Communicate Without Errors

Zifan Hua

November 23, 2022

Table Of Contents

Table of Contents

Introduction

- Channel Capacity
- Zero Error Rate
- Shannon Capacity
- C_n

Table of Contents

Channel

Definition (channel)

A channel has a sender and a receiver.

Channel

Definition (channel)

A channel has a sender and a receiver.
The sender sends a finite sequence of characters to the receiver which is called message.

Channel

Definition (channel)

A channel has a sender and a receiver.
The sender sends a finite sequence of characters to the receiver which is called message.

The receiver receives the sequence of characters and decode it.

Channel

Definition (channel)

A channel has a sender and a receiver.
The sender sends a finite sequence of characters to the receiver which is called message.

The receiver receives the sequence of characters and decode it. However, in the procedure of send and receive, the channel may introduce some errors. For instance, here, b is decoded into c. Worth to note that this may not be "abelian". Which means b have chance to be decoded into c, but c may not have the chance to be decoded into b.

Channel

Definition (confusable)

Given two characters a, b. If a and b have chance to be decoded into a same character say c, we say a and b are confusable.
For a two messages of length n, we say that they are confusable if the two messages can be decoded into the same message.

Channel

Definition (confusable)

Given two characters a, b. If a and b have chance to be decoded into a same character say c, we say a and b are confusable.
For a two messages of length n, we say that they are confusable if the two messages can be decoded into the same message.

Definition (rate of channel)

Given a channel that could send m distinct messages with length n. And could be able to send only one character per unit time. The rate of channel is $\sqrt[n]{m}$.
Here, we do not care about how much characters we have or how many characters we used. We only care about the number of all messages we can send.

Channel

Definition (zero error rate)

Given a channel which could transfer messages of length n.
We want to find the maximum set of messages M that no two of them is confusable.

$$\text{zero error rate} = \max_{M} \sqrt[n]{|M|} \tag{1}$$

Channel

If given a set of characters S, and some of the characters could be confusable.
We want to find

$$\sup_n \{\text{zero error rate of channel with length } n \text{ and } S \text{ as characters}\} \tag{2}$$

Graph Representation Of Characters

Definition (graph)

A graph G is a set of vertices with a set of edges connecting pairs of vertices.

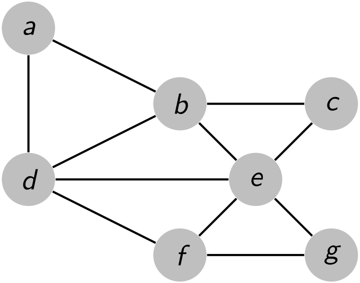


Figure: An example of a graph.

Graph Representation Of Characters

Definition

Given a channel that sending $\{1, 2, \ldots, n\}$ as characters. And some characters i and j could be confused with each other.

Graph Representation Of Characters

Definition

Given a channel that sending $\{1, 2, \ldots, n\}$ as characters. And some characters i and j could be confused with each other.
Then the graph representation of the characters is the graph with vertices $\{1, 2, \ldots, n\}$ and edges (i, j) if and only if i and j could be confused with each other.

$\alpha(G)$

Definition ($\alpha(G)$)

The $\alpha(G)$ represent the maximum number of characters that could not be confused with each other in graph G.

$\alpha(G)$

Definition ($\alpha(G)$)

The $\alpha(G)$ represent the maximum number of characters that could not be confused with each other in graph G.
Given a graph G. Given a subgraph H of G, such that every vertex of H is not connected in G. Then $\alpha(G)$ is the maximum number of vertices of H.

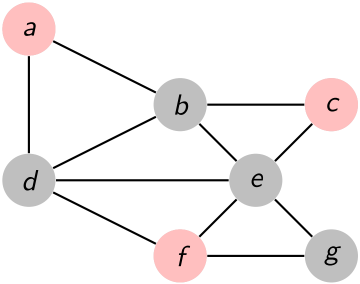


Figure: Example of $\alpha(G)$. Here $\alpha(G) = 3$.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one character or one message.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one character or one message.
Given two graph G and H. The graph product $G \times H$ is the graph with vertices $V(G \times H) = V(G) \times V(H)$ in which (x, y) is adjacent to (x', y') in $G \times H$ if and only if x is adjacent to x' in G and y is adjacent to y' in H.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one character or one message.
Given two graph G and H. The graph product $G \times H$ is the graph with vertices $V(G \times H) = V(G) \times V(H)$ in which (x, y) is adjacent to (x', y') in $G \times H$ if and only if x is adjacent to x' in G and y is adjacent to y' in H.
A graph G product itself for n times will always be denoted by G^n.

