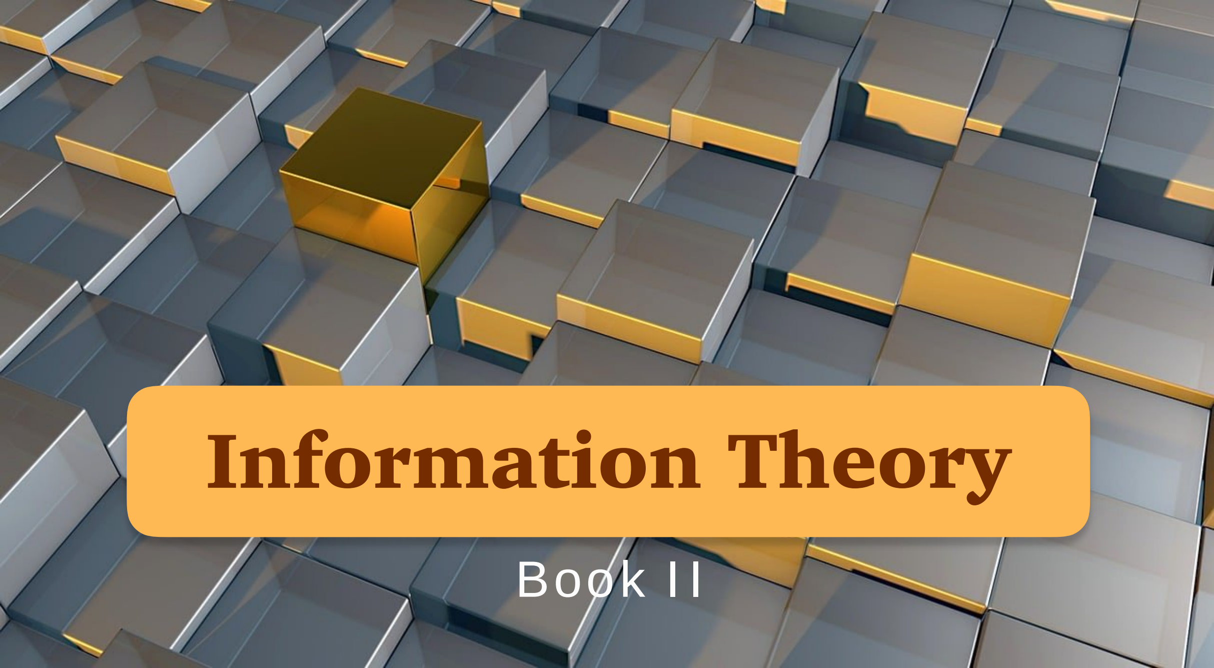
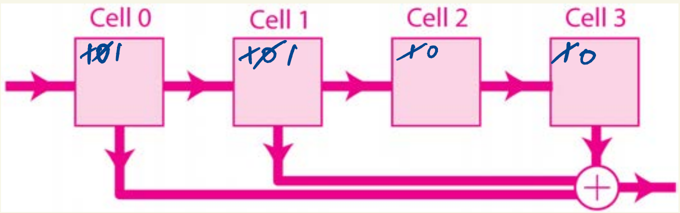


Information Theory

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.



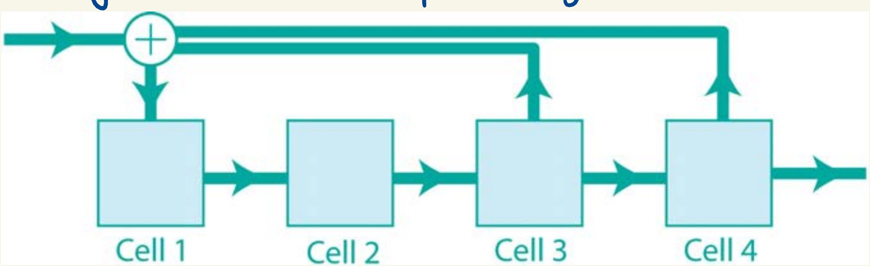
10110...

Eg. consider an input bitstream ~~110011~~011110010... which is encoded by the shift register above to obtain the output bitstream 101100101...

