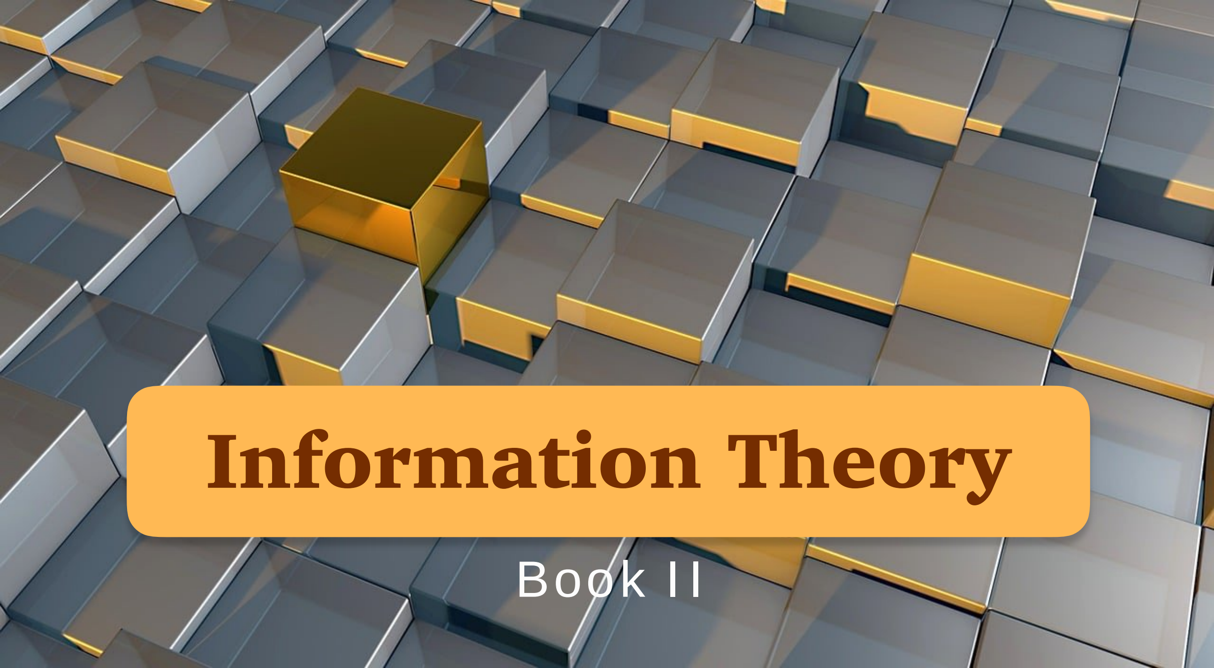
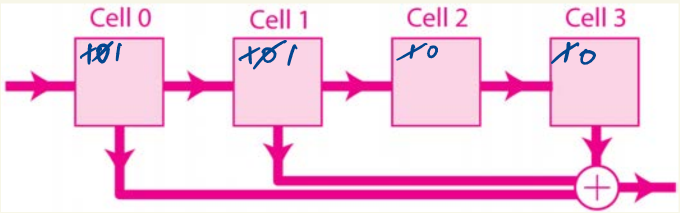


# Information Theory

Book II

Eg. an infinite stream of bits $a_0 a_1 a_2 a_3 a_4 \cdots$ $(a_i \in F)$ can be encoded eg. represent the plaintext bitstream as a $a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \in \mathbb{F}_2[[x]]$

$F[[x]]$ = ring of (formal) power series in $x$ with coefficients in $F$.



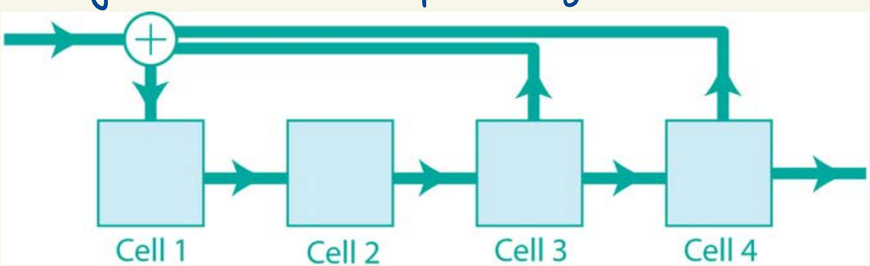

10110...