Properties of $\alpha(g)$

Lemma

$\alpha(G)\alpha(H) \leq \alpha(G \times H)$

Properties of $\alpha(g)$

Lemma

$\alpha(G)\alpha(H) \leq \alpha(G \times H)$

Proof.

Given graph G and H. Let G' and H' be subgraph of G and H such that no vertex of G' or H' is adjacent in G or H, respectively. Then $G' \times H'$ is a subgraph of $G \times H$ such that no vertex of $G' \times H'$ is adjacent in $G \times H$. $\square$

Shannon Capacity

Definition (Shannon Capacity)

The Shannon capacity of a channel is defined on sending a series if characters.

Shannon Capacity

Definition (Shannon Capacity)

The Shannon capacity of a channel is defined on sending a series if characters.
Given a graph G, the Shannon capacity $\Theta(G)$ is defined by

$$\Theta(G) = \sup_n \sqrt[n]{\alpha(G^n)} \tag{3}$$

Shannon Capacity

Definition (Shannon Capacity)

The Shannon capacity of a channel is defined on sending a series if characters.
Given a graph G, the Shannon capacity $\Theta(G)$ is defined by

$$\Theta(G) = \sup_n \sqrt[n]{\alpha(G^n)} \tag{3}$$

We will see that this sequence has an upper bound later.

Orthonormal Representation

Definition (Orthonormal Representation)

Given a graph G with vertices $1, 2, \ldots, n$, the orthonormal representation of G is a set of unit vectors $\{v_1, v_2, \ldots, v_n\}$ such that v_i and v_j are orthogonal if i and j are not adjacent in G.

Orthonormal Representation

Definition (Orthonormal Representation)

Given a graph G with vertices $1, 2, \dots, n$, the orthonormal representation of G is a set of unit vectors $\{v_1, v_2, \dots, v_n\}$ such that v_i and v_j are orthogonal if i and j are not adjacent in G.

Any orthonormal basis of $\mathbb{R}^n$ can be used as an orthonormal representation of G. So, for any graph G, there exist some orthonormal representations.

Tensor Product

Definition (tensor product)

Given two vectors $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$, the tensor product $v \circ w$ is defined by

$$v \circ w = (v_1 w_1, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n) \quad (4)$$

Tensor Product

Definition (tensor product)

Given two vectors $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$, the tensor product $v \circ w$ is defined by

$$v \circ w = (v_1 w_1, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n) \quad (4)$$

Lemma

The inner product of tensor products can be computed by,

$$\langle v \circ w, v' \circ w' \rangle = \langle v, v' \rangle \langle w, w' \rangle \quad (5)$$

Product of Orthonormal Representation

Lemma

Given a graph G with vertices $1, 2, \ldots, n$, and a graph H with vertices $1, 2, \ldots, m$. Then vectors $\{v_i \circ w_j\}$ is an orthonormal representation of $G \times H$.

Theta Function

Given a graph G, the $\theta(G)$ is defined by

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{6}$$

where $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal representation of G, and c is any unit vector does not orthogonal to v_i.

Theta Function

Given a graph G, the $\theta(G)$ is defined by

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{6}$$

where $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal representation of G, and c is any unit vector does not orthogonal to v_i.

Lemma

There always exist such an c and orthonormal representation $\{v_1, v_2, \ldots, v_n\}$ *such that*

$$\theta(G) = \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{7}$$

This could be proved by proving the set of all possible cases of $\{v_1, v_2, \ldots, v_n, c\}$ is compact. And the function $\max \frac{1}{\langle c, v_i \rangle^2}$ is continuous.

