementer sier en $dim(V) = \infty$. Ett vektorrom er endeligdimensjonalt hvis det har en basis som består av et endelig antall vektorer ($dim(V) < \infty$). Hvis basisen ikke har noen vektorer, definerer en $dim(V) = 0$.

La V og W være vektorrom og la $T : V \to W$ være en lineær avbildning. Definerer bildet av T, $Im(T)$, til å være delmengden av W som består av alle bilder av vektorer i V gjennom T:

$$Im(T) = \{T(\mathbf{v}) \mid \mathbf{v} \in \mathrm{V}\}$$

Kjernen til den samme T, $Ker(T)$, er definert som mengden av alle vektorer $\mathbf{v} \in V$ slik at $T(\mathbf{v}) = \mathbf{0}$:

$$Ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

Rangen til den samme T, $rank(T)$, er definert som dimensjonen av $Im(T)$. Hvis $A \in M_{m\times n}(K)$, definer vi rangen til A, til å være rangen av den lineære transformasjonen $L_A : \mathrm{K}^n \to \mathrm{K}^m$.

Kapittel 3

Kodeteori

En kode $\mathcal{C}$ er en delmengde av et vektorrom V over en kropp K. Vektorene i $\mathcal{C}$ kalles for kodeord. Vi vil se på koder der kroppen K er endelig. Dersom $K = \mathbb{Z}_2$ blir koden kalt en binær kode. En kode blir sakt å være av endelig lengde dersom $\mathcal{C} \subseteq K^n$ for en n, $n < \infty$. Kodeord blir representert med vektorer i form av n-tupler. For å nyansere forskjellen mellom ulike kodeord introduseres distanse. Innenfor matematikk eksisterer det flere former for distanse d, for to vektorer $\mathbf{u}$ og $\mathbf{v}$ av lengde n i et vektorrom V.

Euklidsk distanse: $d(\mathbf{u}, \mathbf{v}) = \sqrt{\sum_{i=1}^{n}(u_i - v_i)^2}$, for det reele vektorrommet $\mathbb{R}^n, n < \infty$.

Taxicab distanse: $d(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} \mid u_i - v_i \mid$, for det reele vektorrommet $\mathbb{R}^n, n < \infty$.

Hamming distanse: $d(\mathbf{u}, \mathbf{v}) =$ antall i hvor $u_i \neq v_i$, for vektorrommet $K^n, n < \infty$.

Innenfor kodeteori er Hamming distansen den relevante distansen når en sammenligner kodeord. Minsteavstanden mellom to ulike kodeord i $\mathcal{C}$ er gitt med:

$$\delta(\mathcal{C}) = \min(d(\mathbf{u}, \mathbf{v}) \mid \mathbf{u}, \mathbf{v} \in \mathcal{C}, \mathbf{u} \neq \mathbf{v})$$

Nærliggende til distanse er vekten w, som er antall ikkenull koefisienter i et kodeord $\mathbf{u}$, i K^n. Sammenhengen mellom vekt og distanse er;

$$d(\mathbf{u}, \mathbf{v}) = w(\mathbf{u} - \mathbf{v})$$

.

3.1 Generelle koder

Som nevnt i kapittel 1 finnes det koder av variabel lengde, slik som morsekoden, men disse blir problematisk å behandle matematisk, ser derfor på koder hvor kodeordene er like lange. La K_q være en kropp med q elementer. En kode $\mathcal{C}$ med kodeord av lengde n er en delmengde av K_q^n.

En strekkode $\mathcal{C}$ består av svarte og hvite streker (mellomrom). Tallene 0-9 blir kodet til fire streker (2 av hver) av ulik tykkelse, men den totale lengden er 7. Koden kan ekvivalent skrives med 0 for hvite og 1 for svarte streker, hvor 11 er representert med en dobbelt så tykk svart strek som 1, og 00 er representert med en dobbelt så tykk hvit strek som 0. Har da at $\mathcal{C} \subseteq \mathrm{K}_2^7$ og hvor 0-9, avhengig av struktur, kan kodes på 3 ulike måter som henholdsvis;

	0	1	2	3	4	5	6	7	8	*og*	9
*rad*1	0001101	0011001	0010011	0111101	0100011	0110001	0101111	0111011	0110111		0001011
*rad*2	0100111	0110011	0011011	0100001	0011101	0111001	0000101	0010001	0001001		0010111
*rad*3	1110010	1100110	1101100	1000010	1011100	1001110	1010000	1000100	1001000		1110100

Elementene i rad 3 er lik de i rad 1 bare invertert ($0 \to 1$ og $1 \to 0$), mens de i rad 2 er lik de i rad 3 bare baklengs. Det første sifferet i strekkoden bestemmer hvilken rad de seks neste sifrene skal kodes i:

førte siffer:	0	1	2	3	4	5	6	7	8	*og*	9
radnr:	111111	112122	112212	112221	121122	122112	122211	121212	121221	*og*	122121

mens de siste seks sifrene kodes ifølge rad 3.

Minsteavstanden innad i en rad er 2, mens mellom rader er den 1.

Boken *Fundamentals of Error-Correcting Codes* har strekkode 9780521131704, dette kodes:

$$\overset{9}{1}\overset{7}{2}\overset{8}{2}\overset{0}{1}\underbrace{0111101}_{rad1}\underbrace{0001001}_{rad2}\underbrace{0100111}_{rad2}\underbrace{0110001}_{rad1}\underbrace{0110011}_{rad2}\cdots\underbrace{1110010}_{rad3}\underbrace{1011100}_{rad3}$$

3.1.1 Lineære koder

La K_q være en endelig kropp med q elementer. En lineær kode av lengde n over K er da et undervektorrom av K^n. Hvis $\mathcal{C}$ er en lineær kode av lengde n og $\mathcal{C}$ som K-vektorrom har dimensjon k, sier vi at $\mathcal{C}$ er en $[n,k]$ lineær kode over K_q. Dersom en vet minsteavstanden sier vi at $\mathcal{C}$ er en $[n,k,\delta]$ lineær kode.

Ser på strekkoden som er innført i forrige avsnitt. Hvis en summerer to av kodeordene får en;

$$0001101 + 0011001 = 0010100$$

som ikke er et kodeord. Koden er dermed ikke lineær.