Eg. consider an input bitstream ~~1~~~~1~~~~0~~~~0~~~~1~~~~1~~0 1 1 1 1 0 0 1 0 ... which is encoded by the shift register above to obtain the output bitstream 1011 00101 ...

Compare: this is equivalent to multiplication by $1+x+x^3$:

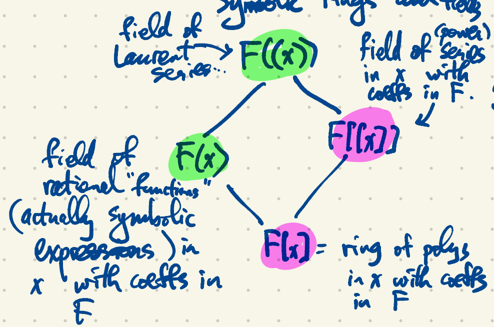
$$(1+x+x^3)(1+x+x^4+x^5+x^7+x^8+x^9+x^{10}+x^{13}+\cdots) = 1 + x^2 + x^3 + x^6 + x^8 + \cdots$$

Decoding of this data is accomplished using backward shift registers eg.




which performs division by $1+x+x^3$ in $\mathbb{F}_2((x))$

Symbolic rings and fields




polynomials vs. polynomial functions

eg. $\mathbb{F}_3 = \{0,1,2\} = \mathbb{Z}/3\mathbb{Z}$

eg. $f(x) = 2+x+x^3 \in \mathbb{F}_3[x]$ is a polynomial of degree 3.

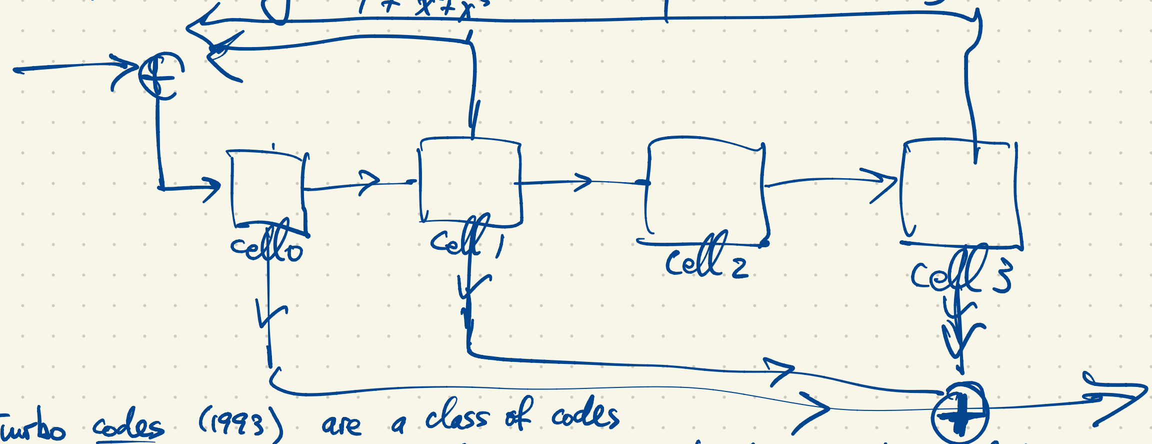
$g(x) = 2+2x \in \mathbb{F}_3[x]$ is a polynomial of degree 1.

| $a$ | $f(a)$ | $g(a)$ |
|---|---|---|
| 0 | 2 | 2 |
| 1 | 1 | 1 |
| 2 | 0 | 0 |

$f(x), g(x)$ are distinct poly's but they represent the same function $\mathbb{F}_3 \to \mathbb{F}_3$.

eg. $f(x) = \dfrac{1+x+x^3}{x+x^2} \in \mathbb{F}_2(x)$

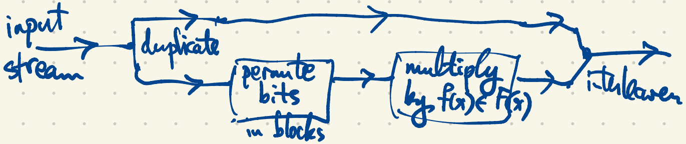
Multiplication by any rational function can be implemented using a single shift register e.g.
multiplication by $\frac{1+x+x^3}{1+x^2+x^3}$ is implemented using the shift register




Turbo codes (1993) are a class of codes used for encoding streams of data using combinations of gates including

- multiplication by a rational function in $F(x)$
- splitters & interleavers
- permutations
- puncturing

eg.