Theta Function

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{8}$$

Theta Function

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{8}$$

Proof.

Let $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_m\}$ be orthonormal representation of G and H and c_v and c_w such that

$$\max\left\{\frac{1}{\langle c_v, v_i\rangle^2} : i = 1, 2, \ldots, n\right\} = \theta(G) \tag{9}$$

and

$$\max\left\{\frac{1}{\langle c_w, w_i\rangle^2} : i = 1, 2, \ldots, m\right\} = \theta(H) \tag{10}$$

□

Proof.

Then

$$
\begin{aligned}
\theta(G \times H) &\leq \max\left\{\frac{1}{\langle c_v \circ c_w, v_i \circ w_j \rangle^2}\right\} && (11)\\
&= \max\left\{\frac{1}{\langle c_v, v_i \rangle^2 \langle c_w, w_j \rangle^2}\right\} && (12)\\
&= \max\left\{\frac{1}{\langle c_v, v_i \rangle^2}\right\} \max\left\{\frac{1}{\langle c_w, w_j \rangle^2}\right\} && (13)\\
&= \theta(G)\theta(H) && (14)\\
& && (15)
\end{aligned}
$$

□

Table of Contents

Theorem (Spectral Theorem)

- *Let T be a symmetric operator on a Euclidean space V.*
  - *There exist an orthonormal basis of V consisting of eigenvectors of T.*
  - *The eigenvalues of T are real numbers.*
- *Let A be a real symmetric matrix.*
  - *There exist an orthogonal matrix P such that $P^T AP$ is a real diagonal matrix.*
  - *The eigenvalues of A are real numbers.*

Table of Contents

Lemma

$$\theta(G) \geq \alpha(G) \tag{16}$$

Lemma

$$\theta(G) \geq \alpha(G) \tag{16}$$

Proof.

Let $\{1, 2, \ldots, k\}$ be the set of vertices of G such that every point is not adjacency in G. And $k = \alpha(G)$
Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal representation of G and c such that

$$\max\left\{\frac{1}{\langle c, v_i\rangle^2}\right\} = \theta(G) \tag{17}$$

Then

$$1 \quad = \quad c^2 \geq \sum_{i=1}^{k} \langle c, v_i\rangle^2 \geq \frac{k}{\theta(G)} \tag{18}$$

□

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{19}$$

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{19}$$

Proof.

$$
\begin{aligned}
\Theta(G) &= \lim_n \sqrt[n]{\alpha(G^n)} && (20)\\
&\leq \lim_n \sqrt[n]{\theta(G^n)} && (21)\\
&\leq \lim_n \sqrt[n]{\theta(G)^n} && (22)\\
&= \theta(G) && (23)
\end{aligned}
$$

$\square$

The final result we want to prove today,

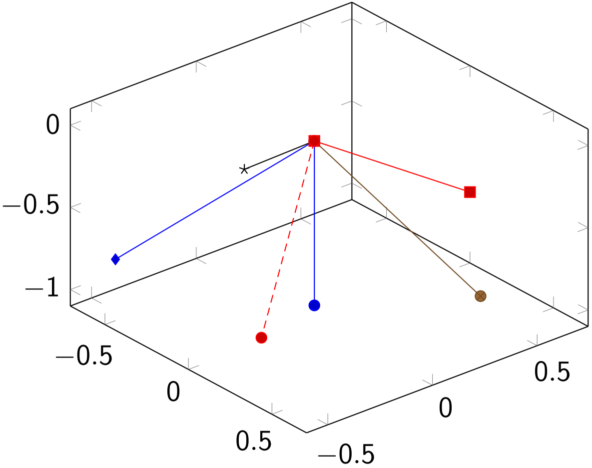
$$\Theta(C_5) \leq \theta(C_5) \leq \sqrt{5} \tag{24}$$

Theorem

$$\theta(C_n) \leq \sqrt{5} \tag{25}$$

Proof.

Consider an umbrella that has a handle and 5 ribs that all have unit length. And also its handle is a unit vector.



□

Proof.

Here, angles between two consecutive ribs are same.

Let v be a rib. And let w be one of the rib that have the largest angle with v. Then, we let the angle between v and w be $\pi/2$.

