DOI: 10.22481/10.22481/intermaths.v2i1.8412
Licença Creative Commons
Artigo | Article

INTERMATHS
ISSN 2675-8318
Vol. 2, N. 1, Jan - Jun 2021, p. 35 – 52

Extensão de Galois $GF(2^6)$ Aplicada na Modelagem do Código Genético

Roberta Siqueira Fernandes

Universidade Federal de Alfenas - MG

robertaf_mat@hotmail.com

Anderson José de Oliveira

Universidade Federal de Alfenas - MG

anderson.oliveira@unifal-mg.edu.br

Galois Extension $GF(2^6)$ Applied in the Genetic Code Modeling

Abstrac

For the scientific community, one of the biggest challenges is to analyze the existence of a mathematical structure related to DNA. Many researches have been carried out involving the genetic code, among other biological phenomena, using Mathematics to contribute to the analysis and description of these theoretical concepts. The codons are formed by a crack of nitrogenous bases, with 64 possible combinations. The nitrogenous bases are adenine, cytosine, guanine and thymine/uracil, which are represented by the letters A, C, G and T/U, respectively, and represent the DNA alphabet and through the bijection of this alphabet with the ring $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ it is possible to obtain 24 permutations, which can be divided into 3 labels (A, B and C). The aim of this work is to present the use of algebra elements in the genetic code modeling. A polynomial, vector and power representation for the genetic code structure will be presented, where an element of the $GF(2^6)$ extension was associated for each codon, since we see that there is an one-to-one association of the codons of the genetic code with an element of the Galois extension. The algebraic characterization for the labels A, B and C of the genetic code was obtained through these representations mentioned above, which can be used in future studies involving the analysis of the genetic code.

Key words: Algebra; DNA; Polynomials.

Resumo

Para a comunidade científica, um dos maiores desafios é analisar a existência de uma estrutura matemática relacionada com o DNA. Muitas pesquisas vêm sendo realizadas envolvendo o código genético, entre outros fenômenos biológicos, utilizando a Matemática para contribuir na análise e descrição desses conceitos teóricos. Os códons são formados por uma trinca de bases nitrogenadas, com 64 combinações possíveis. As bases nitrogenadas são a adenina, citosina, guanina e timina/uracila, que são representadas pelas letras A, C, G e T/U, respectivamente, e representam o alfabeto do DNA e por meio da bijeção desse alfabeto com anel $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ é possível obter 24 permutações, que podem ser divididas em 3 rotulamentos (A, B e C). O objetivo deste trabalho é apresentar a utilização de elementos de álgebra na modelagem do código genético. Serão apresentadas uma representação polinomial, vetorial e por potência para a estrutura do código genético, onde para cada códon foi associado um elemento da extensão $GF(2^6)$, uma vez que vemos que existe uma associação um-a-um dos códons do código genético com um elemento da extensão de Galois. Foi obtida a caracterização algébrica para os rotulamentos A, B e C do código genético, por meio dessas representações mencionadas anteriormente, as quais poderão ser utilizadas em futuros estudos envolvendo a análise do código genético.

Palavras-chave: Álgebra; DNA; Polinômios.

Submetido em: 08 de abril de 2021 – **Aceito em**: 25 de junho de 2021

1 INTRODUÇÃO

O sistema biológico preocupa-se com o armazenamento e transmissão de informação por meio do código genético, e segundo [1], essa forma de comunicação é realizada pelo indivíduo de forma natural. Alguns autores, como [2], [3] e [4], comparam esse processo com a teoria da comunicação, que assim como a genética, cuida do envio de informação, porém sua realização é de forma programada.

Segundo [1], um código corretor de erros é uma forma organizada de adicionar uma informação que se queira armazenar ou transmitir e assim pode-se recuperar, detectar e corrigir erros de determinada informação. Os códigos corretores de erros podem ser classificados em códigos lineares e códigos não-lineares. Como exemplo de classes de códigos não-lineares temos os códigos de Nordstrom-Robinson e Preparata. Já nas classes de códigos lineares temos os códigos cíclicos, os códigos de Hamming, os códigos Reed-Solomon e os códigos BCH (Bose-Chaudhuri-Hocquenghem).

O alfabeto do código genético é dado pelas bases nitrogenadas: adenina (A), citosina (C), guanina (G) e timina (T) ou uracila (U), denotado assim pelo conjunto $N = \{A, C, G, T/U\}$ e pode ser associado ao alfabeto 4-ário na estrutura de anel, denotado por $\mathbb{Z}_4 = \{0, 1, 2, 3\}$.

Segundo [1] e [2], a partir das associações dos elementos do conjunto $N = \{A, C, G, T/U\}$ com os elementos do conjunto $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, obtém-se 24 permutações, apresentadas na Figura 1, que tem por objetivo estabelecer a associação de cada um dos elementos do conjunto N com o elemento correspondente do conjunto $\mathbb{Z}_4$.

Figura 1: Mapeamento para $\mathbb{Z}_4 = \{0, 1, 2, 3\}$.

$$
\begin{bmatrix} A & C & G & T \\ 0 & 1 & 3 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 0 & 3 & 1 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 0 & 2 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 2 & 0 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 1 & 3 & 0 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 3 & 1 & 0 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 0 & 2 & 1 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 2 & 0 & 1 \end{bmatrix}
$$
$$
\begin{bmatrix} A & C & G & T \\ 0 & 1 & 2 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 0 & 3 & 2 & 1 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 0 & 3 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 2 & 3 & 0 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 1 & 0 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 3 & 0 & 1 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 0 & 1 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 2 & 1 & 0 \end{bmatrix}
$$
$$
\begin{bmatrix} A & C & G & T \\ 0 & 2 & 1 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 0 & 2 & 3 & 1 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 3 & 0 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 1 & 3 & 2 & 0 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 0 & 1 & 3 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 2 & 0 & 3 & 1 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 1 & 0 & 2 \end{bmatrix}
\begin{bmatrix} A & C & G & T \\ 3 & 1 & 2 & 0 \end{bmatrix}
$$

Fonte: Elaborado pelos autores.

Sabendo que é desconhecido o mapeamento entre $N \leftrightarrow \mathbb{Z}_4$, uma sequência de DNA pode ser rotulada conforme as 24 permutações entre $N \leftrightarrow \mathbb{Z}_4$. Pode-se organizar esse mapeamento em três conjuntos, denominados rotulamentos *A*, *B* e *C*, contendo 8 permutações cada, de acordo com a caracterização geométrica de cada um, como mostrado na Figura 2, conforme [1] e [2].

Figura 2: Rotulamentos *A*, *B* e *C*.

Rotulamento A

$$\begin{bmatrix} A & C & G & T \\ 0 & 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 1 & 3 & 0 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 0 & 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 3 & 1 & 0 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 0 & 2 & 1 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 2 & 0 & 1 \end{bmatrix}$$

Rotulamento B

$$\begin{bmatrix} A & C & G & T \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 1 & 0 & 3 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 0 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 3 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 0 & 1 & 2 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

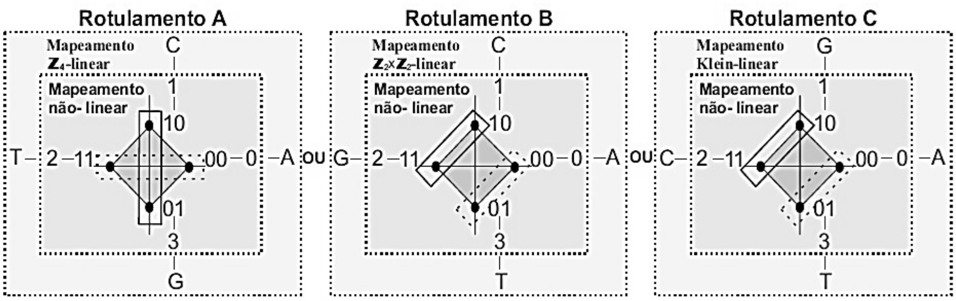
Rotulamento C

$$\begin{bmatrix} A & C & G & T \\ 0 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 0 & 1 & 3 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 0 & 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 2 & 0 & 3 & 1 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 3 & 0 & 2 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 1 & 0 & 2 \end{bmatrix}$$
$$\begin{bmatrix} A & C & G & T \\ 1 & 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} A & C & G & T \\ 3 & 1 & 2 & 0 \end{bmatrix}$$

Fonte: Elaborado pelos autores.

De acordo com [1] e [2], a diferença de cada rotulamento é a associação de complementaridade dos nucleotídeos, como mencionado anteriormente. Temos que *A - T/U* e *C - G*. O rotulamento *A* classifica as sequências como não-lineares ($\mathbb{Z}_4-$ linear), pois, qualquer um dos nucleotídeos para alcançar seu complementar necessita caminhar duas arestas. Já os rotulamentos *B* e *C* classificam suas sequências como lineares ($\mathbb{Z}_2 \times \mathbb{Z}_2$ e Klein-linear) respectivamente, pois, qualquer um dos nucleotídeos para alcançar seu complementar necessita caminhar uma aresta, como apresenta a Figura 3.