For å konstruere kodeordene til den $[n,k]$ lineære koden $\mathcal{C}$ tar en basisen $\{u_1,\ldots,u_k\}$,$u_i \in \mathrm{K}_q^n$ for vektorrommet $\mathcal{C}$ og ordner dem i en tabell som følger:

$$G = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k1} & u_{k2} & \cdots & u_{kn} \end{pmatrix}$$

$$G : K_q^k \to K_q^n$$

Bildene av en vektor i K_q^k ved å bruke G, $Im(\cdot, G) = \mathcal{C}$. Kaller da G en generatormatrise for $\mathcal{C}$, med k linært uavhengige rader, så $rank(G) = k$. Denne G er ikke entydig bestemt. Beskjeder $\mathbf{b} = (b_1,\ldots,b_k)$, $b_i \in K_q^k$ danner sammen med G kodeord $\mathbf{c} = (c_1,\ldots,c_n)$, $c_i \in K_q^n$ ved at

$$\mathbf{c} = \mathbf{b}G.$$

Kodeordene kan også kontrolleres med en paritetsjekkmatrise H, som er en $n \times (n-k)$ matrise. Denne H må tilfredsstille:

$$\mathcal{C} = \{\mathbf{c} \in K_q^n \mid \mathbf{c}H = \mathbf{0}\}$$

Dette er det samme som at $GH = \mathbf{0}$, og at $rank(H) = n - k$. H er heller ikke entydig bestemt.

Personnummeret er en [11,9] lineær kode over $\mathbb{Z}_{11}$. Har da at for fødselsinformasjon $\mathbf{b}$=$(d_1, d_2, m_1, m_2, å_1, å_2, n_1, n_2, n_3)$ blir kodet til $\mathbf{c}$=$(d_1, d_2, m_1, m_2, å_1, å_2, n_1, n_2, n_3, k_1, k_2)$, men for å finne k_1 og k_2 har vi:

$G = \begin{pmatrix} 1&0&0&0&0&0&0&0&0&8&1 \\ 0&1&0&0&0&0&0&0&0&4&10 \\ 0&0&1&0&0&0&0&0&0&5&9 \\ 0&0&0&1&0&0&0&0&0&10&0 \\ 0&0&0&0&1&0&0&0&0&3&9 \\ 0&0&0&0&0&1&0&0&0&2&1 \\ 0&0&0&0&0&0&1&0&0&7&3 \\ 0&0&0&0&0&0&0&1&0&6&6 \\ 0&0&0&0&0&0&0&0&1&9&1 \end{pmatrix}$ og for å kontrollere personnummeret $H = \begin{pmatrix} 3&5 \\ 7&4 \\ 6&3 \\ 1&2 \\ 8&7 \\ 9&6 \\ 4&5 \\ 5&4 \\ 2&3 \\ 1&2 \\ 0&1 \end{pmatrix}$

med rank(G)=9 og rank(H)=2

Den andre kolonnen i paritettsjekkmatrisen H for personnummeret, brukes også til å kontrollere hvorvidt et bankkontonummer er gyldig eller ikke.

3.1.2 Systematiske lineære koder

Systematiske koder er en delmengde av alle $[n, k]$ lineære koder, hvor den første delen av *Mottatt ord* fra figur 1.1, er den samme som *beskjed*. Det vil si at de k første koordinatene i kodeordet $\mathbf{c}$ er de samme som de i beskjeden $\mathbf{b}$ ($c_i = b_i, \forall i \leq k$), mens de resterende koordinatene kontrollerer koden.

Personnummeret er en systematisk kode da de 9 koordinatene i beskjeden er de samme som de 9 første koordinatene i kodeordet.

Generatormatrisen til en systematisk kode er på formen $G = \begin{bmatrix} I_k & A \end{bmatrix}$, hvor I_k er $k \times k$-identitetsmatrisen og A er en $k \times (n-k)$-matrise som koder resten av

koordinatene. Paritetssjekkmatrisen kan da skrives som $H = \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix}$ der A er den samme $k \times (n-k)$-matrisen som for G og I_{n-k} er $(n-k) \times (n-k)$-identitetsmatrisen.

3.1.3 Ekvivalente koder

Som nevnt er ikke generatormatrisen G og paritetssjekkmatrisen H unikt bestemt. Generatormatriser kalles ekvivalente dersom de kan kombineres med en lineærfunksjon T, representert med en inverterbar $k \times k$-matrise. Antall inverterbare $k \times k$ matriser og dermed ekvivalente generatormatriser over en kropp K_p, p primtall er:

$$\text{Antall iverterbare } k \times k\text{-matriser} = \prod_{i=0}^{k-1} (p^k - p^i)$$

Paritetssjekkmatriser kalles ekvivalente dersom de kan kombineres med en lineærfunksjon L, representert med en inverterbar $(n-k) \times (n-k)$-matrise. Har da:

$$Im(TG) = Im(G) \text{ og } Ker(HL) = Ker(H)$$

$$\begin{array}{ccccc} K_q^k & \xrightarrow{G} & K_q^n & \xrightarrow{H} & K_q^{n-k}. \\ \circlearrowleft & & & & \circlearrowleft \\ T & & & & L \end{array}$$

To $[n, k]$ lineære koder $\mathcal{C}$ og $\mathcal{C}'$ er permutasjonsekvivalent hvis det eksisterer en permutasjon av koordinater som sender $\mathcal{C}$ til $\mathcal{C}'$. Dette svarer til at det eksisterer en $n \times n$-matrise P med nøyaktig en ener i hver kolonne og nøyaktig en ener i hver rad. Da vil $\mathcal{C}P = \mathcal{C}'$

For hver lineær $[n, k]$ kode $\mathcal{C}$, finnes det en permutasjonsekvivalent kode $\mathcal{C}'$ som er systematisk.

Partitetssjekkmatrisen en har brukt for personnummeret er ikke systematisk, men om en multipliserer den med den inverterbare matrisen $L = \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$

$$\begin{pmatrix} 3 & 5 \\ 7 & 4 \\ 6 & 3 \\ 1 & 2 \\ 8 & 7 \\ 9 & 6 \\ 4 & 5 \\ 5 & 4 \\ 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{pmatrix} \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right) = \begin{pmatrix} 3 & 5-(2\cdot 3) \\ 7 & 4-(2\cdot 7) \\ 6 & 3-(2\cdot 6) \\ 1 & 2-(2\cdot 1) \\ 8 & 7-(2\cdot 8) \\ 9 & 6-(2\cdot 9) \\ 4 & 5-(2\cdot 4) \\ 5 & 4-(2\cdot 5) \\ 2 & 3-(2\cdot 2) \\ 1 & 2-(2\cdot 1) \\ 0 & 1-(2\cdot 0) \end{pmatrix} = \begin{pmatrix} 3 & 10 \\ 7 & 1 \\ 6 & 2 \\ 1 & 0 \\ 8 & 2 \\ 9 & 10 \\ 4 & 8 \\ 5 & 5 \\ 2 & 10 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

så blir den systematisk.

Paritetsjekkmatrisen kan fortsatt oppdage to feil, men fortsatt ikke dersom feilen skjer i 4 siffer og 10 siffer, med omvendt fortegn modulo 11.

