DATA TRANSMISSION (noisy communication channel)

channel noise $\longrightarrow$ redundancy encoding

m initial message $\rightarrow$ M bits

encoding $w(m) = \underline{x}$ codeword N bits

Redundant coding : $N > M \longrightarrow$ encoding rate $R = M/N$

$0 < R < 1$

<u>codebook</u> : set of all possible codewords (2^M)

transition probability

$\underline{x}$ sent through a channel $\longrightarrow Q(\underline{y}/\underline{x})$

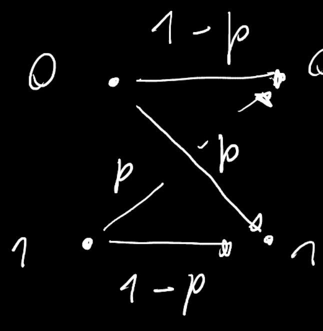
memoryless channels $Q(\underline{y}/\underline{x}) = \sum_{1}^{N}{}_i\, Q(y_i/x_i)$

$y_i \in \mathcal{Y}$

decoding $\rightarrow$ $m' = d(\underline{y})$ $\quad m' = m$?

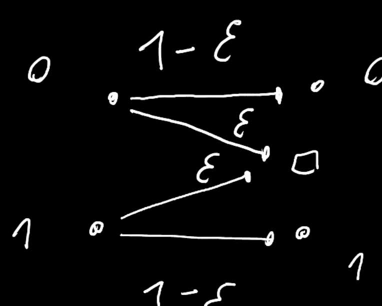
<u>Example of channels</u>

- Binary Symmetric Channel



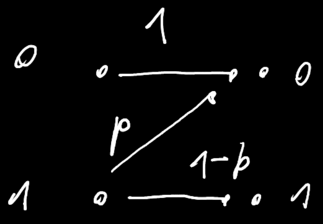
p = crossover (flip) probability

- Binary erasure channel



ε = bit erasure probability

- Z channel (asymmetric)



Channel Capacity C

$$C = \max_{p(x)} \underline{I_{X,Y}} = \max_{p(x)} \sum_{x,y} p(x,y) \cdot \log_2 \left[\frac{p(x,y)}{p(x)p(y)}\right]$$

$p(x,y) = p(x)\,Q(y/x)$; $p(y) = \sum_x p(x,y) = \sum_x p(x)\,Q(y/x)$

$I_{X,Y}$ = reduction in the uncertainty of X due to the knowledge of Y

the higher $I_{X,Y}$, C → the better the channel

- X, Y independent $\xrightarrow{\text{(pure noise)}}$ $I_{X,Y} = C = 0$
- perfect correlation (no noise) $Y = f(X)$ → $I_{X,Y} = H_X$

$$C = \max_{p(x)} \mathcal{H}(p) = 1 \quad (\text{uniform } p(x))$$

$$\boxed{0 \leq C \leq 1}$$

$m_1, m_2, \ldots$ blocks of bits

Measures for transmission fidelity

source = sequence of uncorrelated, unbiased bits

Shannon's source channel separation theorem

input stream parsed in blocks

- Block error probability

for a given input block m

prob. that m is not transmitted correctly

$$P_B(m) = \sum_{\underline{y}} Q(\underline{y} / \underline{x}(m)) \, \mathbb{I}(d(\underline{y}) \neq m)$$

2^M possible input m

- Maximal block error prob.

$$P_B^{max} = \max_m P_B(m) \longrightarrow \text{"worst case"}$$

- Average block error prob.

$$P_B^{av} = \frac{1}{2^M} \sum_m P_B(m) = P_B$$

$$\boxed{\text{Channel coding Theorem}} \quad \left(\begin{array}{c} \text{Shannon} \\ 1948 \end{array} \right)$$

1) given C; $\forall \boxed{R < C}$;

(direct) there exists a sequence C_N; R_N; $P_{B,N}$

such that $R_N \to R$; $C_N \to C$; $P_{B,N} \to 0$

as $N \to \infty$

2) given sequences $C_N \to C$; $R_N \to R$; $P_{B,N} \to 0$

(inverse) $(\text{as } N \to \infty)$

$$\Rightarrow \boxed{R < C}$$

- reliable communication is possible $(P_B = 0)$ $\iff$ $R < C$ $(R = M/N)$

- not trivial at all to have $P_B = 0$ with $R > 0$

(vanishing ~~error~~ block probability by adding a finite redundancy)