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

$$\begin{tikzcd}
 & F((x)) \arrow[dl, dash] \arrow[dr, dash] & \\
F(x) \arrow[dr, dash] & & F[[x]] \arrow[dl, dash] \\
 & F[x] &
\end{tikzcd}$$

field of Laurent series $\rightarrow F((x))$

field of (power) series in x with coeffs in F. $\rightarrow F[[x]]$

field of rational "functions" (actually symbolic expressions) in x with coeffs in F

$F[x]$ = ring of polys in x with coeffs in F

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.

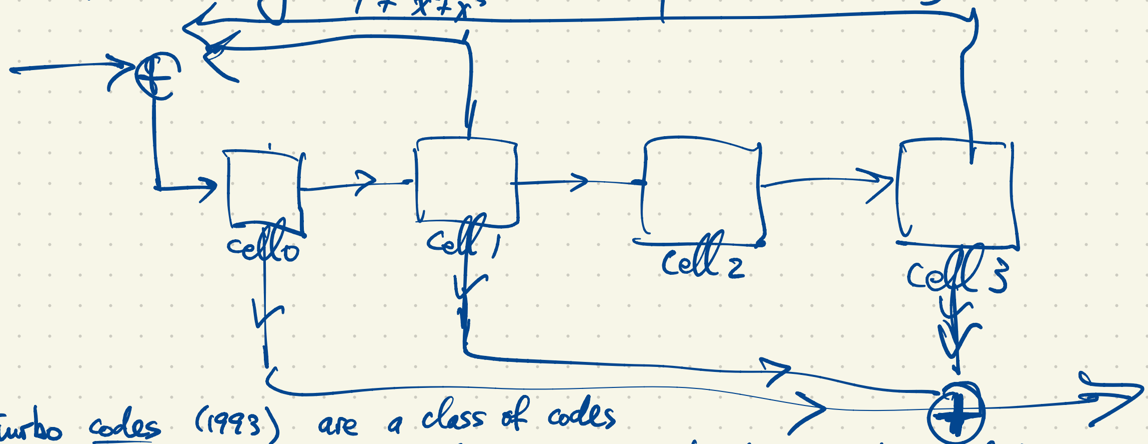
$g(x) = 2 + 2x \in \mathbb{F}_3[x]$ is a polynomial of degree 1.

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

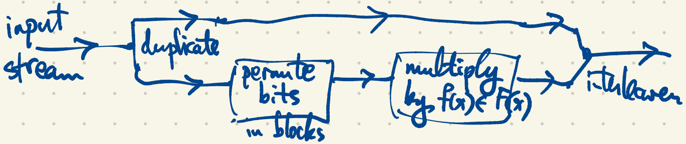
Multiplication by any rational function can be implemented using a single shift register e.g. multiplication by $\frac{1+x+x^3}{1+x^2+x^3}$ is implemented using the shift register



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

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

$$(a+b)^2 = a^2+b^2$$
$$(a+b)^4 = a^4+b^4$$

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

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

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

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	$x, \quad x+1$
2	x^2+x+1
3	$x^3+x+1, \quad x^3+x^2+1$
4	$x^4+x+1, \quad x^4+x^3+1, \quad x^4+x^3+x^2+x+1$
....	....

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 $(a_0,a_1,a_2,a_3,a_4,a_5,a_6) \longmapsto (a_6,a_0,a_1,\dots,a_5)$ $\quad a_i \in F$

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. $\{(0000000)\}$

$\{0000000, 1111111\}$

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

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

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

$\langle 1101000, 0110100, \dots, 1010001 \rangle$ (all cyclic shifts of 1101000 span this code)

Hamming $[7,4,3]_2$ code $\mathcal{H} = \langle 1101000, 0110100, \dots, 1010001 \rangle$

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

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 $E = \mathbb{F}_q$, $\quad x^q - x = \underset{x-0}{x}(x-1)(x-a_2)(x-a_3)\cdots(x-a_q)$

$a_0 = 0,\ a_1 = 1,\ a_2, a_3, \ldots, a_q$ 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

$$\mathbb{F}_8 = \mathbb{F}_2[\alpha] = \{a_0 + a_1\alpha + a_2\alpha^2 : a_0, a_1, a_2 \in \mathbb{F}_2\}$$
$$= \{0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+1, \alpha^2+\alpha, \alpha^2+\alpha+1\}$$

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

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

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

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 β is a root of $f(x)$.