3.2 Paritetssjekk og feiloppdagende koder

Grunnen til at kodeord er lengre enn den opprinnelige beskjeden er at funksjonen som koder beskjeden må være injektiv for at den skal være unikt dekodbar. Lengre sekvenser gir rom for paritetssiffer og de(n) ekstra koordinaten(e) gir informasjon om hele beskjeden. Dette gjør at dersom den kodede beskjeden blir utsatt for støy, er det fortsatt mulig å dekode beskjeden til sin opprinnelige form. Innfører vektoren $\mathbf{y}$ for det mottatte kodeordet. Dersom paritettsjekkmatrisen gir at $\mathbf{y}H \neq \mathbf{0}$ tilhører ikke $\mathbf{y}$ delmengden $\mathcal{C}$ og vi sier at det har skjedd minimum en feil[1]. Dersom en har at $\delta(\mathcal{C}) = 2$, vil enhver enkelt feil bli oppdaget. Dette kan gjøres ved å innføre at vekten av alle kodeord er et partallm dermed blir minsteavstanden også et partall.Ved å tilføre en ekstra bit til ordet av lengde n. En beskjed $b_0 b_1 \dots b_{n-1}$ (hvor $b_i \in \mathbb{Z}_2 : 0 \leq i \leq n-1$) blir da lagt til:

$$b_n = \sum_{i=0}^{n-1} b_i \pmod 2$$

Alternativt kan en en generere kodeordene med $n \times (n+1)$ generatormatrisen

[1]Merk, det kan oppstå feil slik at $\mathbf{y} \in \mathcal{C}$

$G = \begin{bmatrix} & 1 \\ I_n & \vdots \\ & 1 \end{bmatrix}$. Her blir paritetsjekkmatrisen H en $1 \times n$-matrise bestående utelukkende av enere. Dersom det har oppstått ett odde antall endringer i den mottatte beskjeden $\mathbf{y}$, så vil $\mathbf{y}H$=1.

Den amerikanske standardkoden for informasjonsutveksling (ASCII) tar sekvenser av 7-bits og knytter dem opp til ulike elementer. For eksempel har en at;

$$A = (1,0,0,0,0,0,1) \text{ og } D = (1,0,0,0,1,0,1).$$

Så distansen mellom de to elementene er $d(A,D) = 1$, da det bare er den femte koeffisiene som skiller de to.
For å gjøre vekten til et partall tilføres en åttende bit, 0 for A og 1 for D, slik at vi nå har

$$\bar{A} = (1,0,0,0,0,0,1,0) \text{ og } \bar{D} = (1,0,0,0,1,0,1,1)$$

$$w(\bar{A}) = 2 \text{ og } w(\bar{D}) = 4, \text{ så } d(\bar{A},\bar{D}) = 2$$

Nå er distansen mellom de to kodeordene økt til 2 da den femte og åttende koeffisienten er ulike. Siden den åttende biten er lagt til for at vekten skal bli ett partall, er også minsteavstanden 2.

Med paritetssjekk er det da mulig å oppdage dersom det skjer en feil. Avhengig av antall paritetsbits og hvordan de er organisert, er det mulig å oppdage flere enn en feil. Antall paritetsbit for en $[n,k]$ lineær kode er $n-k$. Personnummeret i seksjon 1.3 illustrerer dette. Et annet slik eksempel er International Standard Book Number (ISBN) som gir informasjon om hvilket språk boken har, hvilket forlag som har gitt ut boken og et eget nummer for boken i forlagets lager. Dette nummeret hadde orginalt formen $x_1x_2\ldots x_9x_{10}$ hvor x_{10} er kontrollsifferet og som tilfredsstiller:

$$x_{10} = 11 - \sum_{i=1}^{9}(11-i)x_i \pmod{11}.$$

Merk at dette betyr at en kan ha $x_{10} = 10$, dette markerer en med å gi det romerske

symbolet X for 10.

Boken, *Error Control Coding - From Theory to Practice* av Peter Sweeney, fra 2002, har ISBN: 0-470-84356-X. Dette betyr:
Språk: 0=Engelsk
Forlag: 470=John Wiley & Sons
Bokens nr: 84356
Sjekknr: X
Kan kontrollregne:

$$(10\cdot 0)+(9\cdot 4)+(8\cdot 7)+(7\cdot 0)+(6\cdot 8)+(5\cdot 4)+(4\cdot 3)+(3\cdot 5)+(2\cdot 6)+(1\cdot 10) = 209 \equiv 0 \pmod{11}$$

Sjekknr. stemmer for denne boken.

ISBN ble innført på 1960-tallet [se [ISB12]] da forlag og distributører ønsket å bruke datamaskiner til varetelling og bestilling av bøker. Da trengte de et nummer som ga informasjon og gjorde oppdaget dersom det hadde skjedd en feil. Fra og med 2007 fikk nye bøker et 13 sifret ISBN-nummer som tilfredsstiller

$$\sum_{i=1}^{13} a_i x_i = 0 \pmod{10}, \text{ hvor } a_i = \begin{cases} 1 & \text{når i er oddetall} \\ 3 & \text{når i er partall} \end{cases}$$

Det 13-sifrede ISBN nummeret koresponderer med strekkoden for for boken. De tre første koordinatene angir at det er en bok, mens de øvrige 10 følger samme mønster som for den 10-sifrede koden.

Boken *Fundamentals of Error-Correcting Codes* har
ISBN: 978-0-521-13170-4 og strekkode representert med: 9780521131704.

3.3 Feilkorrigerende kode

Det er altså mulig å finne ut om det har skjedd en feil eller ikke, men det er ikke alltid tid eller mulighet til å sende informasjonen på nytt. Dersom det er snakk

om informasjon som ligger lagret på harddisken til en datamaskin, trenger en å gjenskape den informasjonen som har gått tapt, dersom det er mulig.

Ettersom innføring av paritetssiffer gjorde det mulig å oppdage at det har skjedd feil, kan en få enda mer informasjon om hva som har gått tapt med innføring av flere paritetssifre. Den enkleste måten å innføre en kode som gjør det mulig å se hvor feilen har skjedd er å gjenta hvert siffer flere ganger. Den som mottar koden kan da sjekke hvilket gyldig kodeord som har er nærmest (har kortest distanse) til det mottatte kodeordet. Det enkleste eksemplet på dette er den trippelrepterende koden, hvor koeffisientene repeteres 3 ganger for å kunne oppdage hvor feilen har skjedd. Det vil si at en istedenfor å sende kodeordet (u_1, u_2, u_3, u_4) så sender en $(u_1, u_1, u_1, u_2, u_2, u_2, u_3, u_3, u_3, u_4, u_4, u_4)$. Fordelen her er at dersom det kun skjer 1 feil pr. triplet (u_i, u_i, u_i) er det mulig å rette opp feilen, og dersom det skjer 2 feil i en triplet ser en at noe er feil, men forsøk på å rette opp feilen vil være mislykket.