Then, the 5 ribs of such an umbrella form an orthonormal representation of C_5.

Let d be the vector represent the handle, and $v_1, v_2, \ldots, v_5$ be the 5 ribs. And let γ be the angle between the handle and any rib.

So, by some calculation, we get

$$\begin{aligned}
\theta(C_5) &\leq \max \frac{1}{< d, v_i >^2} && (26) \\
&= \left(\frac{1}{\cos(\gamma)}\right)^2 && (27) \\
&= \sqrt{5} && (28)
\end{aligned}$$

□

Theorem

Let G be a graph with n vertices. Let $A = (a_{i,j})$ be a symmetric matrix such that

$$a_{i,j} = \begin{cases} 1 & \text{if } i = j \text{ or } i \text{ and } j \text{ are not adjacent in } G \\ \text{arbitrary} & \text{otherwise} \end{cases} \tag{29}$$

Then

$$\theta(G) \leq \max\{\text{eigenvalues of } A\} \tag{30}$$

This theorem gives us some way to estimate or even compute the value of $\theta(G)$ in some special cases.

Proof.

Let A be a matrix satisfying the requirement. Let λ be its largest eigenvalue. Then $\lambda I - A$ is positive semidefinite, and thus exits vectors $\{x_1, x_2, \ldots, x_n\}$ such that

$$\lambda \delta_{i,j} - a_{i,j} = \langle x_i, x_j \rangle \tag{31}$$

Let c be a unit vector perpendicular to x_i, and let,[1]

$$u_i = \frac{c + x_i}{\sqrt{\lambda}} \tag{32}$$

We will prove that u_i is an orthonormal representation of G. And also we have

$$\frac{1}{< c, u_i >^2} = \lambda \tag{33}$$

So, $\theta(G) \leq \lambda = \max\{\text{eigenvalues of } A\}$ $\square$

[1] Why such c and $\{x_1, x_2, \ldots, x_n\}$ exist, will need some linear algebra. If we have time, we could come back to here.

Proof.

Then

$$
\begin{aligned}
< u_i, u_i > &= \frac{1}{\lambda}\langle c + x_i, c + x_i \rangle && (34)\\
&= \frac{1}{\lambda}(\langle x_i, x_i \rangle + 1) && (35)\\
&= \frac{1}{\lambda}(\lambda - 1 + 1) && (36)\\
&= 1 && (37)
\end{aligned}
$$

Also, for nonadjacent i and j,

$$
\begin{aligned}
< u_i, u_j > &= \frac{1}{\lambda}\langle c + x_i, c + x_j \rangle && (38)\\
&= \frac{1}{\lambda}(\langle x_i, x_j \rangle + 1) && (39)\\
&= 0 && (40)
\end{aligned}
$$

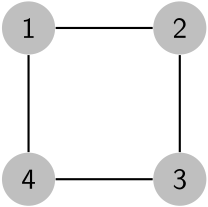
□

Adjacency Matrix

Definition (Adjacency Matrix)

Given a graph G with vertices $\{1, 2, \ldots, n\}$, the adjacency matrix A is defined by

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ is adjacent in } G \\ 0 & \text{otherwise} \end{cases} \tag{41}$$



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

Regular Graph

Definition (regular graph)

A regular graph is a graph where each vertex has the same number of neighbors.

Theorem

Let G be a regular graph, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of its adjacency matrix A. Then

$$\theta(G) \leq n\frac{1-\lambda_n}{\lambda_1 - \lambda_n} \tag{42}$$

This theorem actually gives us a really good estimate of $\theta(G)$ for regular graphs.

Theorem

Let G be a regular graph, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of its adjacency matrix A. Then

$$\theta(G) \leq n\frac{1-\lambda_n}{\lambda_1-\lambda_n} \tag{42}$$

This theorem actually gives us a really good estimate of $\theta(G)$ for regular graphs.
Also, as C_n is a regular graph, and its adjacency graph is a circulant matrix, we could actually compute $n\frac{1-\lambda_n}{\lambda_1-\lambda_n}$ for C_n.
And we could get the result we want

$$\theta(C_n) \leq \frac{n\cos(\pi/n)}{1+\cos(\pi/n)} \tag{43}$$

Here we let J be the matrix with all entries equal to 1.

