06/01/2025

Bases

$|v_1\rangle$ ← Ket notation.

$|v_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $|v_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$|v\rangle = a_1|v_1\rangle + a_2|v_2\rangle$

$= \begin{bmatrix} a_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ ← Linear combination (spans the vector space)

Linear independence

$a_1|v_1\rangle + a_2|v_2\rangle + \dots a_n|v_n| = 0$ } linearly dependent

$\exists a_i \neq 0$ & $a_1, a_2, \dots a_n$ are complex

set of linearly independent elements which span the entire vector space → has same no. of elements.
(called Basis)

10/01/2025

Inner product

$(|v\rangle, |w\rangle)$ or $\langle v|w\rangle$

$(|v\rangle, |v\rangle) \geq 0$ (equal at zero vec)

Dual of $|v\rangle$ is $\langle v|$

Spanning

In a binary field

coeff		
0/1	0 0 1	v_1
0/1	0 1 0	v_2
0/1	1 0 0	v_3
a_i	V_i	

$\sum a_i v_i$

if all a_i's are 0, then $0\,0\,0 \rightarrow$ (all zero case).

$0\,0\,1 \rightarrow v_1$

$0\,1\,0 \rightarrow v_2$

$1\,0\,0 \rightarrow v_3$

$0\,1\,1 \rightarrow 1\,1\,0$

$1\,0\,1 \rightarrow 1\,0\,1$

$1\,1\,0 \rightarrow 0\,1\,1$

$1\,1\,1 \rightarrow 1\,1\,1$

> These are all possible elements of the three dimensional space. All elements are covered through the linear combination of the 3 vectors.

> Subspace of a vector space is such that any operation on a vector in the subspace keeps the result within it.

> Two vector spaces are orthogonal if two vectors, one each taken from a space, give zero on an inner product.

Norm of $|v\rangle$

$\| |v\rangle \| = \sqrt{\langle v|v\rangle}$

if $\| |v\rangle \| = 1$ then vector is normalised

$\frac{|v\rangle}{\| |v\rangle \|}$

Orthonormal vectors :- Both orthogonal and norm = 1.

Outer product

$|w\rangle = \sum_i w_i |i\rangle$ $\qquad |v\rangle = \sum_j v_j |j\rangle$

$A : V \rightarrow W$

$A = I_W A I_V = \sum_{ij} |w_j\rangle\langle w_j|A|v_i\rangle\langle v_i|$

$= \sum \langle w_j|A|v_i\rangle |w_j\rangle\langle v_i|$

$A^\dagger \leftarrow$ adjoint conjugate & transpose

$A^\dagger = (A^*)^T$

$A^\dagger = A$ ← Hermitian matrix

$AA^T = A^T A = I$ ← Unitary operator

$(U|v\rangle, U|w\rangle) = \langle v|U^\dagger U|w\rangle = \langle v|I|w\rangle = \langle v|w\rangle$

(dagger)

Unitary operator conserves the Inner Product. and trace

Tensor product

$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$

$X \otimes Y$ works on 4×1

X & Y works on 2×1

Inner product [A, B]

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$[X, Y] = AB - BA = 2iZ$$

Since it is not zero, A & B do not commute.

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

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Postulates of quantum mechanics - state space

→ Associated to any physical system is a complex vector space with inner product (i.e. Hilbert space) known as state space of the system

→ ~~Set~~ vectors to describe it → state vector $|\psi\rangle$

Evolution of state space of system → unitary transformation

$$|\psi'\rangle = U|\psi\rangle$$

↑ unitary

← Discrete.

Unitary transformation exists for each evolution from state ψ to ψ'.

Continuous

schrödinger

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle$$

↑ Hamiltonian

Lagrange was the first to introduce state space concept. He started x and $\dot{x}$ and related it to Newtonian mechanics $F = ma$.

(Rate of change of momentum)

Eg: Burning of fuel → change in mass.

In classical mechanics, measurement is somewhat accurate. However quantum mechanics involves superposition of states. Any measurement results in a collapse. Hence, it is defined through a collection of measurement operators.

If we count over the measurement (for n number of tries), it results in a probability

$$P_m = \langle\psi| M_m^\dagger M_m |\psi\rangle$$

↑ exhaustive.

State of system collapses into a single state

$$\frac{M_m|\psi\rangle}{\langle\psi|M_m^\dagger M_m|\psi\rangle}$$

; which is heavily dependent on what is measured.

$$\sum M_m^\dagger M_m = I$$

; since all probabilities add to 1.

In Qubit measurement

a & b satisfy completeness

$M_0 = |0\rangle\langle 0|$

$M_1 = |1\rangle\langle 1|$

if $a = b = \frac{1}{\sqrt{2}}$, then $P_0 = P_1$ (probabilities)

If we can prepare a state in superposition of a and b, then we will get to the states with probabilities P_1 and P_2.

Distinguishing quantum states

→ Can only be done for orthogonal states

→ Non-orthogonal states have components that influence each other. In such a case, the probability of the state $|v_i\rangle$ which is prepared, will not be able to reach 1.

→ with a full probability of 1 (one).

The no. of times each state collapses into during measurement, is determined by amplitudes.

Book: Quantum computation and quantum information by Michael A. Nielsen and Isaac L. Chuang

Qubit

Dirac notation $\langle\psi|\psi\rangle$

⇒ where there are 2 very distinct, opposing states such as spin of electron / photon / magnetic resonance but cannot ascertain that state, then 0 and 1 are achieved.

17/01/2025

With probability α we get $|0\rangle$ states & with probability β we get $|1\rangle$ states.

upon measurement, it collapses to 1 state & we cannot retrieve α and β.

Qubit states α and β are unobservable

$|+\rangle$: $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ equiprobability of having $|0\rangle$ and $|1\rangle$ states

Binary Implementations

Transformation: Rotation of qubit by an angle,

Adjustment of $|0\rangle$ to $|+\rangle$ gives rise to mixed-state.

Changing of states is done through energy.

There are merits and demerits of creating qubits in

→ Coherence

It does not stay for long in a state. It is not stable (μs, ms).

The mixture of states is retained for a very short duration for exploiting quantum computing. (Decoherence at the end of the duration)

For a technology.

If decoherence takes place in μs.

If change of states is of order ms

Then no-of computations allowed = 1000.

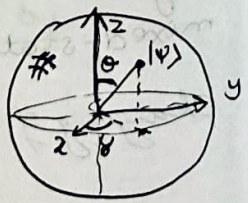
Photon particle based implementation is more suitable for Quantum Communication.

Classical & Quantum computation coexist to obtain a speedup.

SLIDE 27: Some binary implementations → mainly theory; has not been realised as such.

Geometric Intepretation of qubit

$$|\psi\rangle = e^{i\delta}\left(\cos\left(\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}|1\rangle\right)\right)$$



$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

If any state (other than $|0\rangle$ and $|1\rangle$) has a component over another state, the states cannot be distinguished through rotation.

$$|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{bmatrix}$$

$$\frac{\theta}{2} = \frac{\pi}{2} \Rightarrow \theta = \pi \,; \quad \phi = 0$$

for $|0\rangle$

$$\cos\frac{\theta}{2} = 1, \quad \theta = 0 \,; \quad \phi = \frac{\pi}{2}$$

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{bmatrix}$$

Quantum NOT gate

$\alpha^2 + \beta^2 = 1$

I/P: $\alpha|0\rangle + \beta|1\rangle$

O/P: $\alpha|1\rangle + \beta|0\rangle$

probability of getting zeroes becomes probability of getting ones and vice-versa.

Quantum mechanics only follows linear combination otherwise there would be a paradox like time travel or faster than light.

Information travel,

Matter colliding with anti-matter → Removed in linearity.

Results in entropy destruction (violates Second law of thermodynamics).

Q) What is Qubit?

Linear combination

2 states

About Non-linear: Paradox / Violations

QNOT

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$ ← Pauli gates

$$\alpha|0\rangle + \beta|1\rangle \leftarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$X\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$$ ← unitary transformation

$$|\psi'\rangle = \alpha'|0\rangle + \beta'|1\rangle$$

$$|\alpha|^2 + |\beta|^2 = 1 = |\alpha'|^2 + |\beta'|^2$$

$U^\dagger U = I$

$U^\dagger$ ← adjoint

U^t → Transpose then Complex conjugate

Pauli gate

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$|0\rangle$ → unchanged

$|1\rangle$ → flips to $-|1\rangle$

$$U_{or}\, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix}$$

Hadamard gate

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$ sq. root of NOT gate

$$|0\rangle \rightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|1\rangle \rightarrow \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

} Halfway between $|0\rangle$ and $|1\rangle$

$H^2 = I$; not the NOT gate

(★) Applying Hadamard gate twice leaves the state unchanged.