Dersom en ønsker å sende f.eks (1,0,1,0), blir dette kodet til (1,1,1,0,0,0,1,1,1,0,0,0) og sendt over en kanal hvor sannsynligheten for at en vilkårlig feil skjer f.eks er 10%.
Uavhengig av hvilket kodeord motageren får, kan en beregne sannsynligheten for at en korrekt klarer å dekode kodeordet. Ser først på en vilkårlig triplet. Denne klarer en å dekode korrekt dersom det har skjedd 0 eller 1 feil:

$$P(\text{0 eller 1 feil}) = P(\text{0 feil}) + P(\text{1 feil}) = (0,90)^3 + \binom{3}{1}(0,90)^2(0,10) = 0,972$$

En kan anta at de fire triplettene er uavhengig av hverandre, men for å få en korrekt dekoding, må alle ha maximalt 1 feil.

$$P(\text{max 1 feil i hver triplet}) = P(\text{0 eller 1 feil})^4 = 0,972^4 = 0,893$$

Sannsynligheten for at motageren mottar et kodeord han korrekt kan dekode til (1,0,1,0), er 89,3%.

Ulempen med den trippelrepeterende koden er at det er en belastning for ka-

nalen å sende så mye redudant informasjon. Dette tar opp mer lagringsplass og tar mer tid å sende. Innformasjonsraten for en kode beskriver hvor stor andel av koden som er innformasjon, rate=$\frac{k}{n}$. Derfor skal vi nå se på et par tilfeller av koder med en høyere rate som klarer å finne en feil og hvor stor feilen er.[2]

Det er mulige å danne en lineær $[p+1, p-1]$ kode $\mathcal{C}$ over en kropp K av primitiv størrelse p, det vil si K har elementene $\{0, 1, \ldots, p-1\}$, som kan rette opp en feil. Koden er generert av en G og har paritetsjekkmatrise H på formen;

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & p-1 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 1 & p-2 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 2 & p-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & p-3 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & p-2 & 1 \end{pmatrix} \text{ og } H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & p-2 \\ 1 & p-1 \end{pmatrix}$$

Har da at $\mathbf{y}H$=(x_1, x_2). Dersom (x_1, x_2)=(0,0) så har vi et gyldig kodeord i $\mathcal{C}$, dersom $(x_1, x_2)\neq$(0,0) har det oppstått minst en feil. Dersom det ikke er mer enn en feil vil (x_1, x_2) gi informasjon om hvor feilen har skjedd og hva feilen er. Hvis x_1=0 har feilen skjedd i første koordinat ved at $c_1 = y_1 - x_2$. Dersom $x_1 \neq 0$ er størrelsen på feilen x_1, og så finner en koordinaten der det har skjedd en feil ved at $i = 2 + x_2 \cdot x_1^{-1}$, hvor x_1 er slik at $x_1^{-1} \cdot x_1 = 1$. Dette gir at $c_i = y_i - x_1$, hvor y_i er det mottatt i-te koordinatet, mens c_i var sendt.

Dersom vi velger p=7 får vi:

$$\text{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 6 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 5 & 1 \end{pmatrix} \text{ og H} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix}$$

Hvis $\mathbf{b}$=(1 0 1 0 2 1) så får en $\mathbf{c}$=$\mathbf{b}$G=(1 0 1 0 2 1 2 1) (mod 7). Dersom en mottar $\mathbf{y}$=(1 0 1 3 2 1 2 1) kan en kontrollere dette med paritetsjekkmatrisen: $\mathbf{y}H$=(3 6) (mod 7).

$$i = 2 + 6 \cdot 3^{-1} = 2 + 6 \cdot 5 = 4, \text{ der } 3^{-1} = 5 \text{ siden } 3 \cdot 5 = 1 \pmod 7$$

Så vet da $c_4 = y_4 - 3 = 3 - 3 = 0$, den opprinnelige sendte beskjeden var da (1 0 1 0 2 1 2 1).

[2]For binære koder er det nok å finne posisjonen, siden feilen er 1.

En kode med minsteavstand δ kan oppdage dersom det har skjedd opptil $\delta - 1$ feil og kan rette kodeordet dersom det har skjedd opptil $\lfloor \frac{\delta-1}{2} \rfloor$ feil.

3.3.1 Projektiv kode

La K være en kropp med $\mu = p^n$ elementer. Ser på geometrien til ett prokjektivt plan bestående av punkter og linjer. Ett projektivt plan over K tilfredsstiller:

- Planet har $\frac{\mu^3-1}{\mu-1}$ punkter og $\frac{\mu^3-1}{\mu-1}$ linjer.

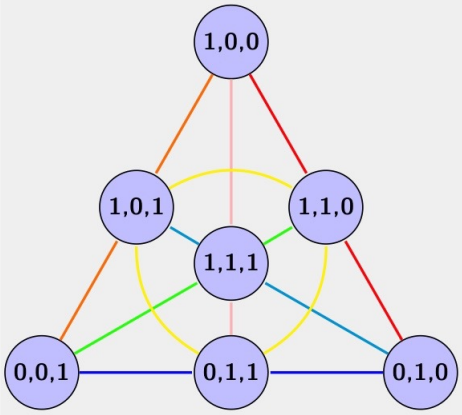
- Distinkte punkter fastsetter en unik linje, og distinkte linjer bestemmer ett unikt punkt

- Hver linje inneholder nøyaktig $\mu + 1$ punkter og hvert punkt har nøyaktig $\mu + 1$ linjer gjennom seg.

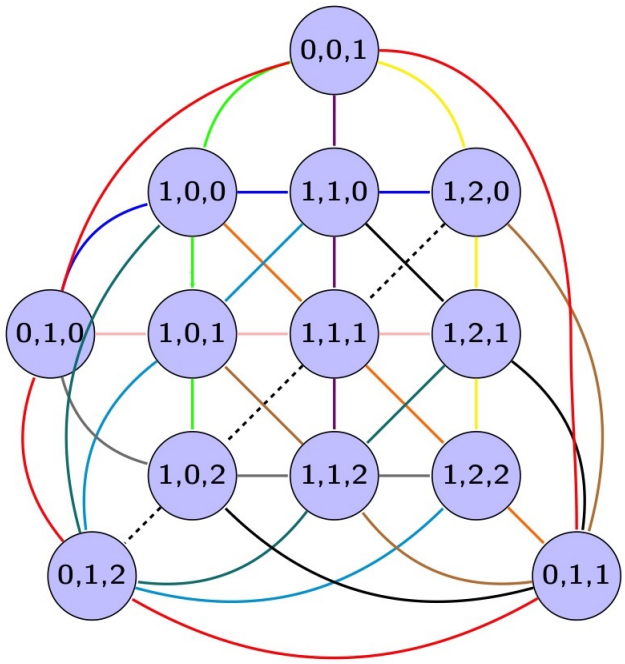
Det enkleste eksemplet på et projektivt plan er Fanoplanet, med orden 2. Dette gir 7 punkter og 7 linjer som systematiseres slik:

linjer \punkter	001	010	011	100	101	110	111
1	1	1	1	0	0	0	0
2	1	0	0	1	1	0	0
3	1	0	0	0	0	1	1
4	0	1	0	1	0	1	0
5	0	1	0	0	1	0	1
6	0	0	1	1	0	0	1
7	0	0	1	0	1	1	0

En mer grafisk fremstilling av dette planet er:



Ved å øke ordenen på det projektive planet til 3 får en 13 punkter og linjer. Dette gir både en mer utfordrende tabell og illustrasjon:



Som en ser blir disse illustrasjonene raskt kompliserte, så en setter strek der.