Proof.

Consider the matrix $J - x(A - I)$ where x is a real number. Then $J - x(A - I)$ actually met the demand of the proceeding theorem. So, its largest eigenvalue is greater or equal then $\theta(G)$.

Let the eigenvalues of A be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. We could actually proof that eigenvalues of $J - x(A - I)$ are $n - \lambda_1 x + x, -\lambda_2 x + x, -\lambda_3 x + x, \ldots, -\lambda_n x + x$.

So, the largest eigenvalue of $J - x(A - I)$ is either $n - \lambda_1 x + x$ or $-\lambda_n x + x$.

Then, choose x to be

$$x = \frac{n}{\lambda_1 - \lambda_n} \tag{44}$$

Then

$$\theta(G) \leq n \frac{1 - \lambda_n}{\lambda_1 - \lambda_n} \tag{45}$$

$\square$

Here we let j be the vector with all entries equal to 1.

Proof.

So, here we proof that eigenvalues of $J - x(A - I)$ are
$n - \lambda_1 x + x, -\lambda_2 x + x, -\lambda_3 x + x, \ldots, -\lambda_n x + x$.
As, G is a regular graph, so suppose every vertex of G has k neighbors, which represent in its adjacency matrix A is that every row or every column of A has exactly $k + 1$ ones and others are zeros.
So, clearly j is a eigenvector of A with eigenvalue $k + 1$. And also j is a eigenvector of J with eigenvalue n.

$\square$

Proof.

Next, we will prove that $\lambda_1 = k + 1$

Suppose, w is a eigenvector of A with eigenvalue $\lambda > k + 1$. Let w_i be the entry of w that have the largest absolute value. And r be the i^{th} row of A. Then

$$\begin{aligned}
|\lambda w_i| &= \left|\sum_{j=1}^{n} r_j w_j\right| && (46)\\
&= \left|\sum_{r_j \neq 0} w_j\right| && (47)\\
&\leq \sum_{r_j \neq 0} |w_j| && (48)\\
&\leq (k+1)\,|w_i| && (49)
\end{aligned}$$

Contradiction, so $\lambda_1 = k + 1$. So, $n - \lambda_1 + x$ is an eigenvalue of $J - x(A - I)$ with eigenvector j. $\square$

Proof.

Then by Spectral Theorem, we could find eigenvectors $\{v_2, v_3, \ldots, v_n\}$ of A corresponding to eigenvalues $\{\lambda_2, \lambda_3, \ldots, \lambda_n\}$.
And also $\{v_2, v_3, \ldots, v_n\}$ are orthogonal to j.
So, $\{v_2, v_3, \ldots, v_n\}$ are also eigenvectors of J with eigenvalue 0.
So, eigenvalues of $J - x(A - I)$ are
$n - \lambda_1 x + x, -\lambda_2 x + x, -\lambda_3 x + x, \ldots, -\lambda_n x + x.$ □

To actually compute the eigenvalues of C_n will use some properties of circulant matrix. I just directly give the result.

$$\begin{aligned} \lambda_1 &= 3 && (50) \\ \lambda_n &= 1 - 2\cos(\pi/n) && (51) \end{aligned}$$

So, we finally get

$$\theta(C_n) \le \frac{n\cos(\pi/n)}{1+\cos(\pi/n)} \tag{52}$$

Table of Contents

Conclusion

- We have proved that $\Theta(C_n)$ is smaller or equal then $n\frac{\cos(\pi/n)}{1+\cos(\pi/n)}$.
- Thus proved that $Theta(C_5) = \sqrt{5}$

Open Questions

- Although we have proved that $\Theta(C_n)$ is smaller or equal then $n\frac{\cos(\pi/n)}{1+\cos(\pi/n)}$, but we still don't know the exact value of $\Theta(C_n)$. Even for $n = 7$
- Is there any good lower bound for $\Theta(C_n)$?
- Is there any patterns for n such that $\Theta(c_n)$ is hard to compute?

Table of Contents