Figura 3: Caracterização Geométrica dos Rotulamentos *A*, *B* e *C*.



Fonte: [1, p. 95]

Em [5] são apresentadas as construções dos reticulados booleanos e dos diagramas de Hasse, associados às permutações 0132 e 2310, onde foi possível obter uma caracterização biológica através das constuções, além de uma análise das diferenças e semelhanças físico-químicas dos aminoácidos, por meio dos cálculos das médias das distâncias de Hamming entre os códons, concluindo que os valores obtidos dos cálculos refetem nos diagramas.

Em [6] é apresentado um modelo no qual são considerados os 64 códons do código genético. Para a construção da tabela do código genético, é feita uma bijeção do alfabeto biológico com a representação $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{00, 01, 10, 11\}$. Nesse estudo, é proposto o mapeamento G - 00, U - 10, A - 01 e C - 11, onde são definidas operações no conjunto das quatro bases do DNA.

Com base nos estudos mencionados anteriormente, pressupõe-se que exista conexão entre os elementos da extensão de Galois $GF(2^6)$ com os códons do código genético. Esta afirmação pode ser comprovada mediante a resposta da seguinte questão: "Como os elementos da extensão de Galois $GF(2^6)$ podem ser aplicados na modelagem do código genético?"

O objetivo deste trabalho é apresentar a relação entre os elementos da extensão de Galois $GF(2^6)$ e os códons do código genético para os rotulamentos *A*, *B* e *C*, a qual poderá ser utilizada no estudo de fenômenos mutacionais em trabalhos futuros. Por meio dessa relação pode-se estabelecer conexões entre Matemática (extensões de Galois), Biologia (código genético) e Engenharia (aplicações em códigos BCH no processo de transmissão da informação).

2 ELEMENTOS DE BIOLOGIA E ÁLGEBRA

Os conceitos apresentados a seguir podem ser encontrados de maneira detalhada em [7], [8], [9], [10], [11] e [12].

2.1 Biologia

A célula é a menor unidade morfofisiológica de um ser vivo. Ela foi descoberta por Robert Hooke (1635-1703), no ano de 1665 ao examinar um pedaço de cortiça em seu microscópio. Hooke percebeu que o pedaço de cortiça era composto por diversas concavidades em forma de poliedro.

As células podem ser apresentadas sob uma variedade de formas e tamanhos. Há aquelas que se movem rapidamente e outras que têm estruturas que se alteram. Além disso, para algumas células o oxigênio é essencial, já para outras ele pode até matar. Existem organismos vivos que possuem uma única célula (unicelulares) e aqueles que são formados por mais de uma célula (pluricelulares).

Os nucleotídeos são unidades de moléculas de um ácido nucleico e que estão presentes no metabolismo como transportador de energia.

Os ácidos nucleicos se encontram em todas as células vivas e têm a função de conter, transmitir e traduzir informações genéticas dos seres vivos. Todos os organismos vivos apresentam ácidos nucleicos na forma de ácido desoxirribonucleico (DNA) e ácido ribonucleico (RNA).

O DNA é um polímero, ou seja, uma longa cadeia de nucleotídeos. Sua função é armazenar informação genética dos seres vivos. O DNA é formado por uma dupla hélice com giro para a direita e parece uma escada de caracol molecular. Cada nucleotídeo do DNA é constituído por um grupo fosfato, uma pentose (desoxirribose) e uma dentre quatro bases nitrogenadas: adenina (A), guanina (G), citosina (C) e timina (T), que estão interligadas por pontes de hidrogênio.

De acordo com a **Regra de Chargaff**, a quantidade total de nucleotídeos de timina + citosina é sempre igual à quantidade total de nucleotídeos de adenina + guanina. A quantidade de timina é sempre igual à quantidade de adenina e a quantidade de citosina é sempre a mesma que de guanina. Porém, a quantidade de adenina + timina não é necessariamente igual à quantidade de guanina + citosina, ou seja, essa proporção pode variar de acordo com cada organismo.

O RNA é um polímero de fita simples e não é formado por uma dupla hélice como o DNA. Possui uma pentose (ribose) em seus nucleotídeos em vez da desoxirribose encontrada no DNA. Logo, a diferença nesses açúcares é a presença ou a ausência de apenas um átomo de oxigênio. O RNA possui a uracila (U) no lugar da timina (T) em sua composição e participa na tradução do DNA em proteína.

Para que ocorra o processo de duplicação do DNA, é necessário o rompimento das pontes de hidrogênio, para que assim aconteça o distanciamento das duas fitas. Com os nucleotídeos presentes nas células e livres, começam a ser atraídos pelas bases do DNA na parte em que as fitas estão separadas, formando assim outro filamento. A formação desse filamento obedece à **Regra de Chargaff**, onde, adenina emparelha com timina e citosina emparelha com guanina. Desta forma, obtém-se duas moléculas de DNA idênticas entre si.

O DNA e o RNA são os responsáveis pela síntese proteica. Para que esse processo ocorra, é necessário que a célula transcreva e traduza a informação do código genético.

Existem três tipos de RNA: o RNAm (RNA mensageiro) que leva a mensagem do DNA, que está no núcleo, ao citoplasma; RNAt (RNA transportador) que é responsável por identificar e transportar os aminoácidos; e o RNAr (RNA ribossômico) que ajuda na decodificação da informação genética.

Cada molécula de DNA se diferencia pela sequência de bases nitrogenadas que elas apresentam, e essa sequência é que vai formar o código genético, ou seja, essa sequência de bases do DNA se relaciona com a sequência correspondente de 20 aminoácidos para a formação das proteínas. Então, essa correspondência permite saber o aminoácido específico de cada códon. Esses códons são formados por uma trinca de bases nitrogenadas com 64 combinações possíveis. As bases são formadas pela adenina, citosina, guanina e timina/uracila, que são representadas pelas letras A, C, G e T/U respectivamente, e representam o alfabeto do DNA. A Tabela 1 apresenta as 64 combinações possíveis compostas pelas quatro bases nitrogenadas.

Tabela 1: Combinações possíveis compostas pelas quatro bases nitrogenadas.

primeira posição	segunda posição				terceira posição
	A	C	G	U	
A	AAA	ACA	AGA	AUA	A
	AAC	ACC	AGC	AUC	C
	AAG	ACG	AGG	AUG	G
	AAU	ACU	AGU	AUU	U
C	CAA	CCA	CGA	CUA	A
	CAC	CCC	CGC	CUC	C
	CAG	CCG	CGG	CUG	G
	CAU	CCU	CGU	CUU	U
G	GAA	GCA	GGA	GUA	A
	GAC	GCC	GGC	GUC	C
	GAG	GCG	GGG	GUG	G
	GAU	GCU	GGU	GUU	U
U	UAA	UCA	UGA	UUA	A
	UAC	UCC	UGC	UUC	C
	UAG	UCG	UGG	UUG	G
	UAU	UCU	UGU	UUU	U

Fonte: Elaborado pelos autores.

2.2 Álgebra

Definição 2.1. *Um grupo é uma estrutura algébrica constituída por um conjunto G e uma operação * que satisfaz as seguintes propriedades: Associativa: $(a*b)*c = a*(b*c)$, $\forall\ a,b,c \in G$; Existência de Elemento Neutro: $\exists$ um elemento $e \in G$ tal que $a*e = e*a = a$, $\forall\ a \in G$; Existência de Simétricos: $\forall\ a \in G$, $\exists$ um elemento $a' \in G$ tal que $a*a' = a'*a = e$. Se o grupo obedecer a propriedade da comutatividade, ou seja, se $a*b = b*a$, $\forall\ a,b \in G$, então esse grupo recebe o nome de grupo comutativo ou abeliano.*

Definição 2.2. *Um anel é uma estrutura algébrica constituída por um conjunto A e duas operações $+$ e $\cdot$ que satisfaz as seguintes propriedades:*

Para a soma: $(A,+)$ é um grupo abeliano; Associativa: $(a+b)+c = a+(b+c)$, $\forall\ a,b,c \in A$; Comutativa: $a+b = b+a$, $\forall\ a,b \in A$; Existência de Elemento Neutro: $\exists$ um elemento $0_A \in A$ tal que $a+0_A = 0_A+a = a, \forall\ a \in A$; Existência de Elemento Oposto: $\forall\ a \in A, \exists -a \in A$, tal que $a+(-a) = (-a)+a = 0_A$;

Para a multiplicação: Associativa: $a\cdot(b\cdot c) = (a\cdot b)\cdot c$, $\forall\ a,b,c \in A$; Distributiva (em relação à adição): $a\cdot(b+c) = a\cdot b + a\cdot c$ e $(b+c)\cdot a = b\cdot a + c\cdot a$; $\forall\ a,b,c \in A$.