- proof (direct) $\to$ <u>RANDOM CODE ENSEMBLE</u>

<u>Qualitative argument</u>: Shannon entropy $H \to$ " 2^H values "

given $\underline{X}(m)$ – $\underline{Y}$ is a random variable with Shannon entropy $H_{\underline{Y}/\underline{X}} = N H_{Y/X}$

$\to 2^{N H_{Y/X}}$ possible channel outputs (given input)

Space of all possible outputs $\longrightarrow 2^{N H_Y}$

for <u>perfect decoding</u>: the number of codewords $< \dfrac{2^{N H_Y}}{2^{N H_{Y/X}}}$

$$2^M = 2^{RN}$$

$$RN < N(H_Y - H_{Y/X})$$

$$\to R < H_Y - H_{Y/X} = I_{X,Y} \leq C \to \boxed{R < C}$$

RANDOM ENERGY MODEL (p-spin model $p \to \infty$)

2^N configurations with energies E_i $1 \leq i \leq 2^N$

E_j are i.i.d random variables

Gaussian PDF $\quad \rho(E) = \frac{1}{\sqrt{\pi N}} \exp\left(-E^2/N\right)$

zero mean

$N/2$ variance (thermodynamic extensive potentials)

given disorder realization (instance / sample)

($\{E_1, \dots, E_{2^N}\}$ each sampled from $\rho(E)$

Boltzmann distrib. $\mu_\beta(j) = \frac{\exp(-\beta E_j)}{Z_N(\beta)}$

itself a random variable

$\mathbb{E}(\cdot)$ = expect. value over sample realization

$\langle \cdot \rangle$ = thermal average (over $\mu_\beta(j)$)

Self-averaging property for X (generic thermodynamic extensive variable)

($\lim_{N\to\infty} \mathbb{P}\left[\left|\frac{X_N}{N} - \mathbb{E}\left(\frac{X_N}{N}\right)\right| \geq \vartheta\right] = 0 \;\; \forall \vartheta > 0$

no sample-to-sample variation for $N \to \infty$

THERMODYNAMICS OF REM

microcanonical ensemble $\longrightarrow$ density of states as a function of

energy density $\longrightarrow$ $\varepsilon = E/N$

$$\rho(\varepsilon) = \sqrt{\frac{N}{\pi}}\, \exp(-N\varepsilon^2)$$

Density of states $\omega(\varepsilon) = \rho(\varepsilon)\, 2^N \simeq$

$(\omega(\varepsilon)\, d\varepsilon$ = # of energy levels between ε and $\varepsilon + d\varepsilon)$

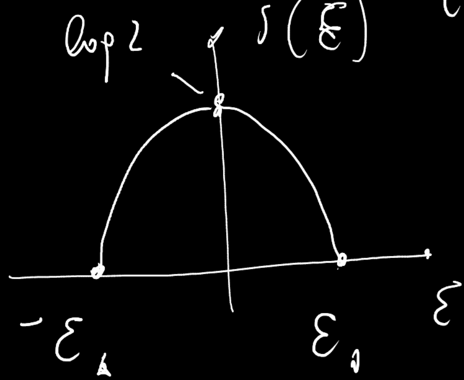
$$\underset{N \gg 1}{\simeq} \exp\big(N(\underbrace{\log 2 - \varepsilon^2})\big)$$

- if $|\varepsilon| < \varepsilon_* = \sqrt{\log 2}$ → exponentially large density of energy levels
- if $|\varepsilon| > \varepsilon_* = \sqrt{\log 2}$

→ density of energy levels exponentially small in N

Microcanonical entropy density $(N \to \infty)$

$$s(\varepsilon) = \begin{cases} \log 2 - \varepsilon^2 & \text{if } |\varepsilon| < \varepsilon_* \\ -\infty & \text{if } |\varepsilon| > \varepsilon_* \end{cases}$$



(no possible states as $N \to \infty$)

$$\boxed{\text{Canonical partition function}}$$

$$Z_N(\beta) \underset{N \gg 1}{\simeq} \int_{-\varepsilon_*}^{\varepsilon_*} \exp\left[N\left(s(\varepsilon) - \beta\varepsilon\right)\right] d\varepsilon$$

$$Z_N(\beta) \underset{N \gg 1}{\simeq} \exp\left[N\phi(\beta)\right]$$

SADDLE POINT

$$\phi(\beta) = \max_{-\varepsilon_* \leq \varepsilon \leq +\varepsilon_*} \left(S(\varepsilon) - \beta \varepsilon \right)$$