# Capacidade de canais simétricos possuindo grupos cíclicos como alfabeto de entrada

Jorge Pedraza Arpasi

***Resumo*— A capacidade de canal $C$, estabelecido por Shannon, é independente de uma possível estrutura interna do alfabeto de entrada do canal $X$. Também, o teorema de codificação para canais com ruido não depende de qualquer estrutura interna de $X$. No entanto, todos os esquemas de codificação lineares precisam que $X$ possua alguma estrutura algébrica, seja de corpo ou anel. A estrutura algébrica de $X$ é induzida a partir do alfabeto da fonte de informação $U$, via o mapeamento de codificação. Neste trabalho estudamos o caso em que $U = \mathbb{Z}_{p^r}$, onde $\mathbb{Z}_{p^r} = \{0, 1, 2, \ldots, p^r - 1\}$ é um grupo cíclico com $p$ primo, e consideramos $C_U$:="capacidade do canal com $U$ como alfabeto de entrada". Pelo teorema de processamento de dados, sabemos que $C_U \leq C$, mas para o caso $U = \mathbb{Z}_{p^r}$ provamos que esta codificação atinge a capacidade de canal, isto é, $C_U = C$.**

***Palavras-Chave*— Capacidade de canal, canais simétricos, integral de Lebesgue, grupos cíclicos**

***Abstract*— The channel capacity $C$, setup by Shannon, is independent from any internal structure of the channel input alphabet $X$. Also, the noisy channel coding theorem do not depend from any internal structure of $X$. However, all the major linear encoding schemes require that $X$ must have some kind of algebraic structure that can be a field or else a ring. This algebraic structure of $X$ is induced from the alphabet of information source $U$ by the encoding mapping. In this work we study the case $U = \mathbb{Z}_{p^r}$, where $\mathbb{Z}_{p^r} = \{0, 1, 2, \ldots, p^r - 1\}$ is a cyclic group with $p$ prime, and consider $C_U$:= "channel capacity with $U$ as input alphabet". By the data processing theorem, we know that $C_U \leq C$, but for the case $U = \mathbb{Z}_{p^r}$ we show that the coding reach the channel capacity, that is, $C_U = C$.**

***Keywords*— Channel capacity, symmetric channels, Lebesgue integral, cyclic groups.**

## I. INTRODUÇÃO

Dados conjuntos finitos e não vazios $X$, $Y$, sejam $\boldsymbol{X}$ e $\boldsymbol{Y}$ variáveis aleatórias (VA), definidas sobre $X$ e $Y$ respectivamente e seja $P(x|y) = P(\boldsymbol{X} = y|\boldsymbol{Y} = x)$ uma distribuição de probabilidade condicional de ocorrência de $y \in Y$ dado $x \in X$. Então, a tripla $(X, Y, P)$ determina um **canal discreto** com alfabeto de entrada $X$, alfabeto de saída $Y$, e matriz de transições de probabilidade $P(y|x)$. Dadas a distribuição conjunta de probabilidade $P(\boldsymbol{X} = x, \boldsymbol{Y} = y) = p(x, y)$, e as marginais $P(\boldsymbol{X} = x) = p(x)$ e $P(\boldsymbol{Y} = y) = p(y)$ a informação mútua das VAs $\boldsymbol{X}$ e $\boldsymbol{Y}$ para estas distribuições é dada por $I(\boldsymbol{X}; \boldsymbol{Y}) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right)$, [1], [2], [3]. Se log esta na base 2, então $I(\boldsymbol{X}; \boldsymbol{Y})$ é medido em bits. Porem, utilizando adequadamente as formulas bayesianas das probabilidades condicionais, podemos obter a fórmula para $I(\boldsymbol{X}; \boldsymbol{Y})$, que depende somente da distribuição marginal $p(x)$ na entrada do canal, e a matriz de transições de probabilidade do canal $P(y|x)$. Esta fórmula é dada por;

$$I(\boldsymbol{X}; \boldsymbol{Y}) = \sum_{x \in X} p(x) \sum_{y \in Y} \left( P(y|x) \log \left( \frac{P(y|x)}{\sum_{z \in X} p(z) P(y|z)} \right) \right) \quad (1)$$

A conveniência da fórmula (1) é que permite definir a capacidade de um canal discreto $(X, Y, P)$, como uma otimização da informação mútua $I(\boldsymbol{X}; \boldsymbol{Y})$ sobre a família das distribuições marginais de probabilidade definidas sobre a VA $\boldsymbol{X}$ do alfabeto de entrada $X$, mais precisamente, a capacidade $C$ do canal $(X, Y, P)$ é:

$$C = \max_{p(x)} \{ I(\boldsymbol{X}; \boldsymbol{Y}) \}. \quad (2)$$

Para canais discretos sem memória com ruído, a capacidade de canal $C$ de qualquer canal discreto sem memória tem a seguinte propriedade: Para cada $\epsilon > 0$ e $R < C$ existem: 1) um código de bloco de tamanho $N$ e taxa $R$, com $N$ suficientemente grande, e 2) um algoritmo de decodificação tal que a probabilidade máxima de erro de decodificação é $< \epsilon$, [1], [4]. Um código de bloco é um mapeamento $\phi : U^L \to X^N$, onde $U$ é o alfabeto da fonte de informação, onde $L$ é o tamanho dos blocos do alfabeto da fonte $U$, e $N$ é o tamanho dos blocos do alfabeto da entrada do canal. Na demonstração construtiva do teorema de codificação acima, a taxa de codificação $R$ é definida por $R = \frac{\log(|U|^L)}{N}$. Note-se que $C$ é independente de qualquer estrutura interna do conjunto $X$. No entanto os códigos mais conhecidos são lineares, no sentido que $\phi : U^L \to X^N$ é uma transformação linear e $U^L$ e $X^N$ são espaços vetoriais sobre algum corpo de Galois adequado, sendo que para o caso binário este corpo é $GF(2^m)$. Isto significa que para o caso dos códigos lineares, deve ser introduzido uma estrutura algébrica interna, sobre o alfabeto $X$ da entrada do canal. Esta estrutura algébrica pode ser *direta* como é o caso dos canais binários com $X = GF(2)$, e pode ser *induzida* como é o caso dos canais M-PSK onde o alfabeto $X$ é uma constelação de pontos do plano $\mathbb{R}^2$. Sobre $X = \{s_0, s_1, \ldots, s_{M-1}\}$ é induzido uma estrutura algébrica com o mapeamento (casamento) $\mu : \mathbb{Z}_M \to X$, onde $\mathbb{Z}_M = \{0, 1, \ldots, M - 1\}$ é o anel de inteiros módulo $M$. Por outro lado, considere a variável aleatória $\boldsymbol{U}$ definida sobre o alfabeto de informação $U$ e a cadeia de Markov $\boldsymbol{U} \to \boldsymbol{X} \to \boldsymbol{Y}$, isto é, a probabilidade conjunta $p(u, x, y) = p(u)P(x|u)P(y|x)$. Pelo teorema do processamento de dados [4], [2], temos $I(\boldsymbol{U}; \boldsymbol{Y}) \leq I(\boldsymbol{X}; \boldsymbol{Y})$. Considerando os elementos $\boldsymbol{x} \in X^N$, $\boldsymbol{u} \in U^L$, $\boldsymbol{y} \in Y^N$ e remodelando o canal $(X, Y, P)$ no canal $(U, Y, Q)$, onde

Universidade Federal do Pampa - UNIPAMPA, Centro Tecnológico de Alegrete, E-mail: jorgearpasi@unipampa.edu.br

$Q = Q(y|u)$ é a matriz de transições de probabilidade induzida por $P_N(\boldsymbol{y}|\phi(\boldsymbol{u})) = P_N(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N} P(y_i|x_i)$, seja $C_U$ a capacidade do canal $(U, Y, Q)$. Temos que $C_U \leq C$. Quando $C_U = C$ dizemos que o esquema de codificação atinge a capacidade de canal, [5], [2]. Para o caso dos canais binários, onde $X = GF(2^m)$ todos os esquemas de codificação lineares binários atingem a capacidade de canal. Em geral para códigos lineares sobre corpos de Galois arbitrários $GF(p^m)$ também é possível atingir a capacidade de canal. Como todo corpo é um grupo, a família dos grupos é muito mais ampla do que a família dos corpos. Isto motiva o estudo de codificação em canais onde $X$ tenha uma estrutura de grupo. Neste caso nem sempre é possível atingir a capacidade de canal. Por exemplo em [5] é mostrado um esquema de codificação de uma família de canais com $X \subset \mathbb{R}^3$, com estrutura de grupo induzida por $\mathbb{Z}_m \oplus \mathbb{Z}_2$ em um esquema de codificação $\phi$ que não atinge a capacidade de canal.

Neste artigo exibiremos uma prova detalhada sobre codificação de um canal cujo alfabeto de entrada $X$ possui uma estrutura de um grupo cíclico da forma $X = \mathbb{Z}_{p^r}$. Mostraremos que para este caso os blocos da fonte $U$ deve ser organizados utilizando os diferentes subgrupos de $\mathbb{Z}_{p^r}$ e que nem sempre são da forma $U^L$, portanto assumiremos que o homomorfismo de codificação $\phi$ esta definido sobre algum grupo $\mathcal{U}$, com a taxa de codificação $R = \frac{\log(|\mathcal{U}|)}{N}$. Mostraremos que para este caso cíclico $\mathbb{Z}_{p^r}$, a capacidade de canal é atingida. Para isto organizamos este trabalho da seguinte maneira: Na seção II descrevemos os fatos mais importantes de medida e integral de Lebesgue. O intuito desta descrição é que permite definir probabilidades, entropia, capacidade de canal e outros conceitos importantes da teoria da informação de uma maneira unificada, sem distinção dos casos discreto e contínuo. Para o caso estudado neste trabalho, $\mathbb{Z}_{p^r}$, não se nota uma dependência da medida de Lebesgue. Mas, o intuito desta apresentação de Lebesgue é para dar uma posterior continuidade sobre grupos genéricos. Na seção III descrevemos canais sem memória e códigos de bloco e um teorema importante de decodificação por máxima verosimilhança. Damos uma definição de capacidade de canal. Na seção IV descrevemos e damos exemplos de canais simétricos. Na Seção V mostramos que o canal $\mathbb{Z}_{p^r}$-simétrico possui uma capacidade igual á de um canal sem estrutura algébrica.