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

(β 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 β 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 β is a root of $f(x)$

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

$0, 1, \beta, \beta^2, \beta^3, \beta^4, 1, \beta, \beta^2, \ldots$ 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, β 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, β root of $g(x) = 1+x+x^3 \in F[x]$, $\beta \in \mathbb{F}_8 = \mathbb{F}_2[\beta]$
Also β^2 - - - - - - - - by Freshman's Dream

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

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

We regard the GV bound as an existence proof only.

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

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

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, $|A| = q$, $A = \{x_1, \cdots, x_q\}$

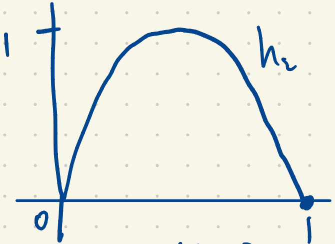
with letter x_1 having frequency $1-p$

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

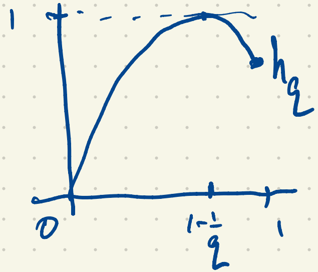
$H(\text{this stream}) = \sum p \log \frac{1}{p} = -\sum p \log p = -(1-p)\log(1-p) - \cancel{(q-1)}\frac{p}{\cancel{q-1}} \cdot \log \frac{p}{q-1} = p \log(q-1) - p \log p - (1-p)\log(1-p)$

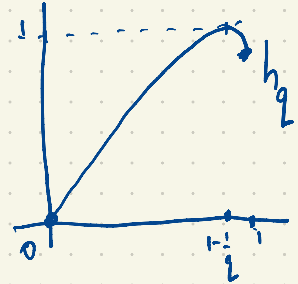
(single char. from)

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



(q = 2)





increasing q

Singleton bound:

$d \le n-k+1$

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

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

$R + \delta \le 1$

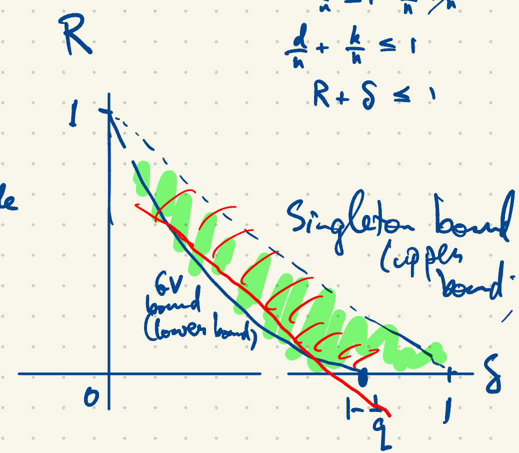
$h_q(x) = x\log_q(q-1) + \frac{\log 2}{\log q} h_2(x)$ Let $x \to 1^-$.

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

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

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

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



For $q \ge 49$ we have a new (1982) lower bound for asymptotically good explicit codes using algebraic geometry (Tsfasman, Vladut, Zink)

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

The 1982 theorem literally says: There exists a family X_i of algebraic curves over $\mathbb{F}_q$ $(i=1,2,3,\dots)$ such that X_i has n_i+1 (rational) points over $\mathbb{F}_q$, genus g_i with

$$\frac{g_i}{n_i} \longrightarrow \frac{1}{\sqrt{q}-1} \quad \text{as } i \to \infty.$$

The Reed-Solomon codes come from the simplest curve of all, the projective line $\mathbb{P}^1 F = F \cup \{\infty\}$ (F: field) of genus 0.

S^2 $\qquad$ $T^2 = S^1 \times S^1$

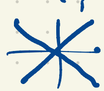
$g=0$ $\qquad$ $g=1$ $\qquad$ $g=2$ $\qquad$ $g=3$

On a curve X, $\Omega_X = \{\text{smooth global differential 1-forms}\}$ is a vector space of dimension $\dim \Omega_X = g$.
The number of $\mathbb{F}_q$-points on the curve (if it's defined over $\mathbb{F}_q$), N_q, satisfies $|N_q - (q+1)| \le 2g\sqrt{q}$

Hasse-Weil bound.