3.3.2 Hammingkoden

Richard Hamming (Se [Ham80]) kom med en forbedring av den trippelrepeterende koden. Ved utviklingen av Hamming (7,4)-koden kom han frem til en måte å kode en beskjed $(b_1, b_2, b_3, b_4); b_i \in \{0, 1\}$ som over til at hvert kodeord har minsteavstand 3, uten å måtte repetere koden 3 ganger. Dette ble gjort ved å innføre 3 paritetsbiter som tilfredsstiller: $b_5 = b_2 + b_3 + b_4$, $b_6 = b_1 + b_3 + b_4$ og $b_7 = b_1 + b_2 + b_4$, alle $\pmod 2$. Denne operasjonen kan enkelt illustreres med følgende matrisemultiplikasjon;

$$\begin{pmatrix} b_1 & b_2 & b_3 & b_4 \end{pmatrix} \begin{pmatrix} 1&0&0&0&0&1&1 \\ 0&1&0&0&1&0&1 \\ 0&0&1&0&1&1&0 \\ 0&0&0&1&1&1&1 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 \end{pmatrix}$$

For å kontrollere at det mottatte kodeordet er korrekt benytter en paritetssjekk matrisen som under og kontrollerer:

$$\begin{pmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 \end{pmatrix} \begin{pmatrix} 0&0&1 \\ 0&1&0 \\ 0&1&1 \\ 1&0&0 \\ 1&0&1 \\ 1&1&0 \\ 1&1&1 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$$

Dersom $\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$, har det ikke skjedd noen feil. Men dersom det har skjedd 1 feil vil koden varsle om posisjonen til denne feilen ved at å gjøre om det binære tallet $(x_1x_2x_3)$ til det desimaltall og få varsel om hvor feilen har skjedd. F.eks hvis en får $(x_1x_2x_3)$=(101) kan en gjøre om 101_2=5, altså er det b_5 som har en feil.

Merk at paritetssjekkmatrisen har de samme 7 radene som koordinatene til det projektive planet av orden 2. Dette er en egenskap hos de projektive kodene. Så har også en [13,10,3] kode over K_3, for planet av orden 3 med å ta koordinatene til de 13 punktene. Her systematisk ordnet:

$$H = \begin{pmatrix} 0&1&1 \\ 0&1&2 \\ 1&0&1 \\ 1&0&2 \\ 1&1&0 \\ 1&2&0 \\ 1&1&1 \\ 1&2&1 \\ 1&1&2 \\ 1&2&2 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{pmatrix} \text{ som gir generatormatrise } G = \begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0&2&2 \\ 0&1&0&0&0&0&0&0&0&0&0&2&1 \\ 0&0&1&0&0&0&0&0&0&0&0&2&0&2 \\ 0&0&0&1&0&0&0&0&0&0&0&2&0&1 \\ 0&0&0&0&1&0&0&0&0&0&0&2&2&0 \\ 0&0&0&0&0&1&0&0&0&0&0&2&1&0 \\ 0&0&0&0&0&0&1&0&0&0&0&2&2&2 \\ 0&0&0&0&0&0&0&1&0&0&0&2&1&2 \\ 0&0&0&0&0&0&0&0&1&0&0&2&2&1 \\ 0&0&0&0&0&0&0&0&0&1&2&1&1 \end{pmatrix}$$

3.3.3 Nye koder fra gamle

To like lange koder kan kombineres til å lage en tredje kode, som er dobbel så lang som de to andre. La $\mathcal{C}_i$ være en $[n, k_i, \delta_i]$ kode for i$\in \{1, 2\}$ over kroppen K_q. Da vil konstruksjonen $(\mathbf{u}|\mathbf{u}+\mathbf{v})$ danne en $[2n, k_1 + k_2, \min(2\delta_1, \delta_2)]$ kode

$$\mathcal{C} = \{(\mathbf{u}, \mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathcal{C}_1, \mathbf{v} \in \mathcal{C}_2\} \subseteq K_q^{2n}.$$

Hvis $\mathcal{C}_i$ har generatormatrise G_i og paritetssjekkmatrise H_i, vil generatormatrisen G og partitetsjekkmatrisen H til $\mathcal{C}$ være

$$G = \begin{bmatrix} G_1 & G_1 \\ 0 & G_2 \end{bmatrix} \text{ og } H = \begin{bmatrix} H_1 & -H_2 \\ 0 & H_2 \end{bmatrix}$$

La $\mathcal{C}_1$ være [7,4,3] hammingkoden og la $\mathcal{C}_2$ være [7,1,7] koden som repeterer beskjeden 7 ganger. Det gir

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

Som er en [14,5,6] kode.

Denne formen for kodekonstruksjon gir grunnlag for de binære Reed-Muller kodene. La m være et positivt heltall og la r være slik at $0 \leq r \leq m$. Konstruerer for hver m, $m+1$ binære koder av lengde 2^m, med notasjon $\mathcal{R}(r,m)$, $0 \leq r \leq m$.

$$\text{De } m+1 \text{ kodene er} \begin{cases} \mathcal{R}(0,m) & \text{er den } 2^m\text{-repeterende koden } \forall m \\ \quad \vdots & \\ \mathcal{R}(m,m) & \text{er hele rommet } K^{2^m} \forall m \end{cases}$$

Hvor de resterende $m-1$ verdiene av r, $0 < r < m$, er $\mathcal{R}(r,m)$ definert induktivt ved at:

$$\mathcal{R}(r,m) = \{\mathbf{u}, \mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathcal{R}(r,m-1), \mathbf{v} \in \mathcal{R}(r-1,m-1)\}$$

Minsteavstanden til $\mathcal{R}(r,m)$ er 2^{m-r}, så en oppdager at det har skjedd feil, dersom det ikke har skjedd mer enn $2^{m-r}-1$ feil, og en kan korrekt rette beskjeden dersom det ikke har skjedd mer enn $2^{m-r-1}-1$ feil. Reed-Muller kodene blir vanligvis ikke systematisk.