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H^2 = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

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

Hadamard

→ Rotate about Y by 90°

→ Rotate about X by 180° (reflection)

X : $\alpha|0\rangle + \beta|1\rangle$ to $\beta|0\rangle + \alpha|1\rangle$

Z : $\alpha|0\rangle + \beta|1\rangle$ to $\alpha|0\rangle - \beta|1\rangle$

H : $\alpha|0\rangle + \beta|1\rangle$ to Hadamard

Common gates (each work on single qubits)

Hadamard $H : \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Pauli X : $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Pauli Y : $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$

Pauli Z : $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Phase : S : $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$

$\pi/8$: (T) : $\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$

All these are unitary transformations

∴

Multiple qubit quantum gates

00, 01, 10, 11

Suppose we have an info system having 2 symbols

$\emptyset$, 1

$P_\emptyset$, P_1 ← probabilities of getting $\emptyset$ and 1 in a bit stream.

Selecting a random number and checking it against a threshold.

If we take 2 bits at a time.

00, 01, 10, 11

probabilities: 0.16, 0.24, 0.24, 0.36

let
$P_0 = 0.4$
$P_1 = 0.6$

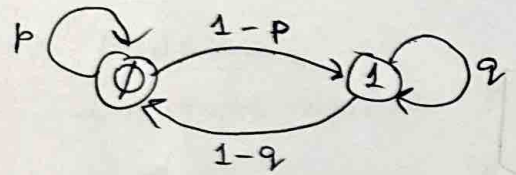
$P_0P_0 + P_0P_1 + P_1P_0 + P_1P_1 = 1$

$P_0^2 + P_0P_1 + P_1P_0 + P_1^2 = 1$

0.16 0.24 0.24 0.36

0.16
0.24
0.24
0.36

system



Dependency on previous state

← STATE TRANSITION MATRIX

p & q : probabilities (threshold) beyond which state change.

~~what is~~

If p & q are too small then observing over long term is beneficial

↓ how do we get to long term?

State transition matrix

$\begin{bmatrix} P_0(k) \\ P_1(k) \end{bmatrix}$ ← prob of 0 & 1 in kth state

k → step/iteration.

$$\begin{bmatrix} P_0(k+1) \\ P_1(k+1) \end{bmatrix} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \begin{bmatrix} P_0(k) \\ P_1(k) \end{bmatrix}$$

$$P_0(k+1) = p\,P_0(k) + (1-q)\,P_1(k)$$

$$P_1(k+1) = (1-p)\,P_0(k) + q\,P_1(k)$$

In the long term $P(k+1)$ and $P(k)$ will lose sense in their independent iterations, it becomes P_0 & P_1, S_0 :-

$$P_0 = p\,P_0 + (1-q)\,P_1$$

$$P_1 = (1-p)\,P_0 + q\,P_1$$

$$\Rightarrow (1-p)\,P_0 = (1-q)\,P_1$$

$\begin{cases} p' = 1-p \\ q' = 1-q \end{cases}$

$$\Rightarrow p'\,P_0 = q'\,P_1$$

also, $P_0 + P_1 = 1$

$$\Rightarrow \frac{P_0}{P_1} = \frac{q'}{p'}$$

$$\Rightarrow \frac{P_0}{P_1} + 1 = \frac{q'}{p'} + 1$$

$$\Rightarrow P_1 = \frac{\bar{p}'}{p'+q'}$$

$$\Rightarrow P_0 = \frac{q'}{p'+q'}$$

} Long term probability of 0's and 1's

The system no longer remains sensitive.

If $p' = 0.1$ and $q' = 0.2$

$$P_1 = \frac{1}{3} \qquad P_2 = \frac{2}{3}$$

Same is for

$p' = 0.000001$ and $q' = 0.00001$

but the point at which this will be achieved is more further away.

EXAM ↑

probability based question long-term

Probability transformations through state transformation diagram.

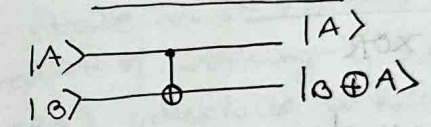
The above change of state is called Markov Process.

In the long run, the Markov process cannot be identified.

This system only considers 1-bit. What about seeing 2 at a time?

20/01/2024

Controlled-NOT



if A=0,	B	O/P
	0	0
	1	1

if ~~A=1~~ (set)	B	O/P
	0	1
	1	0

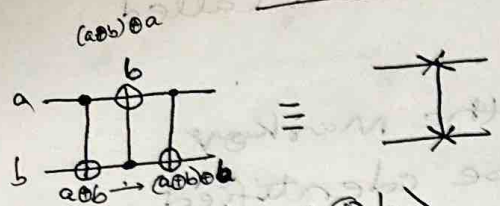
} |A⟩ is Controlled.

controlled qubit ↓ Target qubit ↓

A	B	O/P
|00⟩		|0 0⟩
|01⟩		|01⟩
|10⟩		|11⟩
|11⟩		|10⟩

Swapping.

2 i/p, 2 o/p gates



$|a,b\rangle \rightarrow |a, a\oplus b\rangle$

$\rightarrow |a\oplus(a\oplus b), a\oplus b\rangle = |b, a\oplus b\rangle$

$\rightarrow |a, (a\oplus b)\oplus b\rangle = |b, a\rangle$

Reversibility (slide 44)

CNOT → Generalisation of XOR

U_{CN} ← Controlled NOT ; Unitary

$U_{CN}^{\dagger} U_{CN} = I$

Classical XOR & NAND ← Non invertible, Irreversible

Quantum gates are always invertible ← inverse of unitary matrix is unitary (No info loss & across the gate), i.e. you can get back to the input side.

Basis states

$|+\rangle = \frac{(|0\rangle + |1\rangle)}{\sqrt{2}}$ $\quad |-\rangle = \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$

Rotate 45°

(How?) Photon spin → apply controlled energy → to reach new state.

$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$

$= \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$ ← wrong in slide 43

$\frac{(|\alpha+\beta|)^2}{2}$, $\frac{(|\alpha-\beta|)^2}{2}$ ← probabilities of $|0\rangle$ and $|1\rangle$

Universality - CNOT

→ Swapping is possible.

→ Any logic can be composed of single and CNOT gates. (Realises initial gates)

→ Requires 3 gates for swapping.

The state must remain coherent for the duration of working of the gates (for all 3 stages). Otherwise it makes no sense.

Quantum circuits

→ No loops (Acyclic)

→ FANIN not allowed (2 or more points coming together), because BITWISE-OR is irreversible. (cannot trace $|0\rangle$ or $|1\rangle$ from o/p)

→ FANOUT not allowed, as qubit cannot be copied.

→ Connection (wires) → not physical, →

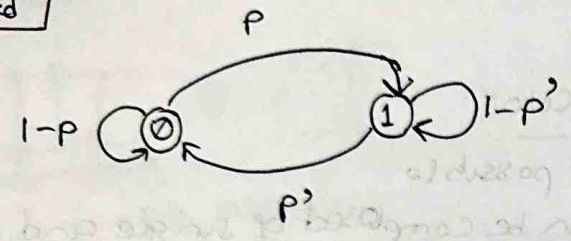
- passage of time
- photon moves from one location to another thru space.
- electron → change state thru energy.

Qubit can't be copied → because upon measurement the qubit state collapses. Amplitudes α and β cannot be copied.

Anti-commutator : $\{A, B\}$

Representation

On the board



long-term prob. (derived earlier)

This is an error channel having very small probability of flipping of bits.