Definição 2.3. *Um anel $(A,+,\cdot)$, onde a operação $\cdot$ é comutativa é denominado de* *anel comutativo.*

Definição 2.4. *Anel das classes de restos módulo m, $\forall\ m > 1$, é o conjunto $\mathbb{Z}_m = \{\overline{0}, \overline{1}, \overline{2}, ..., \overline{m-1}\}$, em relação às operações assim definidas:*

$$\overline{a} + \overline{b} = \overline{a+b},$$
$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

O zero desse anel é classe $\overline{0}$ e oposto de um elemento $\overline{a} \in \mathbb{Z}_m$ é classe $\overline{m-a}$.

Exemplo 2.5. Os anéis da classe de resto módulo m, $\mathbb{Z}_m$ ($m > 1$), são exemplos importantes de anéis finitos.

Considere as tábuas do anel $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, apresentadas nas Tabelas 2 e 3.

Tabela 2: Tábua da adição de $\mathbb{Z}_4$.

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Fonte: Elaborado pelos autores.

Tabela 3: Tábua da multiplicação de $\mathbb{Z}_4$.

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Fonte: Elaborado pelos autores.

Definição 2.6. *Um conjunto K será chamado de corpo se for munido de uma operação de adição $(+)$ e uma operação de multiplicação $(\cdot)$, ou seja, é grupo para adição e é grupo para multiplicação, verificando as seguintes propriedades:*

Para a adição: Associativa: $(a + b) + c = a + (b + c)$, $\forall\ a, b, c \in K$; Comutativa: $a+b = b+a$, $\forall\ a, b \in K$; Existência de Elemento Neutro: $\exists\ 0 \in K$, tal que $a+0 = 0+a = a$, $\forall\ a \in K$; Elemento Inverso: $\forall\ a \in K$, $\exists\ -a \in K$ tal que $a + (-a) = (-a) + a = 0$.

Para a multiplicação: Associativa: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, $\forall\ a, b, c \in K$; Comutativa: $a \cdot b = b \cdot a$, $\forall\ a, b \in K$; Existência de Elemento Neutro: $\exists\ 1 \in K - \{0\}$ tal que $a \cdot 1 = 1 \cdot a = a$, $\forall\ a \in K$; Elemento Inverso: $\forall\ a \in K - \{0\}$ existe $a_1 \in K$, tal que $a \cdot a_1 = a_1 \cdot a = 1$, $\forall\ a \in K$; A multiplicação é distributiva em relação à adição, isto é, $a \cdot (b + c) = a \cdot b + a \cdot c$ e $(b + c) \cdot a = b \cdot a + c \cdot a$, $\forall\ a, b, c \in K$.

Definição 2.7. *Um corpo de Galois possui um número finito de elementos e é representado por $GF(p)$, onde p é um número primo.*

Para qualquer inteiro m é possível estender um corpo primo $GF(p)$ com p elementos, para um corpo estendido $GF(p^m)$ com p^m elementos. Portanto, a ordem de qualquer corpo finito estendido é potência de um primo.

Definição 2.8. *Um polinômio $p(X)$ sobre $GF(2)$ de grau m é dito irredutível sobre $GF(2)$ se ele não for divisível por nenhum outro polinômio sobre $GF(2)$ de grau menor que m, mas maior que zero.*

Exemplo 2.9. Segundo [12], os polinômios: $X^2 + X + 1$; $X^3 + X + 1$; $X^4 + X + 1$ são irredutíveis.

Existem apenas dois polinômios de grau 1 sobre $GF(2)$, que são: X e $X + 1$.

Seja $f_1(X) = X^2 + X + 1$. O Teorema do Resto ([8, p. 297]), nos diz que o fato de $f_1(0) = 1 \neq 0$ e $f_1(1) = 3 = 1 \neq 0$, resulta que f_1 não pode ser divisível por X e nem por $X + 1$. Portanto, $f_1(X)$ é um polinômio irredutível de grau 2.

Seja $f_2(X) = X^3+X+1$. Com um raciocínio análogo, concluímos que $f_2(X)$ não é divisível por X e nem por $X+1$, pois $f_2(0) \neq 0$ e $f_2(1) \neq 0$. Como $f_2(X)$ não é divisível por nenhum polinômio de grau 1, ele não pode ser divisível por um polinômio de grau 2. Lembrando que a única maneira de fatorar $f_2(X)$, sem que um dos fatores seja um polinômio constante, e no caso em que um polinômio tem grau 1 e o outro tem grau 2, para que a soma dos graus resulte em 3 = grau f_2. Portanto, $f_2(X)$ é irredutível sobre $GF(2)$.

Seja $f_3(X) = X^4 + X + 1$. Como $f_3(0) \neq 0$ e $f_3(1) \neq 0$, concluímos que $f_3(X)$ não se fatora como um produto de um polinômio de grau 1 por um polinômio de grau 3. A possibilidade que resta é através do produto de dois polinômios de grau 2. Existem quatro polinômios de grau 2 sobre $GF(2)$, que são: X^2; $1 + X^2$; $X + X^2$ e $1 + X + X^2$. Uma vez que o produto entre eles nunca resulta em $f_3(X)$, concluímos que $f_3(X)$ é um polinômio irredutível de grau 4.

Definição 2.10. *Um polinômio irredutível $p(X)$ de grau m é dito primitivo se o menor inteiro positivo n para o qual $p(X)$ divide $X^n + 1$ é $n = 2^m - 1$.*

2.2.1 Construção dos Corpos de Galois $GF(2^m)$

Nesta seção será apresentado um método para construir corpos de Galois de 2^m elementos $(m > 1)$ do corpo binário $GF(2)$. Começamos com os dois elementos 0 e 1 de $GF(2)$ e um novo símbolo α. Então definimos uma multiplicação "$\cdot$" para introduzir uma sequência de potências de α como segue:

$$
\begin{aligned}
0 \cdot 0 &= 0,\\
0 \cdot 1 &= 0 = 1 \cdot 0,\\
1 \cdot 1 &= 1,\\
0 \cdot \alpha &= 0 = \alpha \cdot 0,\\
1 \cdot \alpha &= 1 = \alpha \cdot 1,\\
\alpha^2 &= \alpha \cdot \alpha,
\end{aligned}
$$

$$\begin{aligned}
&\alpha^3 = \alpha \cdot \alpha \cdot \alpha, \\
&\ \vdots \\
&\alpha^j = \alpha \cdot \alpha \cdot \ldots \cdot \alpha \quad (j - \text{vezes}) \\
&\ \vdots
\end{aligned} \tag{1}$$

Segue da definição de multiplicação precedente que:

$$\begin{aligned}
0 \cdot \alpha^j &= \alpha^j \cdot 0 = 0, \\
1 \cdot \alpha^j &= \alpha^j \cdot 1 = 1, \\
\alpha^i \cdot \alpha^j &= \alpha^j \cdot \alpha^i = \alpha^{i+j}.
\end{aligned}$$

Agora, temos que o seguinte conjunto de elementos sobre os quais a operação multiplicação "$\cdot$" está definida: $F = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^j, \ldots\}$. O elemento 1 é algumas vezes denotado por α^0.

Suponha que o conjunto F contenha apenas 2^m elementos e seja fechado sobre a multiplicação "$\cdot$" definida em (1).

Seja $p(X)$ um polinômio primitivo de grau m sobre $GF(2)$. Assumimos que $p(\alpha) = 0$ (isto é, α é uma raíz de $p(X)$). Como $p(X)$ divide $X^{2^m-1} + 1$, temos: $X^{2^m-1} + 1 = q(X)p(X)$. Se trocarmos X por α, obtemos: $\alpha^{2^m-1} + 1 = q(\alpha)p(\alpha)$. Para $p(\alpha) = 0$, obtem-se $\alpha^{2^m-1} + 1 = q(\alpha) \cdot 0$.

Considere $q(\alpha)$ como um polinômio de α sobre $GF(2)$, segue que $q(\alpha) \cdot 0 = 0$. Como resultado, obtemos: $\alpha^{2^m-1} + 1 = 0$.

Adicionando 1 a ambos os lados de $\alpha^{2^m-1} + 1 = 0$ (usando adição módulo-2), obtemos:

$$\alpha^{2^m-1} = 1 = \alpha^0. \tag{2}$$

Assim, sob a condição que $p(\alpha) = 0$, o conjunto F se torna finito e contendo os seguintes elementos:

$$F^* = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{2^m-2}\}. \tag{3}$$

F^* possui ordem 2^m e os seus elementos não nulos são fechados sob a operação multiplicação "$\cdot$".