Eg. $\mathbb{P}^1 F$ has $N = q+1$ points, $g=0$

For a irreducible plane curve of degree d (defined by a poly. equation of degree d) has genus $g \le \binom{d-1}{2} = \frac{(d-1)(d-2)}{2}$
(equality for smooth curve; $g = \binom{d-1}{2} - \sum_{\text{singular points}} (\ \ )$.

$y^2 = x^2 \iff y = \pm x$  has $2q+1$ points

$y^2 - x^2 = (y+x)(y-x) = 0$

Irreducible conic:

$y = x^2$ $\qquad (t, t^2) \quad t \in F$ $\qquad$ genus $g=0$

plus one point at infinity

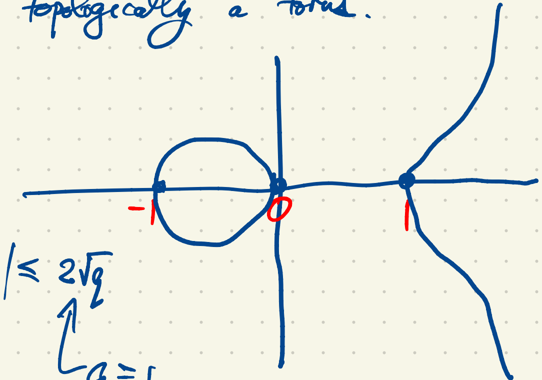
$q+1$ points

Smooth curve of degree $d=3$ has genus $g=\binom{3-1}{2}=1$ is topologically a torus.
(elliptic curve)

eg. $y^2=$ cubic in x with no repeated roots is an elliptic curve.

$y^2=x^3-x = x(x+1)(x-1)$

$g=1$ (torus)



H·W bound: over $\mathbb{F}_q$ the number of points satisfies $|N-(q+1)| \leq 2\sqrt{q}$ ← $g=1$

~~$q>3$~~ $q=$ prime $p\geq 5$

$N = q+1$ if $q=$ prime $p\equiv 3 \mod 4$

$q+1\pm\varepsilon$ if $q=$ prime $p\equiv 1 \mod 4$

$|\varepsilon| \leq 2\sqrt{q}$

Projective line $P^1F = F \cup \{\infty\} = X$

We consider rational functions $f(x)\in F(X)$ defined on a curve X (e.g. $X=P^1F$)

Eg. $F=\mathbb{F}_7=\{0,1,2,3,4,5,6\}$ A B C D E F G ∞ $= X$ (0 1 2 3 4 5 6)

Formal integer-linear combinations of points A, B, C, D, E, F, G, ∞ on X are called <u>divisors</u> as a book-keeping device for keeping track of zeroes and poles of functions on X.

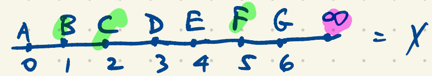
eg. $f(x)=(x-1)(x-2)(x-5)=x^3+3x^2-x-3 = x^3+3x^2+6x+4$ has simple zeroes at B, C, F and a triple pole at ∞

Near ∞, $z=\frac{1}{x}$; $f(x)=f(\frac{1}{z})=\frac{1}{z^3}+\frac{3}{z^2}+\frac{6}{z}+4 = \frac{1+3z+6z^2+4z^3}{z^3}$ so f has a triple pole at $z=0$ (i.e. at $x=\infty$).

The divisor of $f(x)$ is $B+C+F-3\infty =: \mathrm{Div}(f)$ (Sometimes abbreviated (f)).