In a symmetric channel, $p = p'$

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & p' & 1-p' \end{array} = P$$

for 1 bit

Extension:

Two bit

Transition matrix:

for 2 bits

$$\begin{array}{c|cccc} & 00 & 01 & 10 & 11 \\ \hline 00 & (1-p)(1-p) & (1-p)p & & \\ 01 & (1-p)p' & (1-p)(1-p') & & \\ 10 & & & & \\ 11 & & & & \end{array}$$

Taking $(1-p)$ common because first bit remains zero

Treat each bit independently so product of individual bits are in transition matrix.

$$\begin{array}{c|cccc} & 00 & 01 & 10 & 11 \\ \hline 00 & (1-p)\begin{bmatrix} P \end{bmatrix} & & p\begin{bmatrix} P \end{bmatrix} & \\ 01 & & & & \\ 10 & p'\begin{bmatrix} P \end{bmatrix} & & (1-p')\begin{bmatrix} P \end{bmatrix} & \\ 11 & & & & \end{array} = \Pi_2$$

$$\begin{array}{c|cccc} & 00 & 01 & 10 & 11 \\ \hline 00 & (1-p)P & & pP & \\ 01 & & & & \\ 10 & p'P & & (1-p')P & \\ 11 & & & & \end{array}$$

$P \to 2 \times 2$

$\Pi_2 \to 4 \times 4$

$\Pi_2 \neq P^2$

what we need for quantum computing? $\{H, U, M_m\}$

→ Registers (set): Collection of qubits in different states, i.e. an array of states

→ Unitary matrix: which will execute some quantum algorithm

→ measurements: measurement leads to collapse. Doing it multiple times gives a probability.

For an n bit register, 2^n is the dimension of the Hilbert space. $(\mathbb{C}^{2^n})$

Initial state: needs to be prepared → like spin of electron, photon particles, etc.

Transformation:- Those that can change the state of photon particle / spin of electron. eg- oscillating magnetic fields, electric fields.

The algorithm must produce such a transformation so that at the end during measurement, the individual states of $|0\rangle$ and $|1\rangle$ can be measured and not a mixture.

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

Classical

- Digital
- Very difficult to control transformations (very challenging for physical realisation)

Quantum

- Digital + Analog
- Qubit takes arbitrary superposition of $|0\rangle$ and $|1\rangle$ hence the coeffs are continuous complex nos.
- Gate contains unitary group with continuous parameters. (Eg: rotation)
- Parameters are continuous & always contain errors

Unitary matrix single gate exploration

→ Identity $\mathbb{I}$ keeps $|0\rangle$ and $|1\rangle$ the same

$$|0\rangle\langle 0| + |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

→ NOT X flips $|0\rangle$ and $|1\rangle$

$$|1\rangle\langle 0| + |0\rangle\langle 1| = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

→ Phase shift Z: $|0\rangle$ is same $|1\rangle$ rotates by π becomes $|-1\rangle$

$$|0\rangle\langle 0| - |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

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

→ Combination Y.

$|0\rangle \rightarrow |-1\rangle$

$|1\rangle \rightarrow |0\rangle$

$$|0\rangle\langle 1| - |1\rangle\langle 0| = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

2-bit qubit systems

CNOT

$|00\rangle \rightarrow |00\rangle$

$|01\rangle \rightarrow |01\rangle$

$|10\rangle \rightarrow |11\rangle$ } ⊛

$|11\rangle \rightarrow |10\rangle$

$$|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X$$

What are the vectors?

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$|11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

⊛ Based on Transformations

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

← 1-qubit NOT gate

Here NOT only on $|1\rangle$.

CNOT $|00\rangle$ $|00\rangle$

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

CNOT $|01\rangle = |01\rangle$

CNOT $|10\rangle = |11\rangle$

CNOT $|11\rangle = |10\rangle$

} same as transformations

CCNOT : Controlled controlled NOT (Toffoli gate)

- 2 control bits are used.
- Control o/p takes place only when both control bits are switched ON.
- It has 8 states

Here NOT only on $|11\rangle$

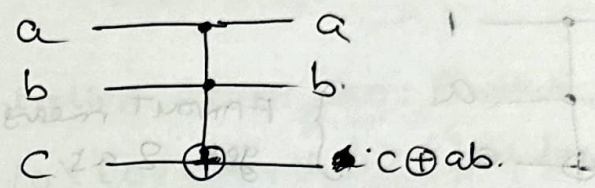
$|110\rangle \rightarrow |111\rangle$
$|111\rangle \rightarrow |110\rangle$

} only these 2 cases are of interest

$$(|100\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|) \otimes I + |11\rangle\langle 11| \otimes X \quad (X = \text{NOT})$$

First quantum gate that gives the flavour of classical computing.

Classical computation using Toffoli gate



Toffoli

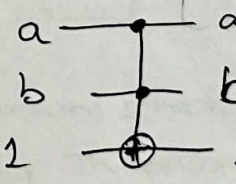
1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0

a	b	c	a'	b'	c'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

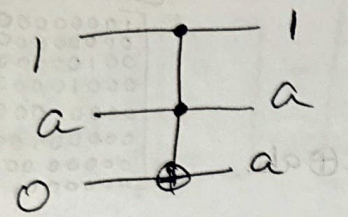
If a and b is one, then CNOT works

c = 1, NAND gate → universal gate can achieve all other classical gates.

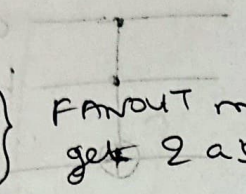
i.e. a & b are not the controls.
a & b are the I/Ps



Toffoli gate



looks like FANOUT



FANOUT means I've got 2 as.

Unitary matrix for SWAP gate

$|00\rangle \to |00\rangle$
$|01\rangle \to |01\rangle$
$|10\rangle \to |10\rangle$
$|11\rangle \to |11\rangle$

$|00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11|$

$$U_{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$ swapped.

$$U_{SWAP}\,|00\rangle = |00\rangle$$

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

$U_{SWAP}\,|01\rangle = |10\rangle$
$U_{SWAP}\,|10\rangle = |01\rangle$.
$U_{SWAP}\,|11\rangle = |11\rangle$

Entanglement of states?

$\alpha|0\rangle + \beta|1\rangle$

Controlled SWAP gate: Fredkin gate

$|0\rangle\langle 0| \otimes I_4 + |1\rangle\langle 1| \otimes U_{SWAP}$

control = 0, no swap

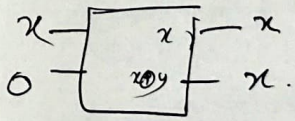
control qubit = 1, swap takes place

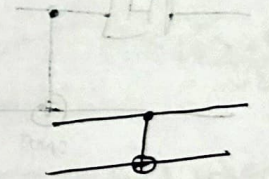
Toffoli & Fredkin gates both originate from single-qubit basic gate.

(Do for) controlled-Z and controlled-phase similarly.

Copying of qubits

classical CNOT





$|\psi\rangle = a|0\rangle + b|1\rangle$, $|0\rangle$ → $a|00\rangle + b|11\rangle$ (inverted due to CNOT)

(can't be copied)

Any unknown state cannot be copied by No Cloning Theorem, ie. if both are $|\psi\rangle = a|0\rangle + b|1\rangle$

Then there will be no a and b.

$(a|0\rangle + b|1\rangle)(a|0\rangle + b|1\rangle)$

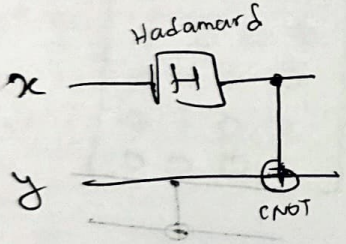
$= a^2|00\rangle + ab|01\rangle + ab|10\rangle + b^2|11\rangle$

Copying means info will be retained. However, measuring this state above will lead to destruction (all is lost)

▷ Probability upon multiple measurements:-

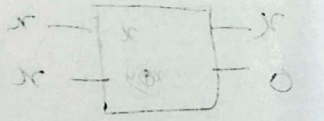
$$\frac{a^2}{a^2+b^2} \qquad \frac{b^2}{a^2+b^2}$$

but we do not intend this, we want to determine exact copy of $|0\rangle$ and $|1\rangle$.



$|\beta_{xy}\rangle$

circuit for copying:



$|00\rangle \to (|00\rangle + |11\rangle)/\sqrt{2} = |\beta_{00}\rangle$

$|01\rangle \to (|01\rangle + |10\rangle)/\sqrt{2} = |\beta_{01}\rangle$

$|10\rangle \to (|00\rangle - |11\rangle)/\sqrt{2} = |\beta_{10}\rangle$

$|11\rangle \to (|01\rangle - |10\rangle)/\sqrt{2} = |\beta_{11}\rangle$

Bell states

Teleportation

27/01/2024

If we have a co-conspirator and share qubits. then upon transformation of one qubit, information about the other is revealed.