Para construir $GF(2^m)$ de $GF(2)$, serão desenvolvidas duas representações para os elementos não nulos de $GF(2^m)$: a representação por potência e a representação polinomial. A representação por potência é conveniente para a multiplicação e a representação polinomial é conveniente para a adição.

Exemplo 2.11. Seja $m = 4$ e considere o polinômio primitivo $p(X) = 1 + X + X^4$ sobre $GF(2)$. Admitindo que α seja uma raíz do polinômio, então $p(\alpha) = 0$, ou seja,

$$0 = 1 + \alpha + \alpha^4 \qquad \Rightarrow \qquad \alpha^4 = 1 + \alpha. \tag{4}$$

Usando a relação (4), pode-se construir um $GF(2^4)$ como se segue:

$$\begin{aligned}
\alpha^5 &= \alpha \cdot \alpha^4 = \alpha(1+\alpha) = \alpha + \alpha^2,\\
\alpha^6 &= \alpha \cdot \alpha^5 = \alpha(\alpha+\alpha^2) = \alpha^2 + \alpha^3,\\
\alpha^7 &= \alpha \cdot \alpha^6 = \alpha(\alpha^2+\alpha^3) = \alpha^3 + \alpha^4 = \alpha^3 + 1 + \alpha = 1 + \alpha + \alpha^3,\\
\alpha^8 &= \alpha \cdot \alpha^7 = \alpha(1+\alpha+\alpha^3) = \alpha + \alpha^2 + \alpha^4 = \alpha + \alpha^2 + 1 + \alpha = 1 + \alpha^2,\\
\alpha^9 &= \alpha \cdot \alpha^8 = \alpha(1+\alpha^2) = \alpha + \alpha^3,\\
\alpha^{10} &= \alpha \cdot \alpha^9 = \alpha(\alpha+\alpha^3) = \alpha^2 + \alpha^4 = \alpha^2 + 1 + \alpha = 1 + \alpha + \alpha^2,\\
\alpha^{11} &= \alpha \cdot \alpha^{10} = \alpha(1+\alpha+\alpha^2) = \alpha + \alpha^2 + \alpha^3,\\
\alpha^{12} &= \alpha \cdot \alpha^{11} = \alpha(\alpha+\alpha^2+\alpha^3) = \alpha^2 + \alpha^3 + 1 + \alpha = 1 + \alpha + \alpha^2 + \alpha^3,\\
\alpha^{13} &= \alpha \cdot \alpha^{12} = \alpha(1+\alpha+\alpha^2+\alpha^3) = \alpha + \alpha^2 + \alpha^3 + 1 + \alpha = 1 + \alpha^2 + \alpha^3,\\
\alpha^{14} &= \alpha \cdot \alpha^{13} = \alpha(1+\alpha^2+\alpha^3) = \alpha + \alpha^3 + \alpha^4 = \alpha + \alpha^3 + 1 + \alpha = 1 + \alpha^3,\\
\alpha^{15} &= \alpha \cdot \alpha^{14} = \alpha(1+\alpha^3) = \alpha + \alpha^4 = \alpha + 1 + \alpha = \underbrace{1}_{voltou!}.
\end{aligned}$$

As representações por potência, polinomial e vetorial para $GF(2^4)$ são apresentadas na Tabela 4.

Tabela 4: $GF(2^4)$ gerado por $p(X) = 1 + X + X^4$.

POTÊNCIA	POLINOMIAL	VETORIAL
0	0	(0000)
$\alpha^0 = 1$	1	(1000)
α^1	α	(0100)
α^2	α^2	(0010)
α^3	α^3	(0001)
α^4	$1+\alpha$	(1100)
α^5	$\alpha+\alpha^2$	(0110)
α^6	$\alpha^2+\alpha^3$	(0011)
α^7	$1+\alpha+\alpha^3$	(1101)
α^8	$1+\alpha^2$	(1010)
α^9	$\alpha+\alpha^3$	(0101)
α^{10}	$1+\alpha+\alpha^2$	(1110)
α^{11}	$\alpha+\alpha^2+\alpha^3$	(0111)
α^{12}	$1+\alpha+\alpha^2+\alpha^3$	(1111)
α^{13}	$1+\alpha^2+\alpha^3$	(1011)
α^{14}	$1+\alpha^3$	(1001)

Fonte: Elaborado pelos autores.

Para cada extensão de Galois existe um número de polinômios primitivos. Como os códons do código genético são formados por uma trinca de bases nitrogenadas, pode-se obter 64 combinações distintas, então será utilizada uma extensão de Galois de grau $r = 6$, a qual possui seis polinômios primitivos, conhecidos na literatura, em [13]:

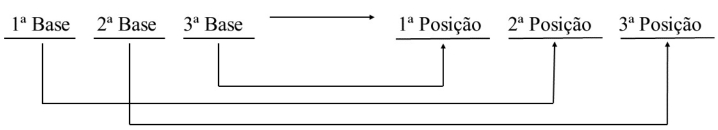
1. $x^6 + x^5 + x^3 + x^2 + 1$,
2. $x^6 + x + 1$,
3. $x^6 + x^5 + x^2 + x + 1$,
4. $x^6 + x^4 + x^3 + x + 1$,
5. $x^6 + x^5 + x^4 + x + 1$,
6. $x^6 + x^5 + 1$.

3 RESULTADOS E DISCUSSÕES

O estudo proposto visa apresentar uma representação polinomial e vetorial para a estrutura do código genético, onde para cada códon será associado um elemento da extensão $GF(2^6)$, uma vez que existe uma associação um-a-um dos códons do código genético com um elemento da extensão de Galois. Para isso, consideramos um polinômio primitivo para a obtenção da extensão e consequente associação com o meio biológico. Além disso, apresentaremos uma aplicação de uma abordagem proposta por [6], para a modelagem do código genético e a relação da mesma com a abordagem da extensão de Galois.

Para a representação polinomial dos códons e para a construção da tabela do código genético, [6], ao contrário de [2], onde o mapeamento é feito através da bijeção do alfabeto biológico $N = \{A, C, G, T/U\}$ com o anel $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, faz uma bijeção do alfabeto biológico com a representação $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{00, 01, 10, 11\}$. Um dos casos propostos por [6], é o mapeamento G - 00, U - 10, A - 01 e C - 11. Portanto, essa representação reflete a permutação 3201 do rotulamento A, de acordo com [13]. Deve-se ressaltar que a representação dos códons não é feita de forma arbitrária, e sim obedecendo a ordem de importância de cada base, onde a terceira base é menos importante e a segunda base é mais importante, como pode ser visto na Figura 4.

Figura 4: Ordem de importância das bases.



Fonte: Elaborado pelos autores.

Tomemos como exemplo, o códon GUC. Então sua representação na forma binária será 110010 (leitura da esquerda para direita), pois, a primeira posição corresponde a terceira base (base C), na segunda posição corresponde a primeira base (base G) e por último, a terceira posição corresponde a segunda base (base U). Assim, o grau aumenta da base menos importante, que é a terceira, para a base mais importante, que é a segunda.

Por fim, o polinômio que representa o códon GUC é $1+\alpha+\alpha^4$.

Neste estudo foi escolhida, de forma aleatória, uma permutação de cada um dos rotulamentos *A*, *B* e *C* e três dos seis polinômios primitivos mencionados na Seção 2.2, a fim de que pudéssemos explorar os conceitos algébricos, geométricos e biológicos, associados aos rotulamentos do código genético.

3.1 Detalhamento do Rotulamento A

Para o rotulamento *A*, foi utilizada a permutação 1203. Escolhemos o polinômio primitivo x^6+x+1 para a geração da extensão de Galois $GF(2^6)$, por meio da qual será feita a associação com os códons do código genético.

A Tabela 5, a seguir, contém as representações por potência, polinomial e vetorial.

Tabela 5: Representações por potência, polinomial e vetorial associados ao polinômio primitivo x^6+x+1.