## II. Probabilidades e Integral de Lebesgue

*Definição 1:* Um sistema de probabilidades consiste numa tripla $(X, \mathcal{A}, P)$ onde, [6]:

1) $X$ é o conjunto dos resultados elementares do experimento, e é chamado de espaço amostral.
2) $\mathcal{A}$ é uma classe de eventos extraído de $\mathcal{P}(X) = \{A \subset X\}$, isto é, se $A \subset X$ então $A \in \mathcal{X}$.
3) Uma *medida de probabilidade* $P : \mathcal{X} \to \mathbb{R}$ tal que;
   a) $P(X) = 1$
   b) $0 \leq P(A) \leq 1$, para todo $A \in \mathcal{X}$.
   c) Se $A, B \in \mathcal{X}$ são disjuntos então $P(A \cup B) = P(A) + P(B)$. □

Os sistemas de probabilidades são classificados para os casos em que $X$ é *finito* e $X$ é *contínuo*. Quando $X$ é finito, isto é, quando a cardinalidade $|X| < \infty$, temos $\mathcal{A} = \mathcal{P}(X)$ e portanto $|\mathcal{P}(X)| = 2^{|X|}$. Neste caso finito, a medida de cada elemento $A \in \mathcal{X}$ é $|A|$, o número de elementos de $A$, e a medida de probabilidade de cada $A \in \mathcal{X}$ é simplesmente $P(A) = \frac{|A|}{|X|}$. Para o caso contínuo $X = \mathbb{R}^n$, o conjunto das partes de $\mathbb{R}^n$, $\mathcal{P}(\mathbb{R}^n)$, além ser *muito grande*, possui elementos muito irregulares que impossibilitam a introdução de alguma medida. É por isso que, neste caso $\mathcal{A}$ é uma subclasse própria de $\mathcal{P}(\mathbb{R}^n)$. Esta subclasse é escolhida de modo que cada elemento seja medível ou **mensurável**. Usualmente $\mathcal{A}$ é constituído por paralelepípedos $n$-dimensionais e as uniões e interseções enumeráveis destes. Observe que para o caso $\mathbb{R}$ os paralelepípedos são os intervalos e para $\mathbb{R}^2$ são os retângulos. Assim, se $X = \mathbb{R}^n$ uma medida de probabilidade de cada $A \in \mathcal{X}$ é dada pela integral de Riemann

$$P(A) = \int_A f(x)dx = \lim_{m\to\infty} \sum_{i=1}^{m} f(x_i)\text{vol}(A_i), \qquad (3)$$

onde $A = \cup_{i=1}^{\infty} A_i$, onde a coleção $\{A_i\}_{i\in\mathbb{N}}$ é tal que cada $A_i \in \mathcal{X}$, $x_i \in A_i$, $\max\{\text{vol}(A_i)\} \to 0$ quando $i \to \infty$, e são dois a dois disjuntos, e $f : \mathbb{R}^n \to [0, \infty)$ é uma função de densidade de probabilidade tal que $\int_{\mathbb{R}^n} f(x)dx = 1$. Observe que a probabilidade $P(A)$, calculada utilizando a integral de Riemann, pode ser interpretada geometricamente como o volume do conjunto $Z = \{(x, y)\ ;\ x \in A, 0 \leq y \leq f(x)\} \subset \mathbb{R}^{n+1}$, ou seja pode-se medir $Z$ como sendo uma união enumerável de paralelepípedos medíveis, dois a dois disjuntos, $Z = \cup_{i\in\mathbb{N}} Z_i$, onde cada $Z_i$ é um paralelepípedo $(n+1)$-dimensional $Z_i = [0, f(x_i)] \times A_i$.

Uma maneira de ampliar a classe de sistemas de probabilidades é ampliando a integral de Riemann. Isto é feito definindo $\mathcal{A}$, em termos das operações de conjuntos ao invés de formas geométricas, no caso os paralelepípedos.

*Definição 2:* Dado um conjunto arbitrário $X$, uma $\sigma$-álgebra de $X$, denotado por $\mathcal{A}$, é um subconjunto de $\mathcal{P}(X)$ tal que:

1) $\mathcal{X}$ não é vazio.
2) Se uma coleção enumerável $\{A_i\}_{i\in\mathbb{N}}$ é tal que cada $A_i \in \mathcal{X}$ então $\cup_{i\in\mathbb{N}}(A_i) \in \mathcal{A}$.
3) Se $A \in \mathcal{X}$ então o complementar $A^c \in \mathcal{X}$. □

Das propriedades 1), 2) e 3) e as leis de DeMorgan decorrem: 4) o conjunto vazio $\emptyset$, e $X$ são elementos de $\mathcal{X}$, 5) se uma coleção enumerável $\{A_i\}_{i\in\mathbb{N}}$ é tal que cada $A_i \in \mathcal{X}$ então $\cap_{i\in\mathbb{N}}(A_i) \in \mathcal{X}$, [7], [8]. Cada elemento $A \in \mathcal{X}$ é subconjunto mensurável de $X$, e o par $(X, \mathcal{A})$ é chamado **espaço mensurável**.