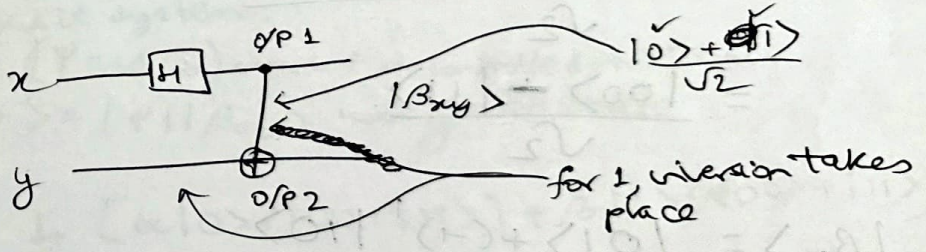
Super-dense coding

while we require 2 classical bits to transmit information, 1 qubit will be enough to perform the same transformation of info.

In Hadamard gate,

$|0\rangle$ becomes $\frac{|0\rangle + |1\rangle}{\sqrt{2}}$

$|1\rangle$ becomes $\frac{|0\rangle - |1\rangle}{\sqrt{2}}$



if $x=0, y=0$

o/p 1) $\frac{|0\rangle + |1\rangle}{\sqrt{2}}$ (0 thru Hadamard.

o/p 2) $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ (control, no inv., inverted)

property

$|00\rangle$ $|10\rangle$

$|01\rangle$ $|11\rangle$

$\Psi_+, \Psi_- / \Phi_+, \Phi_-$

Bell states

$$|\beta_{xy}\rangle = \frac{|0,y\rangle + (-1)^x |1,\bar{y}\rangle}{\sqrt{2}}$$ } General form.

$$|\beta_{00}\rangle = \frac{|00\rangle + (-1)^0 |11\rangle}{\sqrt{2}} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

$$|\beta_{01}\rangle = \frac{|01\rangle + (-1)^0 |10\rangle}{\sqrt{2}} = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$

$$|\beta_{10}\rangle = \frac{|00\rangle + (-1)^1 |11\rangle}{\sqrt{2}} = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

$$|\beta_{11}\rangle = \frac{|01\rangle + (-1)^1 |10\rangle}{\sqrt{2}} = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

These are called EPR pairs

→ If $|\psi\rangle$ contains only one of a ~~and~~ or b, then it is copyable → like $|0\rangle$ or $|1\rangle$.

Teleportation

Deliver $|\psi\rangle$ from A to B.

$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$; unknown α and β

we want to get back $\alpha|0\rangle + \beta|1\rangle$
we can obtain other states through flipping using Pauli gates X and Z.

α and β can be recovered from any of the 4 states.

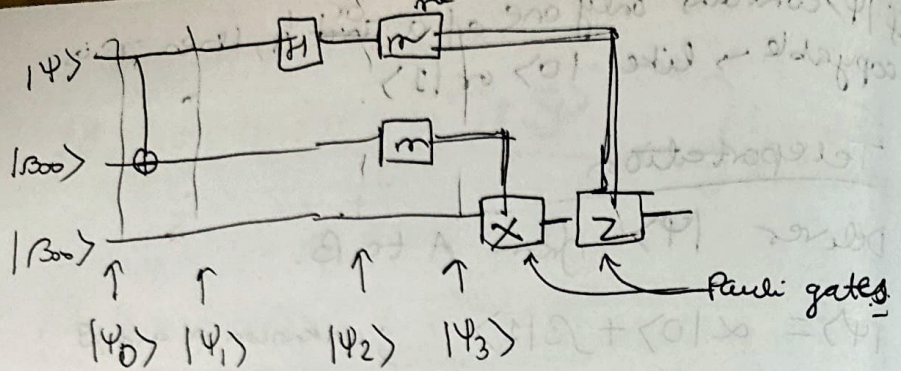
Using Bell states, we are entering into a 3-qubit system.
(ψ and β_0) interact in controlled-NOT system.

$|\psi_0\rangle = |\psi\rangle|\beta_{00}\rangle$ using Bell state

$$= \frac{1}{\sqrt{2}}\left[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)\right]$$

The third addition qubit is given to sender A.

~~$\frac{1}{\sqrt{2}}\alpha(|000\rangle + |100\rangle$~~



CNOT) $|\psi_1\rangle = \frac{1}{\sqrt{2}}\left[\alpha|0\rangle(|00\rangle+|11\rangle) + \beta|1\rangle(|10\rangle+|01\rangle)\right]$

H) $|\psi_2\rangle = \frac{1}{2}\left[\alpha|0\rangle+|1\rangle)(|00\rangle+|11\rangle) + \right.$
$\left.\beta|0\rangle-|1\rangle)(|10\rangle+|01\rangle)\right]$

$|\psi_2\rangle = \frac{1}{2}\left[|00\rangle(\alpha|0\rangle+\beta|1\rangle) + \right.$
$|01\rangle(\alpha|1\rangle+\beta|0\rangle) +$
$|10\rangle(\alpha|0\rangle\mp\beta|1\rangle) +$
$\left.|11\rangle(\alpha|1\rangle-\beta|0\rangle)\right]$

Each of them occurs with equal probability

for $|00\rangle$, α and β can be obtained easily
for others, we need X, Z or combination of both to obtain α and β.

Now, we realise)
A measurement leads us to some deterministic thing (communication of α and β)

i.e. measurement does not result in an entire collapse.

The fact that they shared the EPR pair, so they are able to convey a state.

Does Teleportation work?

- Alice needs to communicate measurement to Bob → over classical channel!?
- Measurement of original qubit leads to collapse.

Bell state helps in realising teleportation, which looked promising, but failed to succeed ultimately.

31/01/2025

→ Any one of these 4 states may occur with equal probability, phase flip is required to obtain a different state from the current.

In case of two distinguishing claims b/w classical and quantum theories, the result obtained through controlled experimentation is accepted.
Eg: Bell's Inequality.

Superdense coding

In classical, whatever information can be transmitted through 2bits, could be done using 1 qubit communication in quantum. & a pre-shared entanglement

In teleportation, the states must be shared.

Here similarly, the entangled states are shared.

To be sent	Apply	Outcome
00	I	$\frac{\|00\rangle + \|11\rangle}{\sqrt{2}}$
01	Z	$\frac{\|00\rangle - \|11\rangle}{\sqrt{2}}$
10	NOT X	$\frac{\|10\rangle + \|01\rangle}{\sqrt{2}}$
11	iY	$\frac{\|01\rangle - \|10\rangle}{\sqrt{2}}$ (i=XZ)

This is done in the sender's end.

Measurement method is applied on the Bell states, & one of these states are obtained. Accordingly we can determine the 2-bit that were sent.

These 4 are the orthogonal states & so they can be measured

on the board

In classical channel → ?

channel capacity is dependent on noise and external factors.

Polarisation of photon particles ← gives the 0-state, 1-state and entanglement. which current state is a photon? Apply a laser beam to change its state.

Binary Symmetric Channel

(0,1) ← distortion of zero to one and vice-versa is equal, with probability 'p'

If p=0.01, then 1 in 100 0's becomes 1 & vice-versa.

Repetitive code ← though not feasible in classical, is the only achievement in quantum thus far.

0 is transmitted

0) 0 0 0

1) 1 1 1

	Probability
0 0 0	$(1-p)^3$
0 0 1	$(1-p)^2 p$
0 1 0	"
1 0 0	"
0 1 1	$(1-p) p^2$
1 0 1	"
1 1 0	"
1 1 1	p^3

$$(p+1-p)^3 = 1^3 = 1 \quad (*)$$

The distortion takes place independently

~~Probability of~~

If we set up a rule,

If ∅ is majority, ∅ was transmitted

If 1 is majority, 1 was transmitted

This rule fails in the last 4 cases

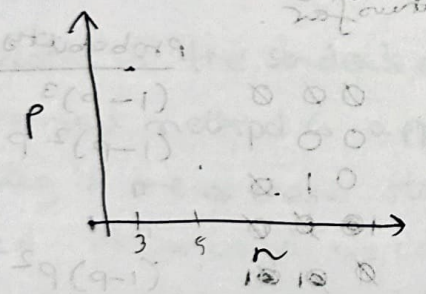
$p^3 + 3p^2(1-p)$ $\quad p = 0.01 = 10^{-2}$

$= 10^{-6} + \frac{3 \times 0.99}{2.97} \times 10^{-4}$ ← Dominates

$\simeq 3 \times 10^{-4}$

3 in 10000 is a significant improvement over 1 out of 100.