-	Potência	Polinomial	Vetorial	-	Potência	Polinomial	Vetorial
00	0	0	(000000)	32	α^{31}	$1+\alpha^2+\alpha^5$	(101001)
01	α^0	1	(100000)	33	α^{32}	$1+\alpha^3$	(100100)
02	α^1	α^1	(010000)	34	α^{33}	$\alpha+\alpha^4$	(010010)
03	α^2	α^2	(001000)	35	α^{34}	$\alpha^2+\alpha^5$	(001001)
04	α^3	α^3	(000100)	36	α^{35}	$1+\alpha+\alpha^3$	(110100)
05	α^4	α^4	(000010)	37	α^{36}	$\alpha+\alpha^2+\alpha^4$	(011010)
06	α^5	α^5	(000001)	38	α^{37}	$\alpha^2+\alpha^3+\alpha^5$	(001101)
07	α^6	$1+\alpha$	(110000)	39	α^{38}	$1+\alpha+\alpha^3+\alpha^4$	(110110)
08	α^7	$\alpha+\alpha^2$	(011000)	40	α^{39}	$\alpha+\alpha^2+\alpha^4+\alpha^5$	(011011)
09	α^8	$\alpha^2+\alpha^3$	(001100)	41	α^{40}	$1+\alpha+\alpha^2+\alpha^3+\alpha^5$	(111101)
10	α^9	$\alpha^3+\alpha^4$	(000110)	42	α^{41}	$1+\alpha^2+\alpha^3+\alpha^4$	(101110)
11	α^{10}	$\alpha^4+\alpha^5$	(000011)	43	α^{42}	$\alpha+\alpha^3+\alpha^4+\alpha^5$	(010111)
12	α^{11}	$1+\alpha+\alpha^5$	(110001)	44	α^{43}	$1+\alpha+\alpha^2+\alpha^4+\alpha^5$	(111011)
13	α^{12}	$1+\alpha^2$	(101000)	45	α^{44}	$1+\alpha^2+\alpha^3+\alpha^5$	(101101)
14	α^{13}	$\alpha+\alpha^3$	(010100)	46	α^{45}	$1+\alpha^3+\alpha^4$	(100110)
15	α^{14}	$\alpha^2+\alpha^4$	(001010)	47	α^{46}	$\alpha+\alpha^4+\alpha^5$	(010011)
16	α^{15}	$\alpha^3+\alpha^5$	(000101)	48	α^{47}	$1+\alpha+\alpha^2+\alpha^5$	(111001)
17	α^{16}	$1+\alpha+\alpha^4$	(110010)	49	α^{48}	$1+\alpha^2+\alpha^3$	(101100)
18	α^{17}	$\alpha+\alpha^2+\alpha^5$	(011001)	50	α^{49}	$\alpha+\alpha^3+\alpha^4$	(010110)
19	α^{18}	$1+\alpha+\alpha^2+\alpha^3$	(111100)	51	α^{50}	$\alpha^2+\alpha^4+\alpha^5$	(001011)
20	α^{19}	$\alpha+\alpha^2+\alpha^3+\alpha^4$	(011110)	52	α^{51}	$1+\alpha+\alpha^3+\alpha^5$	(110101)
21	α^{20}	$\alpha^2+\alpha^3+\alpha^4+\alpha^5$	(001111)	53	α^{52}	$1+\alpha^2+\alpha^4$	(101010)
22	α^{21}	$1+\alpha+\alpha^3+\alpha^4+\alpha^5$	(110111)	54	α^{53}	$\alpha+\alpha^3+\alpha^5$	(010101)
23	α^{22}	$1+\alpha^2+\alpha^4+\alpha^5$	(101011)	55	α^{54}	$1+\alpha+\alpha^2+\alpha^4$	(111010)
24	α^{23}	$1+\alpha^3+\alpha^5$	(100101)	56	α^{55}	$\alpha+\alpha^2+\alpha^3+\alpha^5$	(011101)
25	α^{24}	$1+\alpha^4$	(100010)	57	α^{56}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4$	(111110)
26	α^{25}	$\alpha+\alpha^5$	(010001)	58	α^{57}	$\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$	(011111)
27	α^{26}	$1+\alpha+\alpha^2$	(111000)	59	α^{58}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$	(111111)
28	α^{27}	$\alpha+\alpha^2+\alpha^3$	(011100)	60	α^{59}	$1+\alpha^2+\alpha^3+\alpha^4+\alpha^5$	(101111)
29	α^{28}	$\alpha^2+\alpha^3+\alpha^4$	(001110)	61	α^{60}	$1+\alpha^3+\alpha^4+\alpha^5$	(100111)
30	α^{29}	$\alpha^3+\alpha^4+\alpha^5$	(000111)	62	α^{61}	$1+\alpha^4+\alpha^5$	(100011)
31	α^{30}	$1+\alpha+\alpha^4+\alpha^5$	(110011)	63	α^{62}	$1+\alpha^5$	(100001)

Fonte: Elaborado pelos autores.

Baseado nos cálculos anteriores, a proposta é apresentar uma conexão entre os trabalhos de [1], [2], [6] e [13].

A Tabela 6, apresentada a seguir é composta por:

- 1ª coluna: número representando cada códon.
- 2ª coluna: códons formados pela trinca de bases nitrogenadas.
- 3ª coluna: representação vetorial dos códons, proposta por [6].
- 4ª coluna: representação polinomial proposta por [6].
- 5ª coluna: representação vetorial dos elementos $GF(2^6)$.
- 6ª coluna: representação polinomial dos elementos $GF(2^6)$.

Para a construção das colunas associadas à proposta de [6], é necessário um procedimento baseado na ordem de importância das bases nitrogenadas. Apresentamos a seguir um exemplo, relacionado ao códon AUG.

Exemplo 3.1. Considere o códon AUG. Segundo [6], temos a associação G - 00, U - 10, A - 01 e C - 11. Pela permutação 1203, a associação $\mathbb{Z}_2 \times \mathbb{Z}_2$ é: A - 10, C - 11, G - 00 e U - 01. Então, seguindo a proposta de [6], temos:

$$\text{AUG} \longrightarrow 001001 \longrightarrow x^2 + x^5.$$

Perceba que a representação do códon foi estabelecida da seguinte forma: a primeira posição corresponde a base G (terceira base), representada por 00, já a segunda posição corresponde a base A (primeira base), representada por 10, e por último, a terceira posição corresponde a base U (segunda base), representada por 01. Assim, o grau aumenta da base menos importante, que é a terceira, para a base mais importante, que é a segunda. O procedimento é análogo para a obtenção da representação dos outros códons.

Tabela 6: Rotulamento *A* associado ao $GF(2^6)$ gerado por $p(x) = 1 + x + x^6$.

-	Códon	Vetorial	Polinomial	$GF(2^6)$	Polinomial $GF(2^6)$
1	GGG	000000	0	0	0
2	GGA	100000	1	α^0	1
3	GGU	010000	x	α^1	α^1
4	GGC	110000	$1+x$	α^6	$1+\alpha^1$
5	GAG	000010	x^4	α^4	α^4
6	GAA	100010	$1+x^4$	α^{24}	$1+\alpha^4$
7	GAU	010010	$x+x^4$	α^{33}	$\alpha+\alpha^4$
8	GAC	110010	$1+x+x^4$	α^{16}	$1+\alpha+\alpha^4$
9	GUG	000001	x^5	α^5	α^5
10	GUA	100001	$1+x^5$	α^{62}	$1+\alpha^5$
11	GUU	010001	$x+x^5$	α^{25}	$\alpha+\alpha^5$
12	GUC	110001	$1+x+x^5$	α^{11}	$1+\alpha+\alpha^5$