$F(x) \subset F((x))$ eg. for $F = \mathbb{F}_2 = \{0,1\}$

First method

$$f(x) = \frac{1+x^2+x^5}{x+x^2+x^4} = \frac{1+x^2+x^5}{x(1+x+x^3)} = \frac{1}{x}\left[\frac{1+x^2+x^5}{1+x+x^3}\right] = \frac{1}{x}\left[1+x+x^3+x^5+\cdots\right] = \frac{1}{x}+1+x^2+x^4+\cdots$$

$$\frac{1+x^2+x^5}{1+x+x^3} = 1 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + \cdots$$

$a_1=1$, $a_2=0$, $a_3=1$, $a_4=0$, $a_5=1$

$$1+x^2+x^5 = (1+x+x^3)(1+x+x^3+x^5+\cdots)$$

$(a+b)^2 = a^2+b^2$

$(a+b)^4 = a^4+b^4$

Second method

Geometric series $\frac{1}{1-u} = 1+u+u^2+u^3+u^4+\cdots$

$$\frac{1+x^2+x^5}{1+(x+x^3)} = (1+x^2+x^5)\left(1+(x+x^3)+(x+x^3)^2+(x+x^3)^3+(x+x^3)^4+(x+x^3)^5+\cdots\right)$$

$$= (1+x^2+x^5)\left(1+(x+x^3)+(x^2+x^6)+(x^3+x^5+\cdots)+(x^4+\cdots)+(x^5+\cdots)+\cdots\right)$$

$(x^3+3x^5+3x^7+x^9)$

$$= (1+x^2+x^5)(1+x+x^2+x^4+\cdots)$$

$$= 1+x+x^3+x^5+\cdots$$

$$f(x) = \frac{1}{x}(1+x+x^3+x^5+\cdots) = \frac{1}{x}+1+x^2+x^4+\cdots$$

$F = \mathbb{F}_2 = \{0,1\}$ for the time being

The irreducible (monic) polynomials in $F[x]$:

| degree | irred. polys |
|---|---|
| 1 | $x$, $x+1$ |
| 2 | $x^2+x+1$ |
| 3 | $x^3+x+1$, $x^3+x^2+1$ |
| 4 | $x^4+x+1$, $x^4+x^3+1$, $x^4+x^3+x^2+x+1$ |
| .... | .... |

primitive

not primitive

$x^2$, $x^2+1$, $x^2+x$, $x^2+x+1$ all poly's of degree 2.

$x^2 = x\cdot x$, $x^2+1 = (x+1)(x+1)$, $x^2+x = x(x+1)$

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

See MacWilliams & Sloane, The Theory of Error-Correcting Codes for more extensive lists of irreducible polynomials.

What are all the cyclic (linear) binary codes of length 7? There are exactly 8 of them. (why?)

- subspace of $F^7$, $F = \mathbb{F}_2 = \{0,1\}$
- invariant under cyclic shift $(a_0, a_1, a_2, a_3, a_4, a_5, a_6) \longmapsto (a_6, a_0, a_1, \dots, a_5)$ $a_i \in F$

A linear code $\mathcal{C} \subseteq F^n$ is cyclic iff its dual code $\mathcal{C}^\perp \subseteq F^n$ is also cyclic.

$\dim \mathcal{C} + \dim \mathcal{C}^\perp = n.$

eg. $\{(0000000)\}$

$\{0000000, 1111111\}$

$F^7 \leftarrow g(x)=1$, $h(x)=x^7-1$

$\{\text{words in } F^7 \text{ of even weight}\} = \langle 1100000, 1010000, 1001000, 1000100, 1000010, 1000001\rangle$

$$\begin{array}{r} 1101000 \\ 0101000 \\ \hline 1011100 \end{array}$$

Hamming $[7,4,3]_2$ code $\mathcal{H} = \langle 1101000, 0110100, \dots, 1010001\rangle$ (all cyclic shifts of 1101000 span this code)

$\dim \mathcal{H} = 4$, $|\mathcal{H}| = 2^4 = 16$:

1 codeword of weight 0
7 ... 3
7 ... 4
1 ... 7

Its dual $\mathcal{H}^\perp$, $\dim \mathcal{H}^\perp = 3$ is a $[7,3,4]_2$ - code.

$\mathcal{H}^\perp$ has 1 codeword of weight 0
7 ... 4

$\mathcal{H}^\perp = \mathcal{H} \cap \langle 1111111\rangle^\perp$

$\mathcal{H}' = \langle 1011000, 0101100, \dots, 0110001\rangle$ also $[7,4,3]_2$

$\mathcal{H}'^\perp$ also $[7,3,4]_2$.