If no. of repitions is increased, the p keeps on decreasing



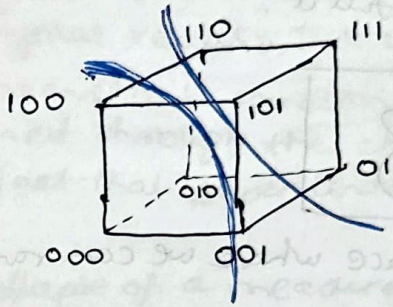
Redundancy

Message rate suffers, ie. to get 1 bit through, we are taking 3 times the time, ∴ Improvement is coming at a cost.

~~Classical has noise~~

Shannon (1948) showed that for every channel there is a term associated called 'channel capacity'.

Geometrical essence



2 out of 8 are valid points

In an n-dimensional space with k valid points, there is a sphere of attraction of recieve words which are at a certain distance from the valid codewords.

The receiver receives 1 out of 2^n codewords.

Let p be the prob. of distortion of 1 bit

Expected probability = np

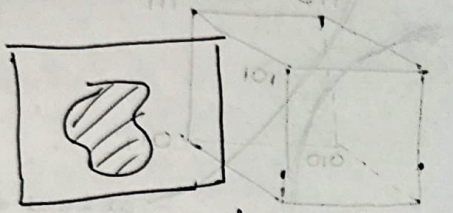
If receiver tries to draw a sphere of radius 'np' around the received codeword, then the receiver can expect to find one valid codeword within that sphere.

However, more than one valid codewords may appear in this sphere. (if there is no restriction on the distribution of valid codewords).

Probabilities of getting none, one or more no. of valid codewords in the sphere can be

determined.

Eg: let there be a pond of uniform depth in a plot of land. what is the ~~probability~~ amt of soil needed to fill it?



This is a 2-D space where we can randomly select points. let

n → No. of points in lake

N → Total no. of random points selected.

$\frac{n}{N}$ ← ratio of lake area to total area, $N \to \infty$

Shannon ~~was~~ tried to find out how many codewords can the channel tolerate?

If we randomly choose codewords out of 2^n possibilities, the errors can be counted out through the above procedure.

No. of valid codewords is restricted by channel capacity, which in turn is dependent on p.

In the super dose coding, 2 bit as 1 qubit has 50% message rate → can be scaled up.

⇒ Measurement only succeeds for ortho-normal states.

→ Entangled state is prepared and distributed among both parties, by a 3rd party.

→ In physical reality, two zeroes are not sent. This means that something has already happened through the channel. (claim for the fact that superdense coding doesn't work)

- Upon collapse of a measurement, some idea about the data is obtained.

The possibilities

→ what type of computations?
Something that is already existent with a classical computer.

→ can it be better?
Improvement of computational complexity.

→ can we implement classical in quantum?
First check if existing gates are achievable. NAND gate universal, but what about complexity?

> Quantum jump :- replacing a high complexity portion of a classical algo with a faster quantum equivalent.

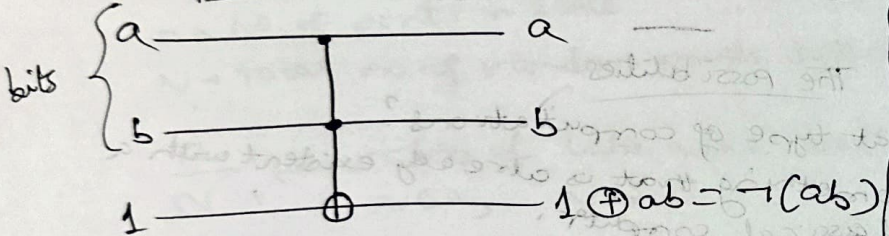
- Fineman: developed full adder circuit in quantum.
- Reversible gates (TOFFOLI GATES) reduce 10% of total energy in classical.

Resonator → like LC (classical)
which has voltage fluctuators to allow reversibility, to get back the energy

EXAM

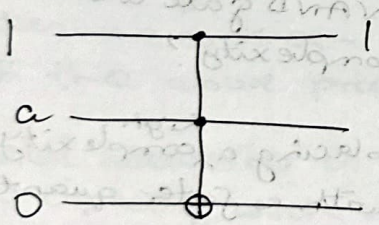
Q) How can I perform some kind of classical computation (say XOR) using Toffoli gate?

NAND using Toffoli:



Not bothered at c: It will flip when both a and b are 1

FANOUT using Toffoli



If $a = \alpha|0\rangle + \beta|1\rangle$

Generating random bits using quantum (Non-deterministic) | classical

prepare qubit in state $|0\rangle$
pass through Hadamard ← Rotates it by 45°
outcome is $|0\rangle$ or $|1\rangle$
measure it, probability is $\frac{1}{2}$ for each state $|0\rangle$ or $|1\rangle$

Quantum parallelism 07/02/2025

$$|x, y\rangle \xrightarrow{U_f} |x, y \oplus f(x)\rangle$$

At $y = 0$, final state $= f(x)$.

How are we bringing parallelism?

$|0\rangle$ is passed through Hadamard gate to get

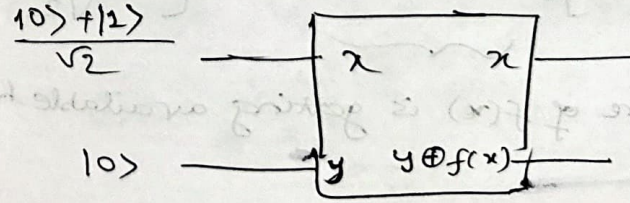
$$\frac{(|0\rangle + |1\rangle)}{\sqrt{2}}$$

Resultant:

$$\frac{|0, f(0)\rangle + |1, f(1)\rangle}{\sqrt{2}}$$; so $f(0)$ and $f(1)$ are computed parallelly & simultaneously.

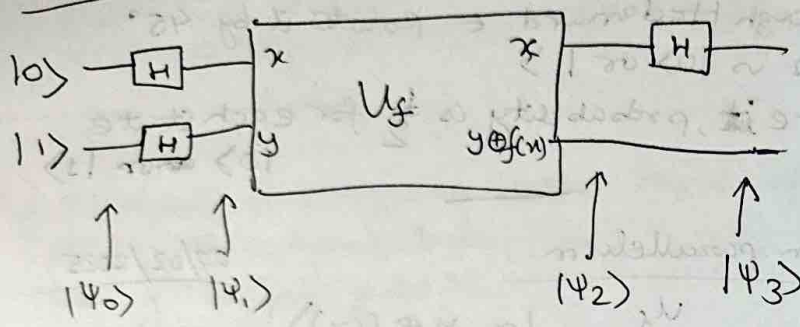
Applying n Hadamard gates in parallel..

$$H^{\otimes 2} = \frac{(|0\rangle + |1\rangle)}{\sqrt{2}} \cdot \frac{(|0\rangle + |1\rangle)}{\sqrt{2}} = \frac{(|00\rangle + |01\rangle + |10\rangle + |11\rangle)}{2}$$



$\frac{1}{\sqrt{2^n}} \sum_x |x\rangle |f(x)\rangle$

Deutsch algorithm



Some property involving $|0\rangle$ and $|1\rangle$ is obtained.
One evaluation can expose the property through measurement (just like superdense coding)

The property: If $f(0) != f(1)$

$$|\psi_0\rangle = |01\rangle$$

$$|\psi_1\rangle = \left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$$

$$|\psi_2\rangle = \pm\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \quad \text{when } f(0)=f(1)$$

$$= \mp\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \quad \text{when } f(0)\neq f(1)$$

$$U_f\left(|x\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]\right) = (-1)^{f(x)}|x\rangle\left[\frac{(|0\rangle-|1\rangle)}{\sqrt{2}}\right]$$

Signature of $f(x)$ is getting available here.

$$|\psi_3\rangle = \pm|f(0)\oplus f(1)\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]$$

Measure this to get 0 or 1, it collapses, yet we can deterministically say what really happened.

Here there is no sharing of qubits unlike superdense coding.

Tweaking both inputs is possible ahead of the Hadamard gate.

Deutsch Jozsa Algorithm

Boolean fn of n variables is balanced or constant. (why?)

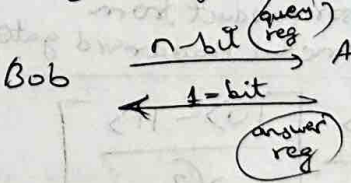
If n=10
0000 ← we can't say if it is balanced.
00000 ← now it is not balanced.

order: $O(n/2)$

How many bits are needed to check for balanced! (Non-balance leads to more prone to attacks from intruders).

In classical, brute force combination check is needed: $O\left(\frac{2^n}{2}+1\right)$

In quantum, exchanging qubits & calculating $f(x)$ through U_f in a single query.