13	GCG	000011	$x^4 + x^5$	α^{10}	$\alpha^4 + \alpha^5$
14	GCA	100011	$1 + x^4 + x^5$	α^{61}	$1 + \alpha^4 + \alpha^5$
15	GCU	010011	$x + x^4 + x^5$	α^{46}	$\alpha + \alpha^4 + \alpha^5$
16	GCC	110011	$1 + x + x^4 + x^5$	α^{30}	$1 + \alpha + \alpha^4 + \alpha^5$
17	AGG	001000	x^2	α^2	α^2
18	AGA	101000	$1 + x^2 +$	α^{12}	$1 + \alpha^2$
19	AGU	011000	$x + x^2$	α^7	$\alpha + \alpha^2$
20	AGC	111000	$1 + x + x^2$	α^{26}	$1 + \alpha + \alpha^2$
21	AAG	001010	$x^2 + x^4$	α^{14}	$\alpha^2 + \alpha^4$
22	AAA	101010	$1 + x^2 + x^4$	α^{52}	$1 + \alpha^2 + \alpha^4$
23	AAU	011010	$x + x^2 + x^4$	α^{36}	$\alpha + \alpha^2 + \alpha^4$
24	AAC	111010	$1 + x + x^2 + x^4$	α^{54}	$1 + \alpha + \alpha^2 + \alpha^4$
25	AUG	001001	$x^2 + x^5$	α^{34}	$\alpha^2 + \alpha^5$
26	AUA	101001	$1 + x^2 + x^5$	α^{31}	$1 + \alpha^2 + \alpha^5$
27	AUU	011001	$x + x^2 + x^5$	α^{17}	$\alpha + \alpha^2 + \alpha^5$
28	AUC	111001	$1 + x + x^2 + x^5$	α^{47}	$1 + \alpha + \alpha^2 + \alpha^5$
29	ACG	001011	$x^2 + x^4 + x^5$	α^{50}	$\alpha^2 + \alpha^4 + \alpha^5$
30	ACA	101011	$1 + x^2 + x^4 + x^5$	α^{22}	$1 + \alpha^2 + \alpha^4 + \alpha^5$
31	ACU	011011	$x + x^2 + x^4 + x^5$	α^{39}	$\alpha + \alpha^2 + \alpha^4 + \alpha^5$
32	ACC	111011	$1 + x + x^2 + x^4 + x^5$	α^{43}	$1 + \alpha + \alpha^2 + \alpha^4 + \alpha^5$
33	UGG	000100	x^3	α^3	α^3
34	UGA	100100	$1 + x^3$	α^{32}	$1 + \alpha^3$
35	UGU	010100	$x + x^3$	α^{13}	$\alpha + \alpha^3$
36	UGC	110100	$1 + x + x^3$	α^{35}	$1 + \alpha + \alpha^3$
37	UAG	000110	$x^3 + x^4$	α^9	$\alpha^3 + \alpha^4$
38	UAA	100110	$1 + x^3 + x^4$	α^{45}	$1 + \alpha^3 + \alpha^4$
39	UAU	010110	$x + x^3 + x^4$	α^{49}	$\alpha + \alpha^3 + \alpha^4$
40	UAC	110110	$1 + x + x^3 + x^4$	α^{38}	$1 + \alpha + \alpha^3 + \alpha^4$
41	UUG	000101	$x^3 + x^5$	α^{15}	$\alpha^3 + \alpha^5$
42	UUA	100101	$1 + x^3 + x^5$	α^{23}	$1 + \alpha^3 + \alpha^5$
43	UUU	010101	$x + x^3 + x^5$	α^{53}	$\alpha + \alpha^3 + \alpha^5$
44	UUC	110101	$1 + x + x^3 + x^5$	α^{51}	$1 + \alpha + \alpha^3 + \alpha^5$
45	UCG	000111	$x^3 + x^4 + x^5$	α^{29}	$\alpha^3 + \alpha^4 + \alpha^5$
46	UCA	100111	$1 + x^3 + x^4 + x^5$	α^{60}	$1 + \alpha^3 + \alpha^4 + \alpha^5$
47	UCU	010111	$x + x^3 + x^4 + x^5$	α^{42}	$\alpha + \alpha^3 + \alpha^4 + \alpha^5$
48	UCC	110111	$1 + x + x^3 + x^4 + x^5$	α^{21}	$1 + \alpha + \alpha^3 + \alpha^4 + \alpha^5$
49	CGG	001100	$x^2 + x^3$	α^8	$\alpha^2 + \alpha^3$
50	CGA	101100	$1 + x^2 + x^3$	α^{48}	$1 + \alpha^2 + \alpha^3$
51	CGU	011100	$x + x^2 + x^3$	α^{27}	$\alpha + \alpha^2 + \alpha^3$
52	CGC	111100	$1 + x + x^2 + x^3$	α^{18}	$1 + \alpha + \alpha^2 + \alpha^3$
53	CAG	001110	$x^2 + x^3 + x^4$	α^{28}	$\alpha^2 + \alpha^3 + \alpha^4$
54	CAA	101110	$1 + x^2 + x^3 + x^4$	α^{41}	$1 + \alpha^2 + \alpha^3 + \alpha^4$
55	CAU	011110	$x + x^2 + x^3 + x^4$	α^{19}	$\alpha + \alpha^2 + \alpha^3 + \alpha^4$
56	CAC	111110	$1 + x + x^2 + x^3 + x^4$	α^{56}	$1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4$
57	CUG	001101	$x^2 + x^3 + x^5$	α^{37}	$\alpha^2 + \alpha^3 + \alpha^5$
58	CUA	101101	$1 + x^2 + x^3 + x^5$	α^{44}	$1 + \alpha^2 + \alpha^3 + \alpha^5$
59	CUU	011101	$x + x^2 + x^3 + x^5$	α^{55}	$\alpha + \alpha^2 + \alpha^3 + \alpha^5$
60	CUC	111101	$1 + x + x^2 + x^3 + x^5$	α^{40}	$1 + \alpha + \alpha^2 + \alpha^3 + \alpha^5$
61	CCG	001111	$x^2 + x^3 + x^4 + x^5$	α^{20}	$\alpha^2 + \alpha^3 + \alpha^4 + \alpha^5$
62	CCA	101111	$1 + x^2 + x^3 + x^4 + x^5$	α^{59}	$1 + \alpha^2 + \alpha^3 + \alpha^4 + \alpha^5$
63	CCU	011111	$x + x^2 + x^3 + x^4 + x^5$	α^{57}	$\alpha + \alpha^2 + \alpha^3 + \alpha^4 + \alpha^5$
64	CCC	111111	$1 + x + x^2 + x^3 + x^4 + x^5$	α^{58}	$\alpha + \alpha^2 + \alpha^3 + \alpha^4 + \alpha^5$

Fonte: Elaborado pelos autores.

3.2 Detalhamento dos Rotulamentos B e C

A seguir, será apresentado o mapeamento do código genético, considerando a permutação 0321 do rotulamento *B*, apresentada na Figura 2. Da mesma forma que o rotulamento *A*, o procedimento para tal mapeamento seguirá uma representação polinomial, uma vetorial e uma por potência, utilizando o polinômio primitivo $x^6 + x^5 + 1$ para a geração da extensão de Galois $GF(2^6)$.

A Tabela 7 apresenta as representações por potência, polinomial e vetorial, construída seguindo o mesmo procedimento do rotulamento *A*, associados aos 64 códons do código genético.

Tabela 7: Rotulamento *B* associado ao $GF(2^6)$ gerado por $p(x) = 1 + x^5 + x^6$.