*Definição 3:* Uma medida positiva sobre um espaço mensurável $(X, \mathcal{A})$, é uma função $\mu : \mathcal{A} \to [0, \infty)$ tal que

$$\mu(\bigcup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i), \qquad (4)$$

para cada coleção $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$ onde os $A_i$ são dois a dois disjuntos. □

O tríplice $\Omega = (X, \mathcal{A}, \mu)$ é chamado de **espaço de medida**. Num espaço de medida, pela equação (4), tem-se $\mu(A) \leq \mu(B)$ se $A \subset B$. Além disso, $\mu(X) \geq \mu(A)$, para todo $A \in \mathcal{A}$. Quando $\mu(X) < \infty$, $\mu$ é dito finito. Por outro lado, $\mu$ é dito

$\sigma$-finito quando $X = \cup_{i\in\mathbb{N}} A_i$ e $\mu(A_i) < \infty$ para cada $A_i$, ou seja, o espaço $\Omega = (X, \mathcal{A}, \mu)$ é de medida $\sigma$-finita quando $A$ é a união enumerável de conjuntos de medida finita. É mostrado, por exemplo em [7], que todo sistema de probabilidades é equivalente a um espaço $\Omega$ com $\mu$ sendo $\sigma$-finito. Para definir a integral de Lebesgue precisamos dos conceitos de função mensurável, função característica, e função degrau.

*Definição 4:* Dado um espaço de medida $\sigma$-finita, $(X, \mathcal{A}, \mu)$, uma função $f \to [0, \infty)$ é dita mensurável se para cada conjunto aberto $V \subset [0, \infty)$ tem-se que $f^{-1}(V)$ é um elemento de $\mathcal{A}$. □

Dado um conjunto arbitrário $X$ e $A \subset X$ função característica de $A$ é definida por

$$\chi_A(x) = \begin{cases} 1, \text{ se } x \in A \\ 0, \text{ se } x \notin A. \end{cases} \tag{5}$$

Se $\chi$ está definido sobre a coleção $\mathcal{A}$ de um espaço de medida $\Omega = (X, \mathcal{A}, \mu)$ $\sigma$-finito, então $\chi$ é mensurável. Por outro lado, uma função degrau de múltiplos valores, que chamaremos simplesmente de função degrau, $s : A \to [0, \infty)$ é uma função cuja imagem possui um numero finito de valores $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset [0, \infty)$. Seja $A_i \subset A$ tal que $A_i = s^{-1}(\alpha_i)$, então a função degrau pode ser colocada de maneira precisa como $s(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$. Notemos que a função de Heaviside, de dois valores, é um caso particular. Se a coleção $\{A_i \ ; \ A_i = s^{-1}(\alpha_i)\}$ esta contido em $\mathcal{A}$ então a função degrau é mensurável. Dado um elemento $A \in \mathcal{A}$ a integral de Lebesgue da função degrau é dada por:

$$\int_A s(x)d\mu(x) = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap X). \tag{6}$$

Toda função mensurável $f$ pode ser aproximada por uma família enumerável de funções degrau $\{s_n\}_{n\in\mathbb{N}}$, [7]. Dada uma função mensurável $f$ definida num espaço $\Omega = (X, \mathcal{A}, \mu)$, de medida $\sigma$-finita, considere $S_f = \{s \ ; \ s \text{ é uma função degrau e } s \leq f\}$, então a integral de Lebesgue de $f$, no elemento $X \in \mathcal{A}$, é definida por:

$$\int_A f(x)d\mu(x) = \sup_{s\in S_f} \int_A s(x)d\mu(x). \tag{7}$$

Por ser a integral de Lebesgue uma generalização da integral de Riemann, ambas as integrais são iguais quando $f$ é contínua. A integral de Lebesgue existe mesmo quando o conjunto de descontinuidades de $f$ tenha medida não nula. Dado um espaço $\Omega = (X, \mathcal{A}, \mu)$, uma função mensurável que define uma medida de probabilidade é a função densidade, que é um função $f : X \to [0, \infty)$ tal que $\int_A f(x)d\mu(x) = 1$. A classe de funções densidade de $\Omega$ é denotado por $\mathcal{D}(\Omega)$. Com isto uma medida de probabilidade $P(A)$="probabilidade do subconjunto $X$", sobre um espaço é definida por

$$P(A) = \int_A f(x)d\mu(x). \tag{8}$$

## III. Canais sem memória e códigos de bloco

Na equação 8, definida sobre o espaço $\Omega = (X, \mathcal{A}, \mu)$, $P(A)$ depende das escolhas de $f$ e $\mu$. Se $|X| < \infty$ a medida é $\mu(x)$ ="número de elementos de $A$", portanto $P(A)$ só depende da densidade $f$, nesta caso $P(A) = \int_A f(x)d\mu(x) = \sum_{x\in A} f(x)$, conforme esperado. Então o formalismo abstrato de espaços de medida, e integral de Lebesgue permite o tratamento de sistemas de probabilidades sem fazer maior distinção entra o caso discreto e o caso contínuo. A entropia de cada $f \in \mathcal{D}(\Omega)$ é definida por

$$H(f) = -\int_A f(x)\log(f(x))d\mu(x) \tag{9}$$

*Definição 5:* Um canal com entrada discreta $X$ e saída contínua $Y$ é uma tripla $(X, Y, P)$ tal que;