Deutsche Algo Internal working

$|\psi_0\rangle = |x, 1\rangle$

$|\psi_1\rangle = |x, H\cdot 1\rangle = \dfrac{|x,0\rangle - |x,1\rangle}{\sqrt{2}}$

$|\psi_2\rangle = \dfrac{|x, 0\oplus f(x)\rangle - |x, 1\oplus f(x)\rangle}{\sqrt{2}}$

$f(x)=0 \;:\; |\psi_2\rangle = \dfrac{(|x,0\rangle - |x,1\rangle)}{\sqrt{2}}$

$f(x)=1 \qquad |\psi_2\rangle = \dfrac{(|x,1\rangle - |x,0\rangle)}{\sqrt{2}}$

$|\psi_2\rangle = (-1)^{f(x)} \dfrac{(|x,0\rangle - |x,1\rangle)}{\sqrt{2}}$

Hadamard Transform

$H|0\rangle = \dfrac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad H|1\rangle = \dfrac{|0\rangle - |1\rangle}{\sqrt{2}}$

$H|x\rangle = (-1)^{x\cdot 0}|0\rangle + (-1)^{x\cdot 1}|1\rangle$

If z assumes the 2 states, then

$(-1)^{xz}|z\rangle$

$H^{\otimes n}|x_1,\ldots,x_n\rangle$ (Tensor product from fundamental Hadamard gate)

$= \sum_z \sum_x \dfrac{(-1)^{x\cdot z + f(x)}|z\rangle}{\sqrt{2^n}} \left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$

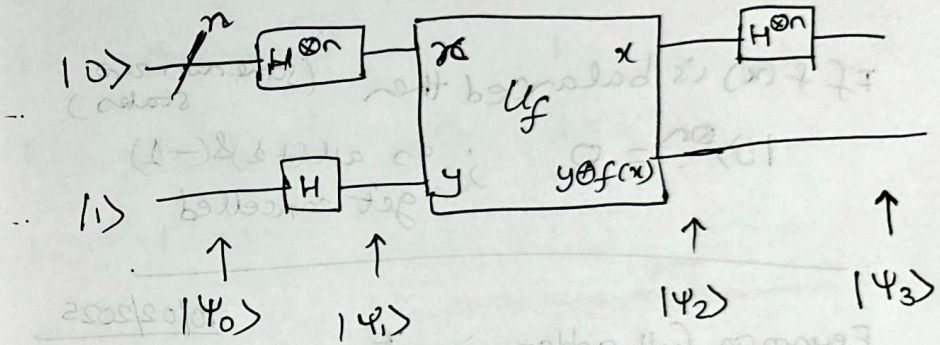
If $|x_1,\ldots,x_n\rangle = |0,\ldots,0\rangle$, as $x\cdot z$ flies away

$\equiv (-1)^{f(x)}|z\rangle$

Deutsch JOSZA Algo working

Q) Query (n-qubit)

A) Answer (1-qubit)



O/p → 0 if f is constant.

order of one evaluation of U_f. (Decoherence will not set in earlier)

$|\psi_0\rangle = |0\rangle^{\otimes n}|1\rangle$

$|\psi_1\rangle = \sum \dfrac{|x\rangle}{\sqrt{2^n}}\left[\dfrac{(|0\rangle - |1\rangle)}{\sqrt{2}}\right]$

$|\psi_2\rangle = \sum \dfrac{(-1)^{f(x)}|x\rangle}{\sqrt{2^n}}\left[\dfrac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$

$H|x\rangle = \sum_z (-1)^{xz}|z\rangle/\sqrt{2}$

$H|x_1,\ldots x_n\rangle = \left[\sum_{z_1,\ldots,z_n} (-1)^{x_1z_1+\cdots x_nz_n}|z_1,\ldots,z_n\rangle\right]/\sqrt{2^n}$

$$|\psi_3\rangle = \sum_z \sum_x \frac{(-1)^{x\cdot z + f(x)}}{2^n} |z\rangle \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right]$$

$$|0\rangle^{\otimes n} \quad \sum_x (-1)^{f(x)}/2^n$$

If $f(x)$ is constant, then (all $|0\rangle$ states)

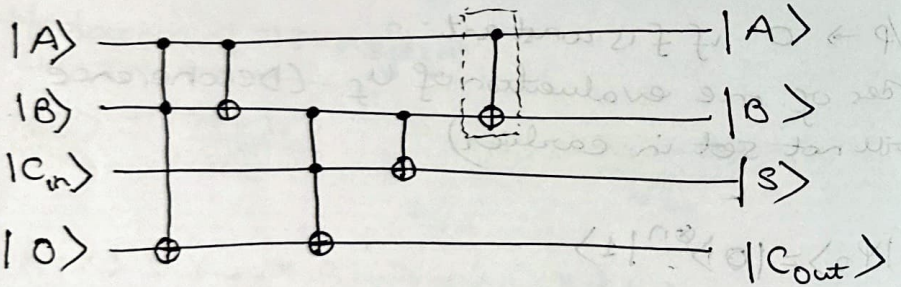
$|0\rangle^{\otimes n} = (-1)$ or $(+1)$

If $f(x)$ is balanced then (some other zero states)

$|0\rangle^{\otimes n} = 0$; as all $(+1)$ & (-1) get cancelled

10/02/2025

Feynman full adder circuit



$|S\rangle = |A \oplus B \oplus C_{in}\rangle$

$|C_{out}\rangle = |(A \wedge B) \oplus |0\rangle\rangle \oplus |(A \oplus B \wedge C_{in}\rangle$

$\|A\rangle$	$\|B\rangle$	$\|C_{in}\rangle$	$\|S\rangle$	$\|C_{out}\rangle$
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Truth Table of Full Adder

PPT midsem

Information channel

A channel has a number of symbols both on the input and output sides.

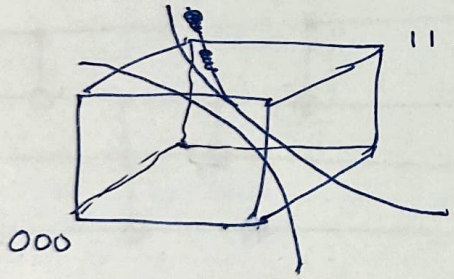
channel matrix:

$$\begin{bmatrix} P(b_1/a_1) & P(b_2/a_1) & \cdots & P(b_s/a_1) \\ P(b_1/a_2) & P(b_2/a_2) & \cdots & P(b_s/a_2) \\ \vdots & & & \\ P(b_1/a_r) & P(b_2/a_r) & \cdots & P(b_s/a_r) \end{bmatrix}$$

For sender: $P(b_j/a_i)$ → FORWARD

For receiver: $P(a_i/b_j)$ → BACKWARD

Receiver computes $P(a_i/b_j)$ for all b_j for each a_i and chooses the highest value.



If prob of flip from 0 to 1 is p, then

$P(0/100) = p(1-p)(1-p)$

$P(1/100) = (1-p)\,p\,p$

If obtained codeword is 100 how to determine what was sent

If $p = 10^{-5}$, then

$P(0/100) < P(1/100)$

Hence we determine that 0 was sent.

Binary symmetric channel

$$\begin{bmatrix} p' & p \\ p & p' \end{bmatrix}$$ ← discussed earlier.

Probability Relation

$$P(b_j) = \sum_{i=1}^{r} P(b_j/a_i)\,P(a_i)$$

↑ Mathematical expectation

$$P(a_i/b_j) = \frac{P(a_i)\,P(b_j/a_i)}{P(b_j)}$$

$$P(a_i/b_j) = \frac{P(b_j/a_i)}{P(a_i)} \sum_{i=1}^{r} P(b_j/a_i)\,P(a_i)$$

Example.

Let

$P(0) = 0.6$

$P(1) = 0.4$

$P(1/0) = 0.1 \Rightarrow P(0/0) = 0.9$

$P(0/1) = 0.2 \Rightarrow P(1/1) = 0.8$

What is prob of 0 and 1 at receiver end?

$P(\text{Rec}=0) = P(\text{sent}=0) \times P(0/0) + P(\text{sent}=1) \times P(0/1)$

$= 0.6 \times 0.9 + 0.4 \times 0.2$

$= 0.54 + 0.08 = 0.62$

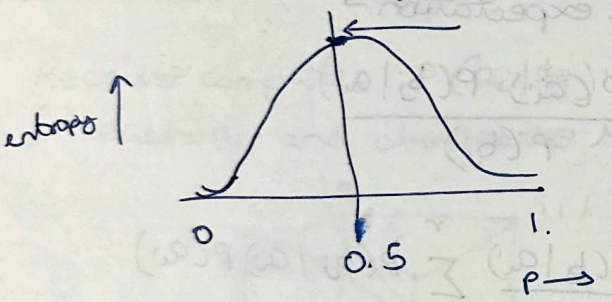
Eg initially 60 out of 100 zeroes were sent but it increased to 62 out of 100 were received.