More complicated: $f(x) = \dfrac{(x-1)^2(x+3)^4(x+5)}{(x+2)(x+1)^3}$

$\mathrm{Div}(f) = 2B + 4E + C - F - 3G - 3\infty$



$z = \frac{1}{x}$

$$f(x) = f(\tfrac{1}{z}) = \frac{(\frac{1}{z}-1)^2(\frac{1}{z}+3)^4(\frac{1}{z}+5)}{(\frac{1}{z}+2)(\frac{1}{z}+1)^3} \cdot \frac{z^7}{z^7} = \frac{(1-z)^2(1+3z)^4(1+5z)}{(1+2z)(1+z)^3 \cdot z^3}$$

triple pole at $z=0$ (i.e. at $x=\infty$)

The degree of $D = \sum_i m_i P_i$ is $\deg D = \sum_i m_i$, $(m_i \in \mathbb{Z})$

For any $f(x) \in F(X)$, $\deg(\mathrm{Div}\, f) = 0$. (equally many poles as zeroes)

Given a divisor $D = \sum_i m_i P_i - \sum_j n_j Q_j$ $(m_i, n_j \geq 1)$,

we consider the vector space $\mathcal{L}(D) = \{ f(x) \in F(X) :$ f has a zero of multiplicity at least m_i at P_i,
f has a pole of order at most n_j at Q_j,
and possibly other zeroes but no other poles$\}$.

In the case of $P^1F = F \cup \{\infty\}$ consider $D = k\infty$, $\deg D = k$.

$\mathcal{L}(k\infty) = \{ f(x) \in F(x) :$ f has a pole of order at most k at ∞; no other poles$\}$

(there can be as many zeroes as you like)

$\dim \mathcal{L}(k\infty) = k+1$. $\qquad \mathcal{L}(D) = \{ f : \mathrm{Div} f + D \geq 0 \}$

$\mathcal{L}(-k\infty) = \{$polynomials in x of degree at most $k\} = \{ a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k : a_i \in F \}$.

has basis $\{1, x, x^2, \cdots, x^k\}$.

The Riemann-Roch theorem gives a relation for determining $\ell(D) = \dim \mathcal{L}(D)$.

$$\underbrace{\ell(D)}_{\text{non-negative}} - \underbrace{\ell(K-D)}_{\text{integers}} = \deg D - g + 1$$

where K is a "canonical divisor"

$\ell(D) \geq \deg D - g + 1$ is Riemann's bound

The genus of a smooth curve X is the dimension $g = \dim \Omega_X$ where Ω_X is the vector space of (globally) smooth differential 1-forms on X.

eg. X = projective line P^1F over F, $P^1F = F \cup \{\infty\}$ $\qquad y = \frac{1}{x}$, $x = \frac{1}{y}$ $\qquad \frac{d(\frac{1}{y})}{dy} = -\frac{1}{y^2}$

A 1-form has the form (shape) $\omega = f(x)\,dx = f(\frac{1}{y})\,d(\frac{1}{y}) = -\frac{f(\frac{1}{y})\,dy}{y^2}$ $\qquad d(\frac{1}{y}) = -\frac{dy}{y^2}$

On P^1F there is no (nonzero) global 1-form

If $f(x)$ is a poly of degree k in x then it has a pole of order k at ∞ (and k zeroes in F).

So $\omega = f(x)\,dx$ has a pole of order $k+2$ at ∞.

$\frac{1}{x^2}dx$ has no pole at ∞ but it has a double pole at the origin.

$= -y^2\frac{dy}{y^2} = -dy$

Div $\omega = \Sigma(\text{zeroes of } \omega) - \Sigma(\text{poles of } \omega)$ the divisor of ω

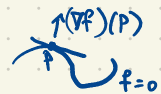
For ω a 1-form on P^1F, $\deg(\omega) = -2$ (2 more poles than divisors, counting multiplicity).

$\Omega_{P^1F} = \{0\}$, $g = \dim \Omega_{P^1F} = 0$.

For an elliptic curve e.g. the curve $X: y^2 = x^3 - x$ has $\Omega_X = \text{span}\{\omega\}$ (ω is a global smooth 1-form)

$g = \dim \Omega_X = 1$.

First, why is X a smooth cubic curve?



$y^2 = x^3$ is singular at $(0,0)$

$f(x,y) = y^2 - x^3$

$\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) = (-3x^2, 2y)$

$X: f = y^2 - x^3 + x = 0$

$\nabla f = (1 - 3x^2, 2y)$

$\begin{cases} 1 - 3x^2 = 0 \\ 2y = 0 \\ y^2 - x^3 + x = 0 \end{cases}$ has no solutions.