-	Códon	Vetorial	Polinomial	$GF(2^6)$	Polinomial $GF(2^6)$
1	AAA	000000	0	0	0
2	AAU	100000	1	α^0	α^0
3	AAC	010000	x	α^1	α^1
4	AAG	110000	$1+x$	α^{58}	$1+\alpha^1$
5	AUA	000010	x^4	α^4	α^4
6	AUU	100010	$1+x^4$	α^{43}	$1+\alpha^4$
7	AUC	010010	$x+x^4$	α^{35}	$\alpha+\alpha^4$
8	AUG	110010	$1+x+x^4$	α^{22}	$1+\alpha+\alpha^4$
9	ACA	000001	x^5	α^5	α^5
10	ACU	100001	$1+x^5$	α^6	$1+\alpha^5$
11	ACC	010001	$x+x^5$	α^{44}	$x+\alpha^5$
12	ACG	110001	$1+x+x^5$	α^7	$1+x+\alpha^5$
13	AGA	000011	x^4+x^5	α^{62}	$\alpha^4+\alpha^5$
14	AGU	100011	$1+x^4+x^5$	α^{57}	$1+\alpha^4+\alpha^5$
15	AGC	010011	$x+x^4+x^5$	α^{52}	$\alpha+\alpha^4+\alpha^5$
16	AGG	110011	$1+x+x^4+x^5$	α^{38}	$1+\alpha+\alpha^4+\alpha^5$
17	UAA	001000	x^2	α^2	α^2
18	UAU	101000	$1+x^2$	α^{53}	$1+\alpha^2$
19	UAC	011000	$x+x^2$	α^{59}	$\alpha+\alpha^2$
20	UAG	111000	$1+x+x^2$	α^{39}	$1+\alpha+\alpha^2$
21	UUA	001010	x^2+x^4	α^{55}	$\alpha^2+\alpha^4$
22	UUU	101010	$1+x^2+x^4$	α^{15}	$1+\alpha^2+\alpha^4$
23	UUC	011010	$x+x^2+x^4$	α^{19}	$\alpha+\alpha^2+\alpha^4$
24	UUG	111010	$1+x+x^2+x^4$	α^{26}	$1+\alpha+\alpha^2+\alpha^4$
25	UCA	001001	x^2+x^5	α^{36}	$\alpha^2+\alpha^5$
26	UCU	101001	$1+x^2+x^5$	α^{45}	$1+\alpha^2+\alpha^5$
27	UCC	011001	$x+x^2+x^5$	α^{23}	$\alpha+\alpha^2+\alpha^5$
28	UCG	111001	$1+x+x^2+x^5$	α^8	$1+\alpha+\alpha^2+\alpha^5$
29	UGA	001011	$x^2+x^4+x^5$	α^{33}	$\alpha^2+\alpha^4+\alpha^5$
30	UGU	101011	$1+x^2+x^4+x^5$	α^{17}	$1+\alpha^2+\alpha^4+\alpha^5$
31	UGC	011011	$x+x^2+x^4+x^5$	α^{30}	$\alpha+\alpha^2+\alpha^4+\alpha^5$
32	UGG	111011	$1+x+x^2+x^4+x^5$	α^{47}	$1+\alpha+\alpha^2+\alpha^4+\alpha^5$
33	CAA	000100	x^3	α^3	α^3
34	CAU	100100	$1+x^3$	α^{34}	$1+\alpha^3$
35	CAC	010100	$x+x^3$	α^{54}	$\alpha+\alpha^3$
36	CAG	110100	$1+x+x^3$	α^{18}	$1+\alpha+\alpha^3$
37	CUA	000110	x^3+x^4	α^{61}	$\alpha^3+\alpha^4$
38	CUU	100110	$1+x^3+x^4$	α^{51}	$1+\alpha^3+\alpha^4$
39	CUC	010110	$x+x^3+x^4$	α^{32}	$\alpha+\alpha^3+\alpha^4$
40	CUG	110110	$1+x+x^3+x^4$	α^{29}	$1+\alpha+\alpha^3+\alpha^4$
41	CCA	000101	x^3+x^5	α^{56}	$\alpha^3+\alpha^5$
42	CCU	100101	$1+x^3+x^5$	α^{37}	$1+\alpha^3+\alpha^5$
43	CCC	010101	$x+x^3+x^5$	α^{16}	$\alpha+\alpha^3+\alpha^5$
44	CCG	110101	$1+x+x^3+x^5$	α^{46}	$1+\alpha+\alpha^3+\alpha^5$
45	CGA	000111	$x^3+x^4+x^5$	α^{42}	$\alpha^3+\alpha^4+\alpha^5$
46	CGU	100111	$1+x^3+x^4+x^5$	α^{21}	$1+\alpha^3+\alpha^4+\alpha^5$
47	CGC	010111	$x+x^3+x^4+x^5$	α^{14}	$\alpha+\alpha^3+\alpha^4+\alpha^5$
48	CGG	110111	$1+x+x^3+x^4+x^5$	α^{25}	$1+\alpha+\alpha^3+\alpha^4+\alpha^5$
49	GAA	001100	x^2+x^3	α^{60}	$\alpha^2+\alpha^3$
50	GAU	101100	$1+x^2+x^3$	α^{31}	$1+\alpha^2+\alpha^3$
51	GAC	011100	$x+x^2+x^3$	α^{40}	$\alpha+\alpha^2+\alpha^3$
52	GAG	111100	$1+x+x^2+x^3$	α^{48}	$1+\alpha+\alpha^2+\alpha^3$
53	GUA	001110	$x^2+x^3+x^4$	α^{41}	$\alpha^2+\alpha^3+\alpha^4$
54	GUU	101110	$1+x^2+x^3+x^4$	α^{13}	$1+\alpha^2+\alpha^3+\alpha^4$
55	GUC	011110	$x+x^2+x^3+x^4$	α^{49}	$\alpha+\alpha^2+\alpha^3+\alpha^4$
56	GUG	111110	$1+x+x^2+x^3+x^4$	α^{11}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4$
57	GCA	001101	$x^2+x^3+x^5$	α^{20}	$\alpha^2+\alpha^3+\alpha^5$
58	GCU	101101	$1+x^2+x^3+x^5$	α^{24}	$1+\alpha^2+\alpha^3+\alpha^5$
59	GCC	011101	$x+x^2+x^3+x^5$	α^{27}	$\alpha+\alpha^2+\alpha^3+\alpha^5$
60	GCG	111101	$1+x+x^2+x^3+x^5$	α^9	$1+\alpha+\alpha^2+\alpha^3+\alpha^5$
61	GGA	001111	$x^2+x^3+x^4+x^5$	α^{50}	$\alpha^2+\alpha^3+\alpha^4+\alpha^5$
62	GGU	101111	$1+x^2+x^3+x^4+x^5$	α^{28}	$1+\alpha^2+\alpha^3+\alpha^4+\alpha^5$
63	GGC	011111	$x+x^2+x^3+x^4+x^5$	α^{12}	$\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$
64	GGG	111111	$1+x+x^2+x^3+x^4+x^5$	α^{10}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$

Fonte: Elaborado pelos autores.

Análogo aos procedimentos apresentados anteriormente, considere a permutação 3120 do rotulamento *C*, da Figura 2. Da mesma forma que o rotulamento *A* e o rotulamento *B*, o procedimento para tal mapeamento seguirá uma representação polinomial, uma vetorial e uma por potência, utilizando o polinômio primitivo $x^6 + x^5 + x^2 + x + 1$ para a geração da extensão de Galois $GF(2^6)$, apresentados na Tabela 8.

Tabela 8: Rotulamento *C* associado ao $GF(2^6)$ gerado por $p(x) = 1 + x + x^2 + x^5 + x^6$.

-	Códon	Vetorial	Polinomial	$GF(2^6)$	Polinomial $GF(2^6)$
1	UUU	000000	0	0	0
2	UUC	100000	1	α^0	α^0
3	UUA	010000	x	α^1	α^1
4	UUG	110000	$1+x$	α^{25}	$1+\alpha^1$
5	UCU	000010	x^4	α^4	α^4
6	UCC	100010	$1+x^4$	α^{37}	$1+\alpha^4$
7	UCA	010010	$x+x^4$	α^{56}	$\alpha+\alpha^4$
8	UCG	110010	$1+x+x^4$	α^{46}	$1+\alpha+\alpha^4$
9	UAU	000001	x^5	α^5	α^5
10	UAC	100001	$1+x^5$	α^{40}	$1+\alpha^5$
11	UAA	010001	$x+x^5$	α^{38}	$\alpha+\alpha^5$
12	UAG	110001	$1+x+x^5$	α^{39}	$1+\alpha+\alpha^5$
13	UGU	000011	x^4+x^5	α^{29}	$\alpha^4+\alpha^5$
14	UGC	100011	$1+x^4+x^5$	α^{49}	$1+\alpha^4+\alpha^5$
15	UGA	010011	$x+x^4+x^5$	α^{24}	$\alpha+\alpha^4+\alpha^5$
16	UGG	110011	$1+x+x^4+x^5$	α^{62}	$1+\alpha+\alpha^4+\alpha^5$
17	CUU	001000	x^2	α^2	α^2
18	CUC	101000	$1+x^2$	α^{50}	$1+\alpha^2$
19	CUA	011000	$x+x^2$	α^{26}	$\alpha+\alpha^2$
20	CUG	111000	$1+x+x^2$	α^{30}	$1+\alpha+\alpha^2$
21	CCU	001010	x^2+x^4	α^{52}	$\alpha^2+\alpha^4$
22	CCC	101010	$1+x^2+x^4$	α^{60}	$1+\alpha^2+\alpha^4$
23	CCA	011010	$x+x^2+x^4$	α^{20}	$\alpha+\alpha^2+\alpha^4$
24	CCG	111010	$1+x+x^2+x^4$	α^{34}	$1+\alpha+\alpha^2+\alpha^4$
25	CAU	001001	x^2+x^5	α^{57}	$\alpha^2+\alpha^5$
26	CAC	101001	$1+x^2+x^5$	α^{41}	$1+\alpha^2+\alpha^5$
27	CAA	011001	$x+x^2+x^5$	α^{47}	$\alpha+\alpha^2+\alpha^5$
28	CAG	111001	$1+x+x^2+x^5$	α^6	$1+\alpha+\alpha^2+\alpha^5$
29	CGU	001011	$x^2+x^4+x^5$	α^{11}	$\alpha^2+\alpha^4+\alpha^5$
30	CGC	101011	$1+x^2+x^4+x^5$	α^8	$1+\alpha^2+\alpha^4+\alpha^5$
31	CGA	011011	$x+x^2+x^4+x^5$	α^{18}	$\alpha+\alpha^2+\alpha^4+\alpha^5$
32	CGG	111011	$1+x+x^2+x^4+x^5$	α^{54}	$1+\alpha+\alpha^2+\alpha^4+\alpha^5$
33	AUU	000100	x^3	α^3	α^3
34	AUC	100100	$1+x^3$	α^{55}	$1+\alpha^3$
35	AUA	010100	$x+x^3$	α^{51}	$\alpha+\alpha^3$
36	AUG	110100	$1+x+x^3$	α^{19}	$1+\alpha+\alpha^3$
37	ACU	000110	x^3+x^4	α^{28}	$\alpha^3+\alpha^4$
38	ACC	100110	$1+x^3+x^4$	α^{23}	$1+\alpha^3+\alpha^4$
39	ACA	010110	$x+x^3+x^4$	α^{10}	$\alpha+\alpha^3+\alpha^4$
40	ACG	110110	$1+x+x^3+x^4$	α^{17}	$1+\alpha+\alpha^3+\alpha^4$
41	AAU	000101	x^3+x^5	α^{53}	$\alpha^3+\alpha^5$
42	AAC	100101	$1+x^3+x^5$	α^7	$1+\alpha^3+\alpha^5$
43	AAA	010101	$x+x^3+x^5$	α^{61}	$\alpha+\alpha^3+\alpha^5$
44	AAG	110101	$1+x+x^3+x^5$	α^{48}	$1+\alpha+\alpha^3+\alpha^5$
45	AGU	000111	$x^3+x^4+x^5$	α^{33}	$\alpha^3+\alpha^4+\alpha^5$
46	AGC	100111	$1+x^3+x^4+x^5$	α^{59}	$1+\alpha^3+\alpha^4+\alpha^5$
47	AGA	010111	$x+x^3+x^4+x^5$	α^{45}	$\alpha+\alpha^3+\alpha^4+\alpha^5$
48	AGG	110111	$1+x+x^3+x^4+x^5$	α^{36}	$1+\alpha+\alpha^3+\alpha^4+\alpha^5$
49	GUU	001100	x^2+x^3	α^{27}	$\alpha^2+\alpha^3$
50	GUC	101100	$1+x^2+x^3$	α^9	$1+\alpha^2+\alpha^3$
51	GUA	011100	$x+x^2+x^3$	α^{31}	$\alpha+\alpha^2+\alpha^3$
52	GUG	111100	$1+x+x^2+x^3$	α^{12}	$1+\alpha+\alpha^2+\alpha^3$
53	GCU	001110	$x^2+x^3+x^4$	α^{32}	$\alpha^2+\alpha^3+\alpha^4$
54	GCC	101110	$1+x^2+x^3+x^4$	α^{44}	$1+\alpha^2+\alpha^3+\alpha^4$
55	GCA	011110	$x+x^2+x^3+x^4$	α^{13}	$\alpha+\alpha^2+\alpha^3+\alpha^4$
56	GCG	111110	$1+x+x^2+x^3+x^4$	α^{15}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4$
57	GAU	001101	$x^2+x^3+x^5$	α^{21}	$\alpha^2+\alpha^3+\alpha^5$
58	GAC	101101	$1+x^2+x^3+x^5$	α^{42}	$1+\alpha^2+\alpha^3+\alpha^5$
59	GAA	011101	$x+x^2+x^3+x^5$	α^{35}	$\alpha+\alpha^2+\alpha^3+\alpha^5$
60	GAG	111101	$1+x+x^2+x^3+x^5$	α^{58}	$1+\alpha+\alpha^2+\alpha^3+\alpha^5$
61	GGU	001111	$x^2+x^3+x^4+x^5$	α^{14}	$\alpha^2+\alpha^3+\alpha^4+\alpha^5$
62	GGC	101111	$1+x^2+x^3+x^4+x^5$	α^{43}	$1+\alpha^2+\alpha^3+\alpha^4+\alpha^5$
63	GGA	011111	$x+x^2+x^3+x^4+x^5$	α^{16}	$\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$
64	GGG	111111	$1+x+x^2+x^3+x^4+x^5$	α^{22}	$1+\alpha+\alpha^2+\alpha^3+\alpha^4+\alpha^5$