$x^7 - 1 \in F[x]$ ← $n$ = length

$x^7 - 1 = \underset{\text{i.e. } x+1}{(x-1)}(x^6+x^5+x^4+x^3+x^2+1) = (x-1)\underbrace{(x^3+x+1)}_{(x-\alpha)(x-\alpha^2)(x-\alpha^4)}\underbrace{(x^3+x^2+1)}_{(x-\beta)(x-\beta^2)(x-\beta^4)}$

actually $x^7+1$ $\quad F = \mathbb{F}_2$

If $E = \mathbb{F}_q$, $x^q - x = \underset{x-I_0}{x}(x-1)(x-a_2)(x-a_3)\cdots(x-a_q)$

$a_0 = 0$, $a_1 = 1$, $a_2, a_3, \ldots, a_q$ are the field elements.

i.e. $x^{q-1} - 1$ has $q-1$ distinct roots which are the nonzero field elements.

If $\alpha \in \mathbb{F}_8$ is a root of $x^3+x+1$

$\mathbb{F}_8 = \mathbb{F}_2[\alpha] = \{a_0 + a_1\alpha + a_2\alpha^2 : a_0, a_1, a_2 \in \mathbb{F}_2\}$

$= \{0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+1, \alpha^2+\alpha, \alpha^2+\alpha+1\}$

Squaring is an automorphism of $\mathbb{F}_8$.

$(u+v)^2 = u^2+v^2$

$(uv)^2 = u^2v^2$

If $f(x) \in \mathbb{F}_p[x]$ is irreducible of degree $d$, then $\mathbb{F}_p[x]/(f(x)) \cong \mathbb{F}_{p^d} = \mathbb{F}_p[\beta]$ where $\beta$ is a root of $f(x)$.

$= \{a_0 + a_1\beta + a_2\beta^2 + \cdots + a_{d-1}\beta^{d-1} : a_i \in \mathbb{F}_p\}$

($\beta$ generates $\mathbb{F}_{p^d} \supset \mathbb{F}_p$ as an algebra)

If in fact $\mathbb{F}_{p^d} = \{0, 1, \beta, \beta^2, \beta^3, \ldots, \beta^{p^d-2}\}$ then we say $\beta$ is a <u>primitive</u> <u>element</u> and we say $f(x)$ is a primitive polynomial.

If $f(x) = x^4+x^3+x^2+x+1$ and $\beta \in \mathbb{F}_{16} = \mathbb{F}_{2^4}$ is a root of $f(x)$ then $\beta^5 = 1$ since $\beta$ is a root of $f(x)$

$\beta^5 - 1 = (\beta-1)\underbrace{(\beta^4+\beta^3+\beta^2+\beta+1)}_{0} = 0$

$0, 1, \beta, \beta^2, \beta^3, \beta^4, 1, \beta, \beta^2, \cdots$ doesn't give all of $\mathbb{F}_{16}$.

There are eight ways to factor $x^7-1 = g(x)h(x)$ in $\mathbb{F}_2[x]$.
In each case $g(x)$ is a generator poly. and $h(x)$ is a parity check poly. for a cyclic code of length 7 over $\mathbb{F}_2 = \{0,1\} = F$

Cyclic (linear) codes $\longleftrightarrow$ ideals in $F[x]/(x^n-1)$

$g(x)=1$, $h(x) = x^7-1$ gives $F^7$
$g(x) = x^7-1$, $h(x)=1$ gives $\{0000000\}$
$g(x) = x+1$, $h(x) = x^6+x^5+x^4+x^3+x^2+x+1$ gives all words of even weight i.e. $\langle 1100000, 1010000, \cdots, 1000001\rangle$
$g(x) = x^6+x^5+\cdots+1$, $h(x) = x+1$ gives $\langle 1111111\rangle = \{0000000, 1111111\}$
$g(x) = 1+x+x^3$, $h(x) = 1+x^2+x^3+x^4$ gives $\mathcal{H}$ $[7,4,3]_2$ code

**BCH bound** : a lower bound for performance of a cyclic code.
Consider a cyclic code of length $n$ over $\mathbb{F}_q$, i.e. an ideal in $\mathbb{F}_q[x]/(x^n-1)$ with gen. poly. $g(x)$, parity check poly. $h(x)$, $x^n-1 = g(x)h(x)$, $g(x)$ primitive, $\beta$ root of $g(x)$ in $\mathbb{F}_{q^r}$, $r = \deg g(x)$, and $\beta, \beta^2, \cdots, \beta^{s-1}$ are roots of $g(x)$, then the code has min. distance $\geq s$.