Generatormatrisen G(r,m) for $\mathcal{R}(r,m)$ er på formen;

$$G(r,m) = \begin{pmatrix} G(r,m-1) & G(r,m-1) \\ 0 & G(r-1,m-1) \end{pmatrix}$$

hvor $G(0,m)$ er $[11\cdots1]$ av lengde 2^m og $G(m,m)$ er I_{2^m}.
$G(r,m)$ er da en $(\sum_{i=0}^{r} \binom{m}{i}) \times 2^m$-matrise.

Generatormatrisen for $\mathcal{R}(2,5)$ er på formen:

$$G(2,5) = \begin{bmatrix} G(2,4) & G(2,4) \\ 0 & G(1,4) \end{bmatrix} = \begin{bmatrix} G(2,3) & G(2,3) & G(2,3) & G(2,3) \\ 0 & G(1,3) & 0 & G(1,3) \\ 0 & 0 & G(1,3) & G(1,3) \\ 0 & 0 & 0 & G(0,3) \end{bmatrix} \ldots etc.$$

Mer spesifikt danner dette en 16×32-matrise

$$G(2,5) = \begin{bmatrix}
1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0 \\
0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0 \\
0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0 \\
0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1 \\
0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0 \\
0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1 \\
0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1 \\
0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0 \\
0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1 \\
0&0&0&0&0&0&0&0&0&0&1&1&0&0&1&1&0&0&0&0&0&0&0&0&0&0&1&1&0&0&1&1 \\
0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1
\end{bmatrix}$$

Reed-Muller koden $\mathcal{R}(1,5)$ ble blant annet brukt av NASA (Se [**?**]) under Mariner 9 til Mars i 1972, for å ta bilder av planeten. Denne koden er generert av en 6×32-matrise. Dette ga mulighet til $2^6 = 64$ ulike beskjeder. Hver beskjed svarte til en pixel, hvor 0=helt hvit, 63=helt svart og ulike nyanser av grå mellom. Hvert kodeord har lengde 32, men siden minsteavstanden er $2^{5-1} = 16$, kunne NASA rette opp bildet dersom det ikke hadde skjedd mer enn $2^{5-1-1} - 1 = 7$ feil.

3.3.4 Sykliske koder

Matematisk bakgrunn

En kommutativ ring R er som en kropp K, bare at en ikke krever at $(K \setminus \{0\}, \cdot)$ er en abelsk gruppe. Ett ideal er en ikketom delmengde I av R hvor:

i) a,b$\in I$ medfører at a-b$\in I$

ii) a$\in I$ og r$\in R$ medfører at ar$\in I$ og ra$\in I$.

En polynomring R bstår av polynom med koeffisienter i K,

$$a_0 + a_1x + a_2x^2 + \cdots + a_mx^m, a_i \in K, m \in \mathbb{N}.$$

Teorem 3.3.1 (Divisjonsalgoritmen og den euklidske algoritmen)**.** *La $f(x)$ og $g(x)$ være polynom i $K[x]$ hvor $g(x)$ er ikkenull.*

i) Divisjonsalgoritmen sier at det eksisterer unike polynom $h(x), r(x) \in K[x]$ slik at

$$f(x) = g(x)h(x) + r(x), \text{ hvor } \deg r(x) < \deg g(x) \text{ eller } r(x) = 0.$$

I tillegg vil gcd($f(x), g(x)$)=gcd($g(x), r(x)$)

ii) Ved å utføre divisjonsalgoritmen gjentatte ganger til restpolynomet $r(x)$ er null, kan vi finne største felles divisor i $f(g)$ og $g(x)$

$$\begin{aligned}
i.e. f(x) &= g(x)h_1(x) + r_1(x), \text{ hvor } \deg r_1(x) < \deg g(x) \\
g(x) &= r_1(x)h_2(x) + r_2(x), \text{ hvor } \deg r_2(x) < \deg r_1(x) \\
r_1(x) &= r_2(x)h_3(x) + r_3(x), \text{ hvor } \deg r_3(x) < \deg r_2(x) \\
&\vdots \\
r_{n-3}(x) &= r_{n-2}(x)h_{n-1}(x) + r_{n-1}(x), \text{ hvor } \deg r_{n-1}(x) < \deg r_{n-2}(x) \\
r_{n-2}(x) &= r_{n-1}(x)h_n(x) + r_n(x), \text{ hvor } r_n(x) = 0.
\end{aligned}$$

Da er gcd($f(x), g(x)$)=$cr_{n-1}(x)$, hvor $c \neq 0 \in K$ er valgt slik at $cr_{n-1}(x)$ er monisk.

iii) Det eksisterer to unike polynomer $a(x)$ og $b(x) \in K[x]$, med $\deg(a(x)) < \deg(g(x))$ og $\deg(b(x)) < \deg(f(x)$ slik at:

$$a(x)f(x) + b(x)g(x) = gcd(f(x), g(x)).$$

Har at $\mathbb{Z}_{n_1 \cdot n_2} \simeq \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}$, når $\gcd(n_1, n_2) = 1$.

Ser på kroppen $\mathbb{Z}_{30}$:

$$\begin{aligned}\mathbb{Z}_{30} &\simeq \mathbb{Z}_6 \times \mathbb{Z}_5 \; textda \gcd(6,5) = 1 \\ &\simeq \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \text{ da } \gcd(2,3) = 1\end{aligned}$$

Uttrykket $x^{p^n} - x$ har ingen multippel rot i noen kroppsutvidelse av kroppen K_p.

$$\frac{d}{dx}(x^{p^n} - x) = p^n x^{p^n - 1} - 1 \equiv -1 \neq 0$$

For $f_i(x)$, irredusible og distinkte polynomer hvor $(x^{p^n} - x) = \prod^m f_i(x)$ har:

$$K_p[x]/(x^{p^n} - x) = K_p[x]/\prod^m f_i(x) \simeq K[x]/f_1(x) \times K[x]/f_2(x) \times \cdots \times K[x]/f_m(x)$$

Faktoriserer $x^{2^4} - x$:

$$x^{16} - x = (x)(x+1)(x^2 - x + 1)(x^4 + x + 1)(x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1)$$

$$\begin{aligned}K_2[x]/(x^{2^4} - x) &\simeq K_2[x]/(x) \times K_2[x]/(x+1) \times K_2[x]/(x^2 + x + 1) \\ &\times K_2[x]/(x^4 + x + 1) \times K_2[x]/(x^4 + x^3 + 1) \\ &\times K_2[x]/(x^4 + x^3 + x^2 + x + 1)\end{aligned}$$

Dette svarer til:

$$K_2[x]/(x^{2^4} - x) \simeq \mathbb{Z}_2 \times \mathbb{Z}_2 \times GF(4) \times GF(16) \times GF(16) \times GF(16)$$