$y^2 = x^3 - x$ points $(x, y, 1)$ of a cubic curve with $z \neq 0$.

$y^2 z = x^3 - xz^2$ points (x, y, z) in homogeneous coords

If $z = 0$: $x = 0$, $y = 1$.

Near this point, $y \neq 0$, divide by y to get $z = x^3$.

$$(x, y, z) \mapsto (\underbrace{\tfrac{x}{y}}_{x'}, 1, \underbrace{\tfrac{z}{y}}_{z'})$$

$$\frac{y^2 z}{y^3} = \frac{x^3 - xz^2}{y^3}$$

$$z' = (x')^3 - x'(z')^2$$

$$y = x^3 - xy^2$$

$(0,0)$ is a smooth point.

$$f = y^2 - x^3 + x = 0$$

$$y^2 = x^3 - x$$

$$dy^2 = d(x^3 - x)$$

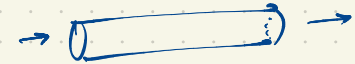
$$2y\,dy = (3x^2 - 1)\,dx$$

$$\omega = \frac{dy}{3x^2 - 1} = \frac{dx}{2y}$$

$$\frac{z\,dy}{3x^2 - z^2} = \frac{dx}{2y}$$

This equation is preserved under scalar multiples $(x, y, \cancel{z}) \mapsto (\lambda x, \lambda y, \cancel{\lambda z})$ $(\lambda \neq 0)$

Shannon's Theorem for noisy channels

Imagine a pipe in which we can send 1 liter of water per second. 

Now for the same pipe imagine that a certain amount of sludge/silt/gravel is carried along at a rate of ε liters per second. This means that only $1-\varepsilon$ liters of water per second can be transmitted by this same channel/pipe.

Amazingly, the same simplistic reasoning applies to send information reliably.

Suppose we transmit information using strings of symbols from an alphabet A, $|A| = q$.

If there were no noise, we could reliably send 1 character (letter in A) per unit time.

If instead error is introduced to the channels having entropy rate $h_q(\varepsilon)$ (characters per unit time)

then the rate at which useful information can be reliably transmitted ($0 \le h_q(\varepsilon) \le 1$)

in this channel is asymptotically $1 - h_q(\varepsilon)$ characters per unit time.

$$|B_d| = \sum_{k=0}^{d} \binom{n}{k}(q-1)^k, \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \text{ as } n \to \infty$$

Stirling's formula

Back to Shannon's first theorem: optimal compression of information for noiseless channel

Source of information is a random variable $X = \begin{cases} x_1 & \text{with prob. } p_1 \\ x_2 & \cdots\ p_2 \\ \vdots \\ x_k & \cdots\ p_k \end{cases}$ $0 \le p_i \le 1$, $\sum p_i = 1$

We ask for the optimal compression of info. from this source using strings over alphabet A, $|A| = q$

A code for this source is a map $X \to A^*$

$$\mathcal{C}: \begin{array}{l} x_1 \longmapsto w_1 \in A^{l_1} \\ x_2 \longmapsto w_2 \in A^{l_2} \\ \vdots \end{array}$$

i.e. w_i is a word in A^* of length l_i.

The expected length of $\mathcal{C}(X)$ is $\sum_{i=1}^{k} p_i l_i$

<u>Theorem</u> $\sum p_i l_i \geq H_q(X) := \sum_i p_i \log_q \frac{1}{p_i} = -\sum_i p_i \log_q p_i$. Moreover, we can asymptotically achieve compression having $\sum_i p_i l_i$ as close as desired to $H_q(X)$.

$\mathcal{C}$ must be an injective map (the code is uniquely decodable). We will discuss the proof under the stronger assumption that $\mathcal{C}$ is prefix-free: none of the codewords $w_1, \dots, w_k$ is a prefix (initial substring) of any of the other codewords.

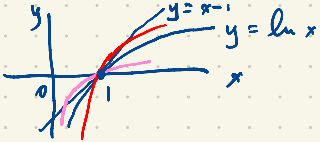
<u>Lemma</u> (Kraft's inequality) $\sum_{i=1}^{k} \frac{1}{q^{l_i}} \le 1$.