For Hamming $[7,4,3]_2$ code, $\beta$ root of $g(x) = 1+x+x^3 \in F[x]$, $\beta \in \mathbb{F}_8 = \mathbb{F}_2[\beta]$
Also $\beta^2$ - - - - - - - - by Freshman's Dream

$$1+\beta+\beta^3 = 0$$
$$(1+\beta+\beta^3)^2 = 1+\beta^2+\beta^6 = 0 = 1+\beta^2+(\beta^2)^3 \implies \mathcal{H} \text{ has min. dist.} \geq 3.$$

BCH : R.C. Bose
Dijen Ray-Chaudhuri
Hocquenghem

The Gilbert-Varshamov Bound (GV-bound): a lower bound for existence of good codes

$A_q(n,d) = \max.|\mathcal{C}|$ s.t. $\mathcal{C} \subseteq A^n$, $|A|=q$ with min. distance $\geq d$ i.e. $d(w,w') \geq d$ for all $w \neq w'$ in $\mathcal{C}$.

Ball of radius $r$ in $A^n$ centered at "0" $\in A^n$ has cardinality $|B_r(0)| = \sum_{k=0}^{r} \binom{n}{k}(q-1)^k$

$e = \lfloor \frac{d-1}{2} \rfloor$ = error-correcting capability.

$1 = |B_0| < |B_1| < |B_2| < \cdots < |B_n| = |A^n| = q^n$

Hamming bound: $A_q(n,d) \leq \dfrac{q^n}{|B_e|}$ : balls of radius $e$ centered at codewords $w \in \mathcal{C}$ are required to be disjoint

$$\bigsqcup_{w \in \mathcal{C}} B_e(w) \subseteq A^n. \Rightarrow |\mathcal{C}| \cdot |B_e(w)| \leq q^n$$

$$\Rightarrow |\mathcal{C}| \leq \frac{q^n}{|B_e(w)|}$$

In the other direction the GV-bound

$$A_q(n,d) \geq \frac{q^n}{|B_{d-1}(0)|}$$

so $\dfrac{q^n}{|B_{d-1}(0)|} \leq A_q(n,d) \leq \dfrac{q^n}{|B_e(0)|}$

We regard the GV bound as an existence proof only.

<u>Proof</u>: Let $\mathcal{C} \subseteq A^n$ be any $q$-ary code with $|\mathcal{C}| = A_q(n,d)$. We claim

$$\bigcup_{w \in \mathcal{C}} B_{d-1}(w) \supseteq A^n.$$

Codes satisfying this condition by greedy construction. But such codes are usually not practical because membership & decoding are not efficient.

If not, there exists $w' \in A^n$, $w' \notin \bigcup_{w \in \mathcal{C}} B_{d-1}(w)$ so $d(w',w) > d-1$ for all $w \in \mathcal{C}$.

But then $\mathcal{C} \cup \{w'\}$ has min. distance $\geq d$. This contradicts the maximality of $\mathcal{C}$ among all $q$-ary codes of length $n$ having min. distance $d$.

So $|\mathcal{C}|\,|B_{d-1}(0)| \geq |A^n| = q^n$. $\square$

Recommended viewing:
YouTube videos on coding & info. theory (including alg. geom. codes) by Mary Wootters

Asymptotic version of GV-bound due to Shannon:

Fix $0<\delta<1$. $|B_{\delta n}(0)| \approx |A^n|^{h_q(\delta)} = q^{n h_q(\delta)}$, $0 \le h_q(\delta) \le 1$.

$$\log_q |B_{\delta n}(0)| \approx n\, h_q(\delta)$$

This is a true asymptotic formula: for fixed $q$ and $\delta \in (0,1)$,

$$\frac{\log_q |B_{\delta n}(0)|}{n\, h_q(\delta)} \longrightarrow 1 \quad \text{as } n \to \infty.$$

$$\log_q |B_{\delta n}(0)| \sim n h_q(\delta).$$

More precisely,

$$n h_q(\delta) - o(1) \le \log_q |B_{\delta n}(0)| \le n h_q(\delta)$$

The qary entropy function

binary entropy function $\quad h_2(q) = -\delta \log_2 \delta - (1-\delta)\log_q(1-\delta) = \delta \log_2 \frac{1}{\delta} + (1-\delta)\log_2 \frac{1}{1-\delta}$