Teorem 3.3.2 (Kinesiske restteorem)**.** *La $A_1, \ldots, A_n$ være ideal i en ring $\mathcal{R}$ med enhet slik at $A_i + A_j = \mathcal{R} \forall i \neq j$. Hvis $x_1, \ldots, x_n \in \mathcal{R}$, så eksisterer et element $x \equiv x_i \pmod{A_i} \forall i$*

Sykliske koder

Sykliske koder er en delmengde av de lineære kodene med følgende egenskap;
Hvis $(c_1, \ldots, c_{n-1}, c_n)$ er ett kodeord, så er $(c_n, c_1, \ldots, c_{n-1})$ også et gyldig kode-

ord. En kan flytte alle koordinatene ett hakk til høyre og den siste koordinaten først ved en syklisk permutasjon.
En syklisk kode er et ideal i restklasseringen $K_q[x]/(x^n-1)$ hvor gcd(n,q)=1.

Ett q-syklotomisk kosett av s modulo q^t-1 er definert som;

$$C_s = \{s, sq, \ldots, sq^{r-1}\} \pmod{q^t-1}, \text{ hvor } sq^r \equiv s \pmod{q^t-1}.$$

Ser på alle de 2-syklotomiske kosettene av s modulo $2^4-1=15$:

$$\begin{aligned}
C_0 &= \{0\} \\
C_1 &= \{1, 2, 2^2, 2^3\} = \{1, 2, 4, 8\} = C_2 = C_4 = C_8, \text{ da } 2^4 = 16 \equiv 1 \mod 15 \\
C_3 &= \{3, 3\cdot 2, 3\cdot 2^2, 3\cdot 2^3\} = \{3, 6, 12, 9\} = C_6 = C_9 = C_{12} \\
C_5 &= \{5, 10\cdot 2\} = \{5, 10\} = C_{10} \\
C_7 &= \{7, 7\cdot 2, 7\cdot 2^2, 7\cdot 2^3\} = \{7, 14, 13, 11\} = C_{11} = C_{13} = C_{14}
\end{aligned}$$

Det er ikke flere 2-syklotomiske kosett modulo 15. Det er vanlig prosedyre å kun nevne de distinkte kosettene med lavest s.

For å faktorisere (x^n-1) over kroppen K_q bruker en kroppsutvidelsen K_{q^t} av K_q som inneholder alle røttene til $xn-1$. Definerer ordenen av q modulo n, $ord_n(q)$, til å være det minste positive heltallet t slik at $q^t \equiv 1 \pmod{n}$. K_{q^t} har en primitiv n'te enhetsrot $\alpha \in K_{q^t}$ slik at:

$$\alpha^n = 1 \text{ og } \alpha^s \neq 1 \forall 0 < s < n$$

$$\begin{aligned}
x^n - 1 &= \prod_{i=0}^{n-1} (x - \alpha^i) \\
&= \prod_{s} \prod_{i \in C_s} (x - \alpha^i)
\end{aligned}$$

Vandermonde

La $\alpha_1, \ldots, \alpha_n$ være elementene i kroppen K. Da vil vi kalle $n \times n$-matrisen

$$V = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{bmatrix},$$

for Vandermonde matrisen.

Teorem 3.3.3 (Vandermondematrisen)**.**

$$detV = \prod_{1 \leq j < i \leq n} (\alpha_i - \alpha_j).$$

Spesielt så er V ikke-singulær hvis $\alpha_1, \ldots, \alpha_n$ er distinkte.

Bevis. : Beviser ved induksjon:

n=2

$$\det \begin{bmatrix} 1 & \alpha_1 \\ 1 & \alpha_2 \end{bmatrix} = \alpha_2 - \alpha_1$$

n-1 Antar ok.

n Lager ett polynom ved å ha variabelen x i den n-te raden.

$$\det \left[\begin{array}{cccc:c} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} & \alpha_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \alpha_{n-1} & \cdots & \alpha_{n-1}^{n-2} & \alpha_{n-1}^{n-1} \\ \hdashline 1 & x & \cdots & x^{n-2} & x^{n-1} \end{array}\right] = \prod_{1 \leq j < i \leq n-1} (\alpha_i - \alpha_j) x^{n-1} + g(x)$$

Her beskriver polynomet $g(x)$ ledd av lavere grad enn $n-1$. Koeffisienten til x^{n-1} stemmer pr. induksjonsantagelse. Vet at hvis $x = \alpha_i, i \in \{1, \ldots, n-1\}$ så er $\det V = 0$ siden to rader da blir lik.

$$\det V = \prod_{1 \leq j < i \leq n-1} (\alpha_i - \alpha_j) h(x)$$

Her er $h(x)$ et monisk polynom av grad $n-1$. Siden $x = \alpha_i, i \in \{1, \ldots, n-1\}$ gir $\det V = 0$, så må $h(\alpha_i) = 0$ for $i \in \{1, \ldots, n-1\}$. Siden $h(x)$ er monisk så er $h(x) = \prod_{i=1}^{n-1}(x - \alpha_i)$ det vil si $h(\alpha_n) = \prod_{i=1}^{n-1}(\alpha_n - \alpha_i)$.

$$\det V = \prod_{1 \leq j < i \leq n-1} (\alpha_i - \alpha_j) h(\alpha_n) = \prod_{1 \leq j < i \leq n} (\alpha_i - \alpha_j)$$

□

BCH-koder

La $g(x)$ være generatorpolynomet for en syklisk kode $\mathcal{C}$ av lengde n over kroppen K_q, der $\gcd(q, n)$=1 og la α være en n-terot av enhet i K_{q^t}, hvor $t = ord_n(q)$. Har at $g(x)$ også er et ideal i $K_q[x]/(x^n - 1)$. Har da at $g(x)$ må dele $x^n - 1$. Ved å velge n på formen $n = p^{t-1}$ har en at $g(x) \mid \prod f_i(x)$. Har da at $g(x)$ må være produktet av ett eller flere av de irredusible polynomene $f_i(x)$. Innfører den definerende mengden T for en syklisk kode $\mathcal{C}$ som:

$$T = \{i \in \{0, 1, \ldots, n-1\} \mid g(\alpha^i) = 0\}$$

Dette medfører at T er unionen av q-syklotomiske kosett, $T = \bigcup_s C_s$. Minsteavstanden til en syklisk kode kan beregnes ved hjelp av Bose-Ray-Chaudhuri-Hocquenghem-begrensningen, BCH-begrensningen. Sier at T inneholder mengden av s påfølgende elementer $\mathcal{S}$ hvis det er en mengde;

$$\mathcal{S} = \{b, b+1, b+2, \ldots, b+s-1\} \mod n \subseteq T.$$

Teorem 3.3.4 (BCH-begrensningen)**.** *La $\mathcal{C}$ være en syklisk kode av lengde n over kroppen K_q der $\gcd(q, n) = 1$ og la α være en primitiv n-te rot av enhet i K_{q^t}, hvor $t = ord_n(q)$ med definerende mengde T. Anta at T har en delmengde $\mathcal{S}$ av d påfølgende elementer. Da er minsteavstanden mellom elementene i $\mathcal{C}$ minst $d+1$.*

Ønsker å konstruere en syklisk kode $\mathcal{C}$ av lengde n over K_q som har både stor minsteavstand og høy dimensjon. BCH-begrensningen sier at minsteavstanden er større enn antall påfølgende elementer i den definerende mengden T. Dermed

kan en garantere stor minsteavstand ved å ha mange påfølgende elementer i den definerende mengden. Dimensjonen til $\mathcal{C}$ er $n-\mid T \mid$, så ønsker færrest mulig elementer i T for å få størst mulig kode.