<u>Proof</u> Elements in $[0,1]$ (real interval) can be written in base q as infinite strings over A as

$0.a_1 a_2 a_3 a_4 \cdots$, $a_j \in A$

Each $w_i \in \mathcal{C}(X)$ determines a subinterval of $[0,1]$ given by all real numbers whose first l_i "digits" agree with w_i i.e. $r \in [0,1]$ s.t. $r \restriction l_i = w_i$. These real numbers form a subinterval of width $\frac{1}{q^{l_i}}$.

These subintervals are disjoint. $\square$

Proof 

$\ln x \le x-1$ for all $x>0$.

$\log_q x = \dfrac{\ln x}{\ln q}$ because $x = q^y$, $y = \log_q x$

$\ln x = y \ln q \Rightarrow y = \log_q x = \dfrac{\ln x}{\ln q}$

$$\ln \frac{1}{p_i q^{l_i}} \le \frac{1}{p_i q^{l_i}} - 1$$

$$p_i \ln \frac{1}{p_i q^{l_i}} \le \frac{1}{q^{l_i}} - p_i \qquad (1 \le i \le k)$$

$$\sum_{i=1}^{k} p_i \ln \frac{1}{p_i q^{l_i}} \le \underbrace{\sum_{i=1}^{k} \frac{1}{q^{l_i}}}_{\le 1} - \underbrace{\sum_{i=1}^{k} p_i}_{=1} \le 0$$

, divide both sides by $\ln q > 0$

$$\sum_{i=1}^{k} p_i \log_q \frac{1}{p_i q^{l_i}} \le 0$$

$$\sum_{i=1}^{k} p_i \left(\log_q \frac{1}{p_i} + \log_q \frac{1}{q^{l_i}} \right) \le 0$$

$\log_q \frac{1}{q^{l_i}} = -l_i$

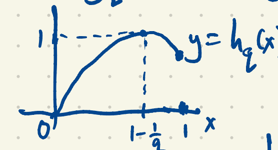
$$\underbrace{\sum_{i=1}^{k} p_i \log_q \frac{1}{p_i}}_{H(X)} \le \underbrace{\sum_{i=1}^{k} p_i l_i}_{E(\text{length of a codeword})} \qquad \square$$

Outline of Shannon's second coding theorem (source coding for noisy channels)

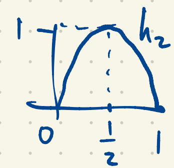
If $0 \le \varepsilon \le 1$ then $|B_{\varepsilon n}(v_0)| \approx q^{n h_q(\varepsilon)}$

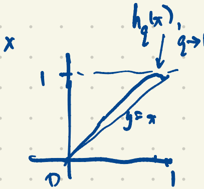
$$\overset{\shortparallel}{\{v \in A^n : d(v, v_0) \le \varepsilon n\}}$$

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



$\frac{1}{n} \log_q |B_{\varepsilon n}(v_0)| \sim h_q(\varepsilon)$ as $n \to \infty$ (q fixed)





Hamming Bound: for $\mathcal{C} \subseteq A^n$ e-error correcting

$$|\mathcal{C}||B_e| \le |A^n| = q^n.$$

$$\log_q |\mathcal{C}| + \log_q |B_e| \le n$$

$$R = \underbrace{\frac{1}{n} \log_q |\mathcal{C}|}_{\text{info. rate}} + \frac{1}{n} \log_q |B_e| \le 1 \implies R \le 1 - \frac{1}{n} \log_q |B_e|$$

As $n \to \infty$, q fixed

$$R \le 1 - h_q(\varepsilon)$$

relative error

$e = \varepsilon n$ ($0 < \varepsilon < 1$ fixed)

$d \approx 2e$

$d = 2e$ or $2e+1$

$e = \lfloor \frac{d-1}{2} \rfloor \approx \frac{d}{2}$

$d = \delta n$ relative distance

If $\mathcal{C}$ is linear, $\dim \mathcal{C} = k \le n$, info. rate $\frac{k}{n} = \frac{1}{n} \cdot \log_q(q^k)$