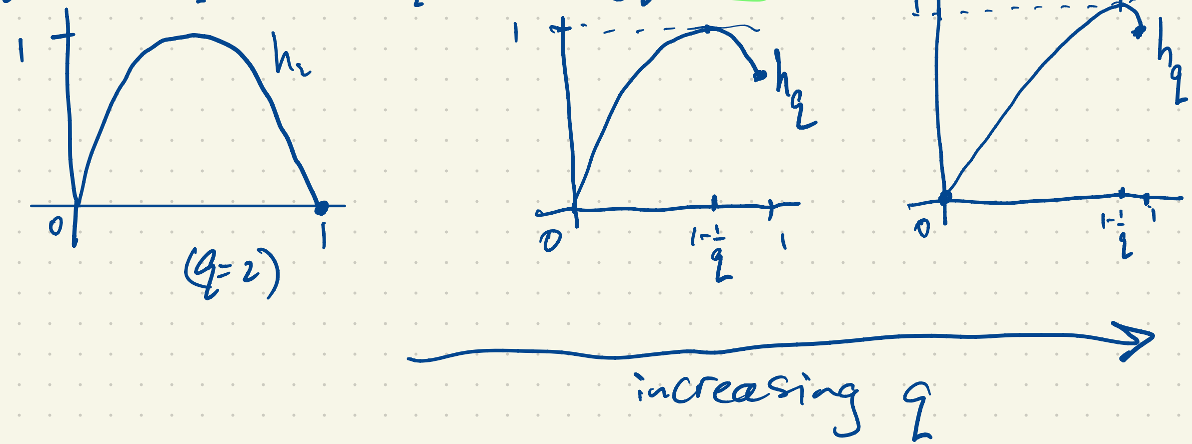
Eg. consider a random stream of information coming from letters in $A$, $|A|=q$, $A=\{x_1,\dots,x_q\}$

with letter $x_1$ having frequency $1-p$

$x_i$ .. .. $\frac{p}{q-1}$ $(2 \le i \le q)$ $\quad$ & $(1-p) + \frac{p}{q-1} + \frac{p}{q-1} + \dots + \frac{p}{q-1} = 1$.

$$H(\underset{\text{single char. from}}{\text{this stream}}) = \sum p \log \frac{1}{p} = -\sum p \log p = -(1-p)\log(1-p) - \cancel{(q-1)}\frac{p}{\cancel{q-1}} \cdot \log \frac{p}{q-1} = p\log(q-1) - p\log p - (1-p)\log(1-p)$$

$$h_q(\delta) = \delta \log_q(q-1) - \delta\log_q\delta - (1-\delta)\log_q(1-\delta)$$




$h_q(x) = x\log_q(q-1) + \frac{\log 2}{\log q} h_2(x)$   Let $x \to 1^-$.

$h_q(x) \to \log_q(q-1)$ as $x \to 1^-$.

For long codes ($n \gg 0$) over a fixed alphabet $|A| = q$, we consider the information rate $R = \dfrac{\log_q|C|}{n} = \dfrac{k}{n}$ in the case of an $[n,k]_q$-code

relative distance $\delta = \frac{d}{n}$

relative error-correcting capability $\frac{e}{n} \approx \frac{d}{2n}$

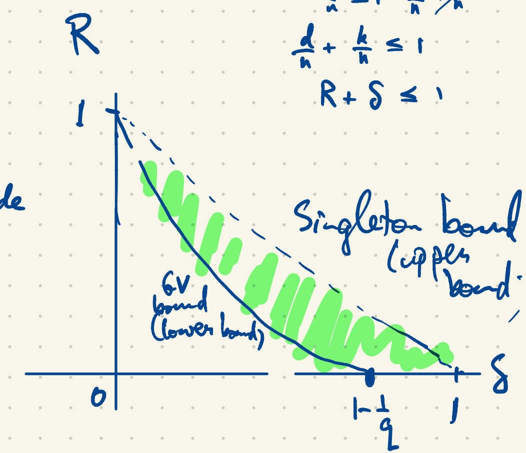
Singleton bound:

$d \le n-k+1$

$\frac{d}{n} \le 1 - \frac{k}{n} + \cancel{\frac{1}{n}}$

$\frac{d}{n} + \frac{k}{n} \le 1$

$R + \delta \le 1$




$R \ge 1 - h_q(\delta)$