En BCH-kode $\mathcal{C}$ over K_q av lengde n og konstruert distanse δ er en syklisk kode med definerende sett

$$T = C_b \cup C_{b+1} \cup \cdots \cup_{b+\delta-2}$$

hvor C_i er det q-syklotomiske kosett modulo n som inneholder i. Generatorpolynomet $g(x)$ til koden er på formen:

$$g(x) = \prod_{i \in T} (x - \alpha^i)$$

Generatorpolynomet har også formen $g(x) = \sum_{i=0}^{n-k} g_i x^i$ hvor $k = n - \deg(g(x))$. Denne formen for generatorpolynomet gir opphav til generatormatrisen G av størrelse $k \times n$:

$$G = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & \cdots & g_{n-k} & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & \cdots & g_{n-k-1} & g_{n-k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & \cdots & g_{n-k} \end{bmatrix}$$

Ved å benytte polynomet $h(x)$ som tilfredsstiller $h(x)g(x) = x^n - 1$, som kan skrives på formen $h(x) = \sum_{i=0}^{k} h_i x^i$ kan en konstruere paritetssjekkmatrisen H av størrelse $n \times (n-k)$:

$$H = \begin{bmatrix} h_k & 0 & \cdots & 0 \\ h_{k-1} & h_k & \ddots & \vdots \\ h_{k-2} & h_{k-1} & \ddots & 0 \\ \vdots & \vdots & \vdots & h_k \\ h_0 & h_1 & \vdots & \vdots \\ 0 & h_0 & \vdots & \vdots \\ \vdots & \ddots & \ddots & h_1 \\ 0 & \cdots & 0 & h_0 \end{bmatrix}$$

Ønsker å konstruere en BCH-kode av lengde 15 over K_2. $15 = 2^4 - 1$, så ser på kroppsutvidelsen $GF(16) = \mathbb{Z}_2/f(x)$, gitt ved det irredusible polynomet $f(x) = x^4 + x + 1$, hvor α er restklassen til x og $f(\alpha) = 0$, så $\alpha^4 = \alpha + 1$. Dersom koden skal rette 1 feil, må minsteavstanden være 3, hvilket betyr at T må ha to påfølgende elementer. Ettersom $C_1 = C_2$ (se side 33) kan vi bruke $T = C_1 \cup C_2 = C_1 = \{1, 2, 4, 8\}$. Får da generatorpolynomet:

$$\begin{aligned} g(x) &= (x-\alpha)(x-\alpha^2)(x-\alpha^4)(x-\alpha^8) \\ &= (x-\alpha)(x-\alpha^2)(x-(\alpha+1))(x-(\alpha+1)^2) \\ &= x^4 + x + 1 \end{aligned}$$

Dette gir en [15,11,3] syklisk kode med generatormatrise:

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

Har fra side 32 at

$$x^{15} - 1 = (x+1)(x^2 - x + 1)\underbrace{(x^4 + x + 1)}_{g(x)}(x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1)$$

Dermed vil paritetssjekkmatrisen ta utgangspunkt i polynomet $h(x)$:

$$\begin{aligned} h(x) &= (x+1)(x^2 - x + 1)(x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1) \\ &= x^{11} + x^8 + x^7 + x^5 + x^3 + x^2 + x + 1 \end{aligned}$$

Dersom den sykliske koden av lengde 15 over K_2 skal rette 2 feil, må minsteavstanden være 5, hvilket betyr at T må ha fire påfølgende elementer. Ettersom $C_1 = C_2 = C_4$ kan vi bruke $T = C_1 \cup C_3 = \{1, 2, 4, 8\} \cup \{3, 6, 9, 12\}$. Får da generatorpolynomet:

$$\begin{aligned} g(x) &= (x-\alpha)(x-\alpha^2)(x-\alpha^4)(x-\alpha^8)(x-\alpha^3)(x-\alpha^6)(x-\alpha^9)(x-\alpha^{12}) \\ &= (x^4+x+1)(x-\alpha^3)(x-\alpha^2(\alpha+1))(x-\alpha(\alpha+1)^2)(x-(\alpha+1)^3) \\ &= (x^4+x+1)(x^4+x^3+x^2+x+1) \\ &= x^8+x^7+x^6+x^4+1 \end{aligned}$$

Dette gir en [15,7,5] syklisk kode med generatormatrise:

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

Bibliografi

[Bay98] John Baylis. Error-Correcting Codes. Chapman and Hall Mathematics, 1998.

[Bur11] David M. Burton. Elementary Number Theory. McGraw Hill Education, 2011.

[Fra03] John B. Fraleigh. A First Course in Abstract Algebra. Pearson Education, 2003.

[Ham50] Richard Wesley Hamming. Error detecting and error correcting codes. Bell System Technical Journal, 29:147–160, 1950.

[Ham80] Richard Wesley Hamming. Coding and Information Theory. Prentice Hall, 1980.

[ISB12] Isbn users' manual, January 2012.

[Mor09] International morse code, October 2009.

[PSS94] P.B.Bhattacharya, S.K.Jain, and S.R.Nagpaul. Basic Abstract Algebra. Cambridge university press, 1994.

[Sel64] Ernst Sejersted Selmer. Personnummerering i norge: Litt anvendt tallteori og psykologi. Nordisk matematisk tidsskrift, 12:36–44, 1964.

[Sha48] Claude Elwood Shannon. A mathemathical theory of communication. Bell System Technical Journal, 27:379–443, 1948.

[SHFS03] Arnold J. Insel Stephen H. Friedberg and Lawrence E. Spence. Linear Algebra. Prentice Hall, 2003.

[Sin00] Simon Singh. The Code Book. Fourth Estate, 2000.

[Swe02] Peter Sweeney. Error Control Coding, from theory to practice. John Wiley and sons Ltd., 2002.