This is because 1's are becoming 0's with higher probability, ie. $P(0|1)$

Joint Probability

$$P(a_i, b_j) = P(b_j|a_i)P(a_i)$$
$$= P(a_i|b_j)P(b_j)$$

Entropy $= p_i \log \frac{1}{p_i}$



At 0.62, the entropy is less than that at 0.6. Hence, the channel is stealing some information. The channel adds its own flavour of uncertainty.

There is an A posterior Entropy for every b_j.

$$H(A/B) = \sum_B P(b)\, P(A/b)$$

Joint Entropy

$$H(A, B) = \sum_j \sum_i P(a_i, b_j) \log \frac{1}{P(a_i, b_j)}$$

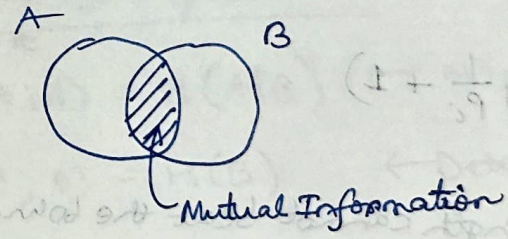
$$H(A, B) = H(B/A) + H(A)$$

mutual Information

$$MI = H(A) - H(A/B).$$

Diff b/w orgrial uncertainty of A and given that B is received what is the uncertainty of A.

$$I(A, B) = H(A) + H(B) - H(A, B)$$



Channel capacity is where the mutual Information maximises

Aposteriori entropy is conditioned on the received signal.

Shannon's First Theorem

If we know the probability of occurence of symbols, then we know entropy.
we can choose symbols having high probability and code them with smaller alphabet length, and vice versa.

Then the average length of the transmission reduces.
This is bounded by the info entropy.

length of symbol

$$L_i = \left\lceil \log \frac{1}{P_i} \right\rceil$$

← taking floor leads to loss of no. of symbols for alphabets.

Avg length

$$\bar{L} = \sum P_i L_i$$

$$= \sum P_i \left(\log \frac{1}{P_i} + 1 \right)$$

$$= H$$

i.e. the avg. length cannot beat the bound of info entropy.

For each b_j there is an optimal code

L_{ij} ← length assigned to i^{th} symbol but conditioned by the length of the j^{th} received symbol.

a priori → $H(A)$

aposteriori → $H(A/b_j)$ Output symbols occur with probs $P(b_j)$

$$H(A/B) = \sum_B P(b) H(A/b)$$

$$= \sum_j$$

Mutual Information

Diff b/w entropy at the source and the aposteriori entropy

$$I(A;B) = \sum_j \sum_i P(a_i, b_j) \log \frac{P(a_i, b_j)}{P(a_i) P(b_j)}$$

Special types of channels

$$I(A;B) = H(A/B)$$ ← Noiseless

$$I(A;B) = H(B)$$ ← Deterministic

→ Cascading : Not Important

$$I(A;B) = H(B) - H(B/A)$$

Shannon's theorem for channels

$H(A|B) \leq L < H(A|B)+1$

if every b_j has its own code, then n^{th} extension

Channel capacity

Noise creates distortion and there is no way to recover.

Recovery can be done through redundance by taking more alphabets. However, this increases time of message being sent.

→ Repeating codes : has already been implemented.

Channel capacity ⇒ where mutual information maximises.

Dependent on probability distribution of symbols on input side. So if the probs are equal, it maximises.

How the channel works on 0 and 1 is a property of the channel itself.

In a Binary Symmetric channel, with error prob 'p'

Capacity = $1 - H(p)$

Signal to Noise Ratio in dB is measured, not distortion.

In reality, 0 does not flip to 1, in physical terms, it is the change of SNR ratio over a threshold.

Uniform channels

one symbol in each row has a very high probability. other cases are low probability.

In BSC, 0 remains 0 at high probability and becomes zero at very low prob.

$$\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$ ← Rows are permutations of each other.

$$I(A;B) = H(B) - H(B/A)$$

$$= H(B) - \sum_A P(a) \underbrace{\sum_B P(b|a) \log \frac{1}{P(b|a)}}_{W}$$

↑ because every row is a permutation of other rows

$$= H(B) - W \sum P(a)$$

For noiseless, $W = 0$

'w' represents the whiteness of the noise. At every frequency, same amplitude; over a large range of frequencies.

If SNR is high, signal dominates.

For uniform

$I(A;B) = H(B) - w$

{ H(B) is a channel, i.e. receiver side probs

~~H(B) becomes log r~~.

For uniform channel,

$P(0|1) = P(1|0)$

Hence H(B) can attain max value of log r.; where output contains 'r' no. of symbols.

$$I(A;B) = \log r + \sum_B P(b/a) \log \frac{1}{P(b/a)}$$

---

n-bit error detecting binary code

I/P symbols = n-1

parity bit = 1.

$q = 2^{n-1}$ (No of I/P symbols)

No. of errors = k

Q → prob of error.

P → prob of no error

$C(n,k)\, Q^k\, P^{n-k}$; error count

$$w = \sum_{k=0}^{n} C(n,k)\, P^{n-k}\, Q^k \log \frac{1}{P^{n-k} Q^k}$$

$$= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} P^{n-k} Q^k (n-k) \log \frac{1}{P}$$

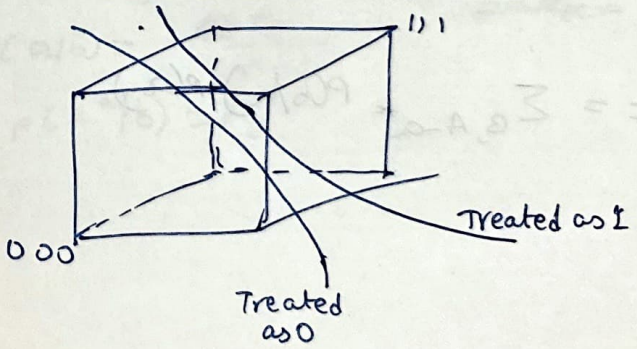
$$= nP \sum_{k=0}^{n} \frac{n!(n-1)!}{k!(n-k-1)!} P^{n-k-1} Q^k \log \frac{1}{P}$$

$$= nP (P+Q)^{n-1} \log \frac{1}{P}$$

$P+Q = 1$

$$= nP \log \frac{1}{P}$$

Decoding in Noisy channel



Decision rule based on Hamming distance.

Mapping of every o/p symbol to an i/p symbol.

- Message Rate
- Decision Rule

Prob of error

- Sum of all cases leading to error
- Every set of decision rules with have its corresponding prob of errors

Maximum likelihood decision rule (a^*) based on what is received.

$$\frac{p(b_j/a^*)\,p(a^*)}{p(b_j)} \geq \frac{p(b_j/a_i)\,p(a_i)}{p(b_j)}$$

$$P_E = \sum_B p(E/b)\,p(b)$$

(error)

Joint occurence of a^* together with b is $\bar{P}_E = 1 - P_E$

$$P_E = \sum_{B,\,A-a^*} p(b/a)\,p(a)$$

The Fano Bound

Decision rule applies to the type of redundancy added, individually. probability of success is higher for more redundancy but time taken hugely increases.

$$H(P_E) = P_E \log \frac{1}{P_E} + \bar{P}_E \log \frac{1}{\bar{P}_E}$$

$$H(P_E) + P_E \log (r-1) = P_E \log \frac{r-1}{P_E} + \bar{P}_E \log \frac{1}{\bar{P}_E}$$

$$= \sum_{B,\,A-a^*} p(a,b) \log \frac{r-1}{P_E} + \sum_B p(a^*,b) \log \frac{1}{\bar{P}_E}$$

$$H(A/B) = \sum_{B,\,A-a^*} p(a/b) \log \frac{1}{p(a/b)} + \sum_B p(a^*,b) \log \frac{1}{p(a^*/b)}$$

.

$$H(A/B) \leq H(P_E) + P_E \log (r-1)$$; each case is equally erroneous.