1) Um espaço de medida $\Omega_X = (X, \mathcal{A}, \mu)$, onde $X$ é um conjunto finito,
2) Um espaço de medida $\sigma$-finita $\Omega_Y = (Y, \mathcal{B}, \nu)$, onde $Y \subset \mathbb{R}^n$,
3) Uma família de densidades de probabilidade $\{P_x, x \in X\}$ tal que $P_x = P(.|x) \in \mathcal{D}(\Omega_Y)$. Esta família de densidades condicionais também é chamada de transições de probabilidade do canal. [5]. □

Quando o canal é **sem memória**, considerando os produtos de espaços $\Omega_X^N = (X^N, \mathcal{A}^N, \mu^N)$, $\Omega_Y^N = (Y^N, \mathcal{B}^N, \nu^N)$, o canal estendido $(X^N, Y^N, P_N)$ é dado por $P_N(\boldsymbol{y}|\boldsymbol{x}) = \Pi_{i=1}^N P(y_i|y_i)$ onde $\boldsymbol{x} = (x_1, x_2, \ldots, x_N) \in X^N$ e $\boldsymbol{y} = (y_1, y_2, \ldots, y_N) \in Y^N$.

Um **codificador de bloco** de comprimento $N$, para o canal $(X, Y, P)$, consiste num conjunto finito $\mathcal{U}$ e um mapeamento $\phi : \mathcal{U} \to X^N$ com taxa de codificação $R = \frac{\log|\mathcal{U}|}{N}$. O decodificador é um mapeamento mensurável $\psi : Y^N \to \mathcal{U}$. Um esquema de codificação é o par codificador e decodificador. Sobre o conjunto $\mathcal{U}$ considere a distribuição de probabilidade uniforme $\boldsymbol{U}$ e seja $\boldsymbol{X}^N$ a variável aleatória $\boldsymbol{X}^N = \phi(\boldsymbol{U})$. Também considere a variável aleatória $\boldsymbol{Y}^N$, sobre $Y^N$, definida pela densidade condicional $P_N(\boldsymbol{y}|\boldsymbol{x})$, donde a densidade marginal é dada por $P_{\boldsymbol{Y}^N}(\boldsymbol{y}) = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} P_{\boldsymbol{Y}^N}(\boldsymbol{y}|\phi(u))$. Se $\hat{\boldsymbol{U}} = \psi(\boldsymbol{U})$ é a estimativa do decodificador, a probabilidade de erro do evento $\hat{\boldsymbol{U}} \neq \boldsymbol{U}$ é denotada por $p_e(\phi, \psi) = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} p_e(\Phi, \psi|u)$ onde $p_e(\Phi, \psi|u) = \int_{Y^N} \chi_A(\boldsymbol{y})d\nu_N(\boldsymbol{y})$, e $A = \psi^{-1}(\mathcal{U} \setminus \{u\})$. É bem conhecido que o esquema de decodificação que minimiza a probabilidade de erro é a chamada de máxima verosimilhança ML (Maximum Likelihood) o mesmo que é definido por $\psi_{ML}(\mathbf{y}) = \max_{u\in\mathcal{U}} P_N(\mathbf{y}|\phi(u))$. Duas consequências imediatas do esquema ML são;

A capacidade de um canal $(X, Y, P)$ genérico, utilizando a medida de Lebesgue, é definido por

$$C = \max_{p\in\mathcal{D}(\Omega_X)} \left\{ \sum_{x\in X} p(x) \int_Y P(y|x) \frac{P(y|x)}{\sum_{z\in X} p(z)P(y|z)} d\nu(y) \right\} \tag{10}$$

## IV. Canais simétricos e grupos

Dado um grupo $G$ com identidade $e_G$ e dado um conjunto arbitrário $X$, é dito que $G$ atua sobre $X$ quando existir um mapeamento sobrejetor $\varphi : G \times X \to X$ tal que:

- $\varphi(e_G, x) = x$, para todo $x \in X$,
- $\varphi(g_1 g_2, x) = \varphi(g_1, \varphi(g_2, x))$, para todo $g_1, g_2 \in G$ e para todo $x \in X$.

O mapeamento $\varphi$ é denominado *ação* de $G$ sobre $X$. A ação $\varphi$ é dita transitiva quando para cada $x, y \in X$ existir $g \in G$ tal que $y = \varphi(g, x)$. A ação $\varphi$ é dita *simplesmente transitiva* quando para cada $x, y \in X$ existir um único $g \in G$ tal que $y = \varphi(g, x)$.

Por outro lado, considere o conjunto $Y$ e o espaço de medida $\Omega = (Y, \mathcal{B}, \nu)$. Um ação $\varphi$ do grupo $G$ sobre $Y$ é dita *ação isométrica* quando $\nu(\varphi(g, B)) = \nu(B)$, para todo $B \in \mathcal{B}$ e para todo $g \in G$.

*Definição 6:* Um canal $(X, Y, P)$ é dito $G$-simétrico se existir um grupo $G$ tal que;

1) Existe uma ação simplesmente transitiva de $G$ sobre $X$
2) Existe uma ação isométrica de $G$ sobre $Y$
3) $P(y|x) = P(\varphi(g, y)|\varphi(g, x))$ para todo $g \in G$, para todo $x \in X$ e para todo $y \in Y$. □

*Exemplo 1:* Considere $G = \mathbb{Z}_2 = \{0, 1\}$, o grupo aditivo binário com a operação soma modulo 2. Esta classe de canais $\mathbb{Z}_2$-simétricos são conhecidos, nas diferentes publicações como canais BIOS (Binary-Input-Output-Symmetric). Uma subclasse de canal BIOS, muito conhecido, é o BSC (Binary Symmetric Channel), onde $X = Y = \{0, 1\} \subset \mathbb{R}$ e $P(1|0) = P(0|1) = \epsilon$, $P(0|0) = P(1|1) = 1 - \epsilon$. Para implementar uma ação $\varphi$ é necessária uma representação adequada dos pontos 0 e 1. Esta representação é dada por $X = Y = \{1, -1\}$, nesta representação $0 \mapsto 1$ e $1 \mapsto -1$. O grupo aditivo $(\mathbb{Z}_2, +)$ é representado pelo grupo multiplicativo $S^0 = \{1, -1\} \subset \mathbb{R}$, isomorfo com $\mathbb{Z}_2$, conforme observamos nas seguintes tabelas:

| $\mathbb{Z}_2$,+ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| $S^0, \cdot$ | 1 | -1 |
|---|---|---|
| 1 | 1 | -1 |
| -1 | -1 | 1. |

Então, $\varphi : S^0 \times \{1, -1\} \to \{1, -1\}$ dada por $\varphi(gx) = g.x$ é uma ação, mais ainda, podemos verificar que $\varphi$ é uma ação simplesmente transitiva. O espaço de medida $(X, \mathcal{A}, \mu)$ na entrada do canal é dado por $\mathcal{A} = \{\emptyset, \{0\}, \{1\}, X\}$ e $\mu(A) = |A|$ para todo $A \in \mathcal{A}$. Analogamente, o espaço de medida $(Y, \mathcal{B}, \nu)$ na saída do canal é dado por $\mathcal{B} = \{\emptyset, \{0\}, \{1\}, Y\}$ e $\mu(B) = |B|$ para todo $B \in \mathcal{B}$. Como $|\varphi(g, B)| = |B|$ para cada $B \in \mathcal{B}$ e para cada $g \in S^0$ então $\nu(\varphi(g, B)) = \nu(B)$. Portanto $\varphi$ é uma ação isométrica na saída do canal.

Um outra subclasse de canal simétrico BIOS é o canal BEC (Binary Erasure Channel) onde $X = \{0, 1\}$, $Y = \{0, 1, 2\}$, $P(1|0) = P(0|1) = 0$ e $P(2|0) = P(2|1)$. Neste caso a implementação da ação simplesmente transitiva do grupo $\mathbb{Z}_2$ sobre $X$ é realizada representando $\mathbb{Z}_2$ por $S^0$ e representando o alfabeto de entrada $X$ de maneira similar para o subcaso BSC. Por outro lado a ação isométrica de $\mathbb{Z}_2$ sobre $Y$ é implementada representando o alfabeto de saída por $Y = \{1, 0, -1\}$, onde $0 \mapsto 1$, $1 \mapsto -1$ e $2 \mapsto 0$. Então, $\varphi : S^0 \times \{1, -1\} \to \{1, -1, 0\}$ dada por $\varphi(gx) = g.x$ é uma ação isométrica. □

*Exemplo 2:* Considere um conjunto finito $X$ de cardinalidade $m \geq 2$ e algum $\epsilon \in [0, 1]$. O canal simétrico $m$-ário é definido pelo tríplice $(X, X, P)$, onde $P(y|x) = 1 - \epsilon$ se $y = x$ e $P(y|x) = \frac{\epsilon}{m-1}$ em outro caso. Neste canal, a probabilidade de recepção do simbolo correto a partir do simbolo transmitido $x$ é $1 - \epsilon$, enquanto que a recepção errónea do simbolo $x$ tem probabilidade $\epsilon$, sendo esta probabilidade uniformemente distribuída em $X - \{x\}$. O caso especial $m = 2$ é BSC. O canal simétrico $m$-ário possui o maior nível de simetria possível, pois é $G$-simétrico para cada grupo $G$ de ordem $|G| = m$. Para verificar isto é suficiente observar que cada grupo $G$ atua atua sobre ele mesmo de maneira simplesmente transitiva. Outra família de canais que desfrutam este alto nível de simetria são os canais aditivos Gaussianos com constelação ortogonal, de energia equivalente, de $m$ sinais como entrada $X$. A verificação da simetria pode ser vista [9] Notemos que sempre que $m = p^r$, para algum primo $p$ e algum $r \in \mathbb{N}$ o grupo $G$ pode ser escolhido como sendo o produto direto de $r$ copias do grupo cíclico $\mathbb{Z}_p$, isto é, $(\mathbb{Z}_p)^r = \mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \cdots \oplus \mathbb{Z}_p$ que é o grupo aditivo do corpo de Galois $GF(p^r)$ □

## V. Canal $\mathbb{Z}_{p^r}$-simétrico

Um grupo abeliano finito $G$ pode ser decomposto na forma $G = (\mathbb{Z}_{p_1^{r_1}})^{s_1} \oplus (\mathbb{Z}_{p_2^{r_2}})^{s_2} \oplus \cdots \oplus (\mathbb{Z}_{p_k^{r_k}})^{s_3}$, onde os $p_i$ são primos dois a dois diferentes, os $r_i$ e $s_i$ são números naturais maiores do que zero e cada $\mathbb{Z}_{p_i^{r_i}}$ é um grupo cíclico de ordem $p_i^{r_i}$, [10], [11]. Se $G$ é cíclico então $G = \mathbb{Z}_{p_1^{r_1}} \oplus \mathbb{Z}_{p_2^{r_2}} \oplus \cdots \oplus \mathbb{Z}_{p_k^{r_k}}$ e o elemento $g = g_1 g_2 \ldots g_k$, onde $g_i$ é gerador de $\mathbb{Z}_{p_i^{r_i}}$ é tal que se $N = p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k}$ então $g^n \neq 0$ se $n \in \{1, 2, \ldots N - 1\}$ e $g^N = 0$. Analisaremos a capacidade de canal para o caso especial $\mathbb{Z}_{p^r}$ onde $p$ é primo e $r$ é um inteiro positivo.

*Lema 1:* Se $\phi : \mathcal{U} \to (\mathbb{Z}_{p^r})^N$ é um homomorfismo injetor de grupos, então $\mathcal{U} = (\mathbb{Z}_p)^{k_1} \oplus (\mathbb{Z}_{p^2})^{k_2} \oplus \cdots \oplus (\mathbb{Z}_{p^r})^{k_r}$, onde $k_1, k_2, \ldots, k_n$ são inteiros tais que $\sum_{j=1}^{r} j k_j \leq rN$.

*Demonstração:* Como $\phi$ é homomorfismo injetor, $\mathcal{U}$ é isomorfo com algum subgrupo de $(\mathbb{Z}_{p^r})^N$. Porem, todos os subgrupos de $\mathbb{Z}_{p^r}$ são da forma $\{0\}$, $\mathbb{Z}_p$, $\mathbb{Z}_{p^2}$, ..., $\mathbb{Z}_{p^{r-1}}$, $\mathbb{Z}_{p^r}$. Portanto todos os subgrupos de $(\mathbb{Z}_{p^r})^N$ devem ser da forma $(\mathbb{Z}_p)^{k_1} \oplus (\mathbb{Z}_{p^2})^{k_2} \oplus \cdots \oplus (\mathbb{Z}_{p^r})^{k_r}$ sujeitos à cardinalidade $|\mathcal{U}| = p^{k_1} p^{2k_2} \ldots p^{rk_r} = p^{\sum_{j=1}^{r} j k_j} \leq p^{rN}$, ou seja, $\sum_{j=1}^{r} j k_j \leq rN$.

■

*Lema 2:* Considere a família de sub-grupos $\mathcal{U}_l$ dado por $\mathcal{U}_l = (\mathbb{Z}_p)^{k_1} \oplus (\mathbb{Z}_{p^2})^{k_2} \oplus \cdots \oplus (\mathbb{Z}_{p^l})^{k_l} \oplus p(\mathbb{Z}_{p^{(l+1)}})^{k_{(l+1)}} \oplus \cdots \oplus p^{(r-l)}(\mathbb{Z}_{p^r})^{k_r}$ então $\phi(\mathcal{U}_l) \subset (p^{r-l}\mathbb{Z}_{p^r})^N$

*Demonstração:* Note que $\mathcal{U}_r = \mathcal{U}$ e $\mathcal{U}_j \subset \mathcal{U}_i$ se $j \leq i$. Os componentes da forma $p^j \mathbb{Z}_{p^{l+j}}$, para $j = 1, 2, \ldots, r - l$ são isomorfos com $\mathbb{Z}_{p^l}$, logo $(p^j \mathbb{Z}_{p^{l+j}})^{k_{l+j}} \cong (\mathbb{Z}_{p^l})^{k_{l+j}} \subset (\mathbb{Z}_{p^l})^N$ e $\mathcal{U}_l \cong (\mathbb{Z}_p)^{k_1} \oplus (\mathbb{Z}_{p^2})^{k_2} \oplus \cdots \oplus (\mathbb{Z}_{p^l})^{k_l} \oplus (\mathbb{Z}_{p^l})^{k_{l+1}} \oplus \cdots \oplus (\mathbb{Z}_{p^l})^{k_r}$. Portanto $\phi(\mathcal{U}_l)$ é isomorfo a um grupo contido em $(\mathbb{Z}_{p^l})^N$ e $\phi(\mathcal{U}_l) \subset (p^{r-l}\mathbb{Z}_{p^r})^N$. Finalmente, $|\mathcal{U}_l| = p^{k_1} p^{2k_2} \ldots p^{lk_l} p^{lk_{l+1}} \ldots p^{lkr} = p^{\sum_{j=1}^{l} j k_j} p^{l \sum_{j=1}^{r-l} k_{l+j}} = p^{\sum_{j=1}^{l} j k_j + l \sum_{j=1}^{r-l} k_{l+j}} = p^{\sum_{j=1}^{l} j k_j + l \sum_{j=l+j}^{r} k_j}$.

■

*Teorema 1:* Num canal $\mathbb{Z}_{p^r}$-simétrico $(X, Y, P)$ a capacidade de canal de Shannon é atingida se o mapeamento codificador é um homomorfismo $\phi : \mathcal{U} = (\mathbb{Z}_{p^r})^K \to (\mathbb{Z}_{p^r})^N$

*Demonstração:* Como o mapeamento codificador $\phi$ é injetor, pelo Lema 1, o grupo $\mathcal{U}$ deve ser isomorfo com $(\mathbb{Z}_p)^{k_1} \oplus (\mathbb{Z}_{p^2})^{k_2} \oplus \cdots \oplus (\mathbb{Z}_{p^r})^{k_r}$, com $\sum_{j=1}^{r} j k_j \leq rN$. Dai, a cardinalidade $|\mathcal{U}| = p^{\sum_{j=1}^{r} j k_j}$ e a taxa de codificação

$R = \frac{\log(|\mathcal{U}|)}{N}$ é dada por;

$$R = \frac{\log(p)}{N} \sum_{j=1}^{r} j k_j \tag{11}$$

Considere os subgrupos $G_l = p^{r-l}\mathbb{Z}_{p^r}$, $l \in \{1, 2, \ldots, r\}$. Note que $G_r = \mathbb{Z}_{p^r}$. Se $l = r-1$ temos $G_{r-1} = p\mathbb{Z}_{p^r} \cong \mathbb{Z}_{p^{(r-1)}}$. Em geral $G_l \cong \mathbb{Z}_{p^l}$. Considere o $l$-ésimo sub-canal $(p^{r-l}\mathbb{Z}_{p^r}, Y, P)$ e denote por $C_l$ sua capacidade.

Seja $R_l$ a taxa de codificação do sub-canal $(p^{r-l}\mathbb{Z}_{p^r}, Y, P)$. Então; $R_l = \frac{\log(|\mathcal{U}|)}{N}$. Pelo Lema 2, $|\mathcal{U}_l| = p^{\sum_{j=1}^{l} j k_j + l \sum_{j=l+j}^{r} k_j}$, então: $R_l = \frac{\log(p)}{N}(\sum_{j=1}^{l} j k_j + l \sum_{j=l+1}^{r} k_j)$. Como $l \leq r$ então $1 \geq \frac{l}{r}$ e $1 \geq \frac{j}{r}$ para todo $j = l+1, l+2, \ldots, r$, tem-se $\sum_{j=1}^{l} j k_j \geq (\frac{l}{r}) \sum_{j=1}^{l} j k_j$ e $l \sum_{j=l+1}^{r} k_j \geq \frac{l}{r} \sum_{j=l+1}^{r} j k_j$. Logo

$$R_l \geq \left(\frac{l}{r}\right) \frac{\log(p)}{N} \sum_{j=1}^{r} j k_j \tag{12}$$

Combinando as equações (11) e (12) obtemos $R \leq \frac{r}{l} R_l$. Por outro lado, a taxa de codificação do sub-canal $R_l$ deve ser menor ou igual do que a capacidade do sub-canal $C_l$. Portanto $R \leq \min_{l=1,2,\ldots,r}\{\frac{r}{l} C_l\}$. Sendo que $R$ atinge este mínimo quando $\mathcal{U} = (\mathbb{Z}_{p^r})^K$ com $K = \frac{RN}{r \log(p)}$.

■

## VI. Conclusões

Neste trabalho foi discutido que o valor da capacidade de canal $C$ pode diminuir para alguns esquemas de codificação. Na maioria dos esquemas de codificação linear conhecidos mantem o valor de $C$. Nestes casos é dito que a codificação atinge a capacidade de canal. Uma observação que foi feita nesta discussão é que quaisquer que seja o esquema de codificação o valor $C$ jamais pode ser superado, isto decorre do teorema de processamento de dados. Neste trabalho também foram dadas adaptações generalizadoras dos princípios de probabilidade utilizando a medida de Lebesgue. A consequência disto é que entropia, informação mútua, capacidade de canal e outros conceitos da teoria da informação também são generalizados. Finalmente, mostramos que para o caso particular de codificação sobre grupos cíclicos da forma $\mathbb{Z}_{p^r}$ a capacidade de canal é atingida. Na continuidade desta linha de pesquisa, sobre capacidade de canal, as próximas estruturas algébricas a serem consideradas deverão ser grupos cíclicos em geral, grupos abelianos, e grupos não abelianos.

## Referências


[1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 28, pp. 379–423, 1948.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley and Sons, 1968.

[3] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Piscataway NJ: Wiley InterScience, 2006.

[4] D. J. C. Mackay, *Information Theory, Inference, and Learning Algorithms*. United Kingdom: Cambridge University Press, 2005.

[5] G. Como and F. Fagnani, "The capacity of abelian group codes over symmetric channels," *IEEE Trans. Inform. Theory*, vol. IT 45, no. 01, pp. 3–31, 2009.

[6] S. Haykin, *Comunication Systems*, 4th ed. Wiley and Sons, 2001.

[7] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill, 1986.

[8] F. Morgan, *Geometric Measure Theory*, 4th ed. San Diego, CA USA: Elsevier Inc., 2009.

[9] R. Gallager, "Low density parity check codes," Ph.D. dissertation, Massachussets Institute of Technology, MIT, Cambrige MA, 1963.

[10] J. J. Rotman, *An Introduction to the Theory of the Groups*, 4th ed. New York: Springer Verlag, 1995.

[11] M. Hall, *The Theory of Groups*. New York: Mac Millan, 1959.