Fonte: Elaborado pelos autores.

Foram apresentadas construções para os rotulamentos A, B e C, por meio da representação usando elementos da extensão $GF(2^6)$ conectando com a proposta de [6]. Desta forma, vemos uma interessante caracterização algébrica do código genético, de vital importância para uma série de procedimentos ainda pouco estudados envolvendo fenômenos mutacionais.

4 CONSIDERAÇÕES FINAIS

A modelagem do código genético por meio de ferramentas matemáticas tem sido alvo de diversas pesquisas. Nesses trabalhos, os autores buscam encontrar relações entre essas áreas e como uma pode influenciar na outra. Essas relações remetem a uma questão muito relevante, que é a interdisciplinaridade, uma vez que Matemática e Biologia estão sendo trabalhadas de maneira conjunta.

Buscamos dessa forma, conectar elementos que a princípio pareciam muito distantes, sendo estes o código genético (Biologia) e estruturas algébricas (Matemática). Por meio desta relação, obtemos a caracterização algébrica de cada um dos rotulamentos associados ao código genético, usando uma representação polinomial, vetorial e por potência, através de um polinômio primitivo em cada um dos casos analisados. Com isso, observa-se uma conexão entre os elementos da extensão de Galois $GF(2^6)$ e os códons do código genético.

Relacionando Matemática e Biologia, vemos um interessante elo entre o código genético e extensões de Galois, que torna-se relevante em diversas áreas de pesquisa, como engenharia genética, uma vez que pode servir como base para estudos relacionados a fenômenos mutacionais.

AGRADECIMENTOS

Os autores agradecem à Universidade Federal de Alfenas - UNIFAL - MG.

REFERÊNCIAS

[1] L. C. B. D. Faria, "Existência de códigos corretores de erros e protocolos de comunicação em sequência de DNA", Doutorado em Engenharia Elétrica, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Campinas, jul. 2011.

[2] A. S. L. Rocha, "Modelo de sistema de comunicação digital para o mecanismo de importação de proteínas mitocondriais através de códigos corretores de erros", Doutorado em Engenharia Elétrica, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Campinas, fev. 2010.

[3] M. E. D. Gonzalez, "Modelagem da síntese de proteínas e sua estrutura organizacional através de códigos corretores de erros", Doutorado em Engenharia Elétrica, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Campinas, fev. 2017.

[4] B. P. Rocha and A. J. de Oliveira, "Aplicações de estruturas algébricas na conexão entre sistemas de comunicação padrão e genética", *Sigmae*, vol. 6, no. 2, pp. 1-14, 2017.

[5] R. S. Fernandes and A. J. de Oliveira, “Caracterização das propriedades dos aminoácidos por meio do diagrama de Hasse associado ao rotulamento A do código genético”, *Brazilian Electronic Journal of Mathematics*, vol. 2, n^o. 4, pp. 81-100, 2021.

[6] R. Sánchez, E. Morgado, and R. Grau, “Gene algebra from a genetic code algebraic structure”, *Journal of Mathematical Biology*, vol. 51, n^o. 4, pp. 431-457, 2005.

[7] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Biologia Molecular da Célula*. 5 ed. Porto Alegre: ArtMed, 2010.

[8] H. H. Domingues and G. Iezzi, *Álgebra Moderna*. 4 ed. São Paulo: Atual, 2003.

[9] A. J. F. Griffiths, S. R. Wessler, R. C. Lewontin, W. M. Gelbart, D. T. Suzuki, and J. H. Miller, *Introdução à genética*. Rio de Janeiro: Guanabara Koogan, 2006.

[10] A. Hefez and C. S. Fernandez, *Introdução à Álgebra Linear*. 1 ed. Rio de Janeiro: SBM, 2012.

[11] J. Hib and E. M. F. Robertis, *Bases da biologia celular e molecular*. Rio de Janeiro: Guanabara Koogan, 2006.

[12] S. Lin and J. D. Costelo Jr., *Error Control Coding: Fundamentals and Applications*. São Paulo: Englewood Cliffs: Prentice Hall, 1983.

[13] A. J. de Oliveira, “Análise algébrica dos rotulamentos associados ao mapeamento do código genético”, Mestrado em Engenharia Elétrica, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Campinas, fev. 2012.

BREVE BIOGRAFIA

Roberta Siqueira Fernandes https://orcid.org/0000-0002-5286-704X

Mestra em Estatística Aplicada e Biometria na Universidade Federal de Alfenas, na área de Matemática Aplicada. Finalizou a graduação em Licenciatura-Matemática pela Universidade Federal de Alfenas. Fez parte Programa Institucional de Bolsas de Iniciação à Docência (Pibid) como bolsista. Também, participou do projeto “Conversas Matemáticas” do Programa Universidade Aberta à Terceira Idade (UNATI) da UNIFAL-MG como voluntária, e do projeto “Laboratório de Ensino em Educação Matemática na formação de professores” como bolsista.

Anderson José de Oliveira https://orcid.org/0000-0002-5467-9827

Graduado em Licenciatura Plena em Matemática pelo Centro Universitário do Sul de Minas, (2006). Pós-graduado em Matemática Empresarial, (2007) e Docência em Ead, (2008), pelo Centro Universitário do Sul de Minas. Mestre em Engenharia Elétrica, (2012), pela UNICAMP. Doutor em Engenharia Elétrica, (2017), pela UNICAMP. Atuou como professor universitário de 2007 a 2009 no Centro Universitário do Sul de Minas e como professor de Matemática, no Centro Integrado Sesi Senai, de São Gonçalo do Sapucaí para o Ensino Médio, de 2008 a 2009. Atualmente é professor efetivo do Instituto de Ciências Exatas, Departamento de Matemática, da Universidade Federal de Alfenas - UNIFAL-MG. Trabalha com pesquisas na área de Matemática Aplicada (Códigos Corretores de Erros e Estruturas Algébricas Aplicados na Modelagem do Código Genético, Equações Diferenciais Fuchsianas e Geometria Hiperbólica Aplicadas no Processo de Transmissão de Informação em Sistemas de Comunicação).