Condition of equality : $x = 1$

$$p(a/b) = \frac{P_E}{r-1}$$

$$p(a^*/b) = P_E$$

How many codewords can we put in an n-dimensional hyperspace.

n^{th} extention with r symbols

How many messages can be put into an n-dimensional hyperspace and yet have low prob of errors.

Reliable decoding

$$M < 2^{nC}$$

$$\text{message rate} = \frac{\log M}{n}$$

$$M = 2^{n(C+\epsilon)}, \quad (\epsilon \to 0)$$

For n^{th} extension,

M.I $= H(A^n) - H(A^n/B^n) \leq nC$

($H(A^n)$: Not in PDF; nC: channel capacity)

$$p = \frac{1}{M}$$

$$H(A^n) = \sum \frac{1}{M} \log M = \log M = n(C+\epsilon)$$

$$n(C+\epsilon) - nC = n\epsilon \leq H(A^n/B^n)$$

Fano bound

$$n\epsilon \leq H(A^n/B^n) \leq H(P_E) + P_E \log(q-1)$$

max value of entropy is 1

$$n\epsilon \leq 1 + P_E (nC + C\epsilon)$$

$$P_E \geq \frac{n\epsilon - 1}{nC + n\epsilon} > \frac{\epsilon - \frac{1}{n}}{C + \epsilon} \geq \frac{\epsilon}{C}$$

So, prob of error is bounded away from a quantity, which is not desired.

Hence, P_E cannot be kept small.

Fano introduces channel equivocation and mutual info.

Pf of of shannon's main theorem for BSC

$$C = 1 - H(p)$$

Choose 'M' messages from 2^n codewords.

with replacement

Repeating this multiple times gives measure of prob. of errors, but it becomes unacceptable.

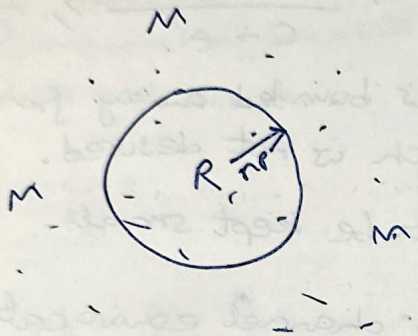
$$M = 2^{(n*(C-\epsilon))} \quad ; \text{say } n = 26$$

Eraser channel → some bits are removed, those are considered as 'x'.

This selection is done randomly

Then an appropriate decision rule needs to be chosen.

Say received word is R.



A hypersphere with certain radius is constructed around R. The radius of hypersphere is kept increasing till atleast one valid codeword appears within the sphere.

No. of bits = n

~~Expected~~ error prob per bit = p

expected error = np.

What is min radius to get a valid codeword?

Say n=100, p=0.1, radius=10.

So, around np radius around R, we should get a valid codeword → this should be the decision rule.

There could be ambiguity in choice of codewords due to various sources of error :-

→ more than one

→ wrong one

- Bounded by channel capacity → Radius.
- We cannot stuff too many M's in the radius region. (gets congested)

17/02/2025

Proof of Shannon's main theorem for BSC

Error prob = p

Capacity ⇒ $C = 1 - H(p)$

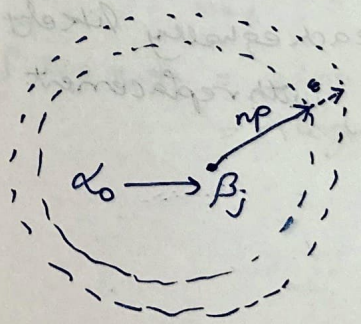
$M = 2^{n(C-\epsilon)}$ ← valid codewords that we are choosing.

⇒ How many M out of 2^n is safe for transmission, to keep the error probability small.

If M is clustered, then there would be a space where no codewords are present, while in another space there would be many. Constructing a valid codeword is a challenge. But choice of M is not important here.

Let, $\alpha_0 \xrightarrow{I/P} \beta_j \xrightarrow{O/P}$

- prob of error = p
- Expected no. of bits in error = np



np is avg dist b/w α_0 and

ϵ → small value added to radius in order to atleast get one codewo

Sources of error
→ x_0 is not in sphere
→ more than one codeword in the sphere

$$P_E = Pr\,x_0 \notin S(n,P_e) + Pr\,x_0 \in S(n,P_E) \times Pr\,(\text{atleast another codeword} \in S(n,P_E))$$

$$Pr\,x_0 \in S(n,P_E) \leq 1$$

$$P_E \leq Pr\,x_0 \notin S(n,P_E) + Pr\,(\text{atleast another codeword} \in S(n,P_E))$$

$$\Rightarrow P_E \leq \underbrace{Pr\,x_0 \notin S(n,P_E)}_{\text{prob. of More than } n(P+\epsilon) \text{ errors}} + \sum_{x_i \neq x_0} Pr\,x_i \in S(n,P_E)$$

$$P_E \leq \delta + \underbrace{\sum Pr\,x_i \in S(n,P_E)}_{\text{prob. that other codewords in the sphere.}}$$

$(M-1)$

Random code → M out of 2^n is chosen randomly
Getting an avg over all such possible codes.
↳ Avg. random code
i.e. avg. over all kinds of distributions of M, etc.

→ 2^{nM} possible codes selected each equally likely.
$\therefore Prob = \frac{1}{2^{nM}}$; with replacement

Avg. probability of error

$$P_E \leq \delta + (M-1)\,P(a \in S(r))$$
$$\leq \delta + M\,P(a \in S(r))$$

No. of points that can be counted in the sphere out of 2^n, gives the proportion of codewords out of 2^n.

$N(r)$ → count of codewords in sphere
r ← Hamming distance.

$$P(a \in S(r)) = \frac{N(r)}{2^n}$$

for BSC,

$N(r)$ = Sum of no. of ways r no of bits can change

$$= 1 + {}^nC_1 + {}^nC_2 + \dots + {}^nC_r$$

$$N(r) = \sum_{k=0}^{r} {}^nC_k$$

2

Stirling approximation

$$n! \approx n^n e^{-n} \sqrt{2\pi n}$$

$$\log n! = \sum_{k=1}^{n} \log_e k$$

$$\int_1^n \log x\,dx = x\log x - x$$
$$= n\log n - n + 1$$

$C(n, \lambda n) \approx \left[\frac{1}{\lambda^{\lambda n}(1-\lambda)^{(1-\lambda)n}}\right] \left[\frac{}{2\pi\lambda(1-\lambda)n}\right]^{1/2}$

$\lambda = p$

$2^{nH(\lambda)}$ $\qquad$ $\text{const}\,(\lambda)\, n^{-1/2}$

$$\sum_{k=0}^{\lambda n} {}^{n}C_k \leq 2^{nH(\lambda)}$$

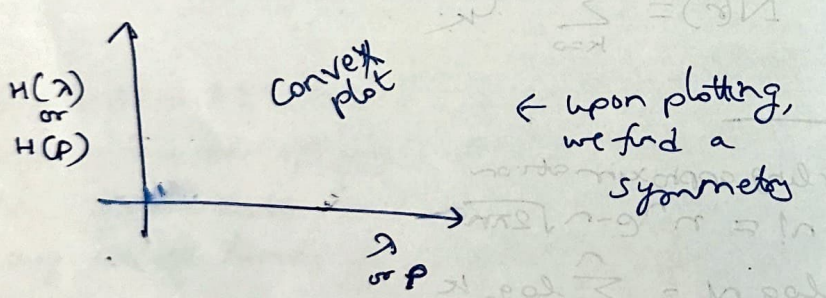
Stirling (NOT imp.)

$N(r) \leq 2^{nH(\lambda)}$

$P(a \in S(r)) \leq 2^{-n(1-H(\lambda))}$

$$P_E \leq \delta + M\, 2^{-n(1-H(\lambda))}$$

for BSC; $C = 1 - H(p)$



← upon plotting, we find a symmetry

anything of the form

$$H(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}.$$

$$P_E \leq \delta + M.2^{-n(C-\epsilon_3)}$$

slightly less than channel capacity.

$$M = 2^{n(C-\epsilon_1)}$$

$$\tilde{P}_E \leq \delta + M.2^{-n(G-\epsilon_3)}.$$ Avg. case

Larger n, smaller $\tilde{P}_E$.

Increasing redundancy may improve error probability but no. of codewords M out of 2^n is very low, like 2 out of 8, or 2 in 32, etc

---

Generalisation of Shannon's Theorem

→ Hamming distance metric → X
→ Counting no. of messages in sphere → difficult
→ Calc. channel capacity -
→ for non-white noise, sphere becomes oval

---

alternatives