

Caracterização Geométrica das Permutações Associadas aos Rotulamentos do Código Genético

Edmara Viana da Silva, Anderson José de Oliveira

Matemática-Licenciatura

Universidade Federal de Alfenas

edmaraviana076@gmail.com, anderson.oliveira@unifal-mg.edu.br

1. Introdução

O estudo da modelagem matemática do código genético é algo que diversos pesquisadores vem desenvolvendo e a busca de conexões entre essas duas importantes áreas (Matemática e Biologia) se torna relevante na busca de representações matemáticas do que o mundo biológico possui e as possibilidades de análises das propriedades e características dos mesmos. Rocha [1] e Faria [2] fazem a proposta de um modelo de um sistema de codificação e decodificação do mecanismo de importação de proteínas mitocondriais. Através do mapeamento das sequências de DNA, utiliza-se um código BCH sobre a estrutura de anel para a geração de sequências de DNA. O mapeamento foi feito por meio da bijeção de um alfabeto biológico das bases nitrogenadas $N = \{A, C, G, T/U\}$, em que A indica adenina, C citosina, G guanina e T/U timina e uracila, respectivamente, com o alfabeto matemático do anel $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. Esse mapeamento $N \rightarrow \mathbb{Z}_4$ consiste de 24 permutações, que podem ser divididas em três rotulamentos: A, B e C, de acordo com as formas geométricas que produzem um diferente nível de não-linearidade para as sequências reproduzidas. As permutações associadas ao rotulamento A levam ao mapeamento \mathbb{Z}_4 -linear; as permutações associadas ao rotulamento B levam ao mapeamento $\mathbb{Z}_2 \times \mathbb{Z}_2$ -linear, enquanto que a permutações associadas ao rotulamento C levam ao mapeamento Klein-linear. O objetivo deste trabalho é apresentar detalhadamente a representação geométrica de cada uma das permutações associadas aos rotulamentos A, B e C de forma a analisar as similitudes e diferenças entre tais representações.

2. Modelagem do Código Genético

Segundo Rocha [1], a modelagem do código genético é feita da seguinte maneira: é feita uma analogia entre o alfabeto 4-ário na saída da fonte com o conjunto de nucleotídeos, denotado por $N = \{A, C, G, T/U\}$, correspondendo as bases adenina (A), citosina (C), guanina (G), e timina (T) ou uracila (U). Similarmente, o alfabeto 4-ário é denotado por $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ para a estrutura de anel. Como o mapeamento entre $N \rightarrow \mathbb{Z}_4$ é desconhecido, temos que toda sequência de DNA é rotulada por cada uma das 24 permutações entre $N \rightarrow \mathbb{Z}_4$. Observamos que para cada sequência reproduzida pelo código existem oito palavras-código correspondentes a 8 permutações. Empregando o rotulamento recíproco, $\mathbb{Z}_4 \rightarrow N$ em cada uma dessas 8 palavras-código temos como resultado 8 palavras-código iguais em termos de nucleotídeos e aminoácidos. Este fato resulta em três conjuntos contendo oito permutações cada um. Classificamos o mapeamento utilizado no codificador, em termos de rotulamentos A, B e C. Esses rotulamentos estão relacionados às formas geométricas que produzem um diferente nível de não-linearidade para as sequências reproduzidas. A representação binária associada a cada um desses rotulamentos é 0 - 00; 1 - 10; 2 - 11; 3 - 01. Todavia, a associação da complementaridade dos nucleotídeos A - T e C - G com os rotulamentos é o que os diferenciam. No caso do rotulamento A, qualquer um dos nucleotídeos para alcançar o seu complementar necessita caminhar duas arestas, enquanto que nos rotulamentos B e C, basta caminhar uma aresta somente. Todas as permutações associadas ao rotulamento A caracterizam o mapeamento como \mathbb{Z}_4 -linear; as permutações associadas ao rotulamento B caracterizam o mapeamento como $\mathbb{Z}_2 \times \mathbb{Z}_2$ -linear; enquanto que as permutações associadas ao rotulamento C caracterizam o mapeamento como Klein-linear. O rotulamento A classifica as sequências como não-lineares (\mathbb{Z}_4 -linear), enquanto que os rotulamentos B e C as classificam como lineares ($\mathbb{Z}_2 \times \mathbb{Z}_2$ -linear e Klein-linear). A Figura 1 a seguir apresenta o mapeamento do código genético e os rotulamentos associados ao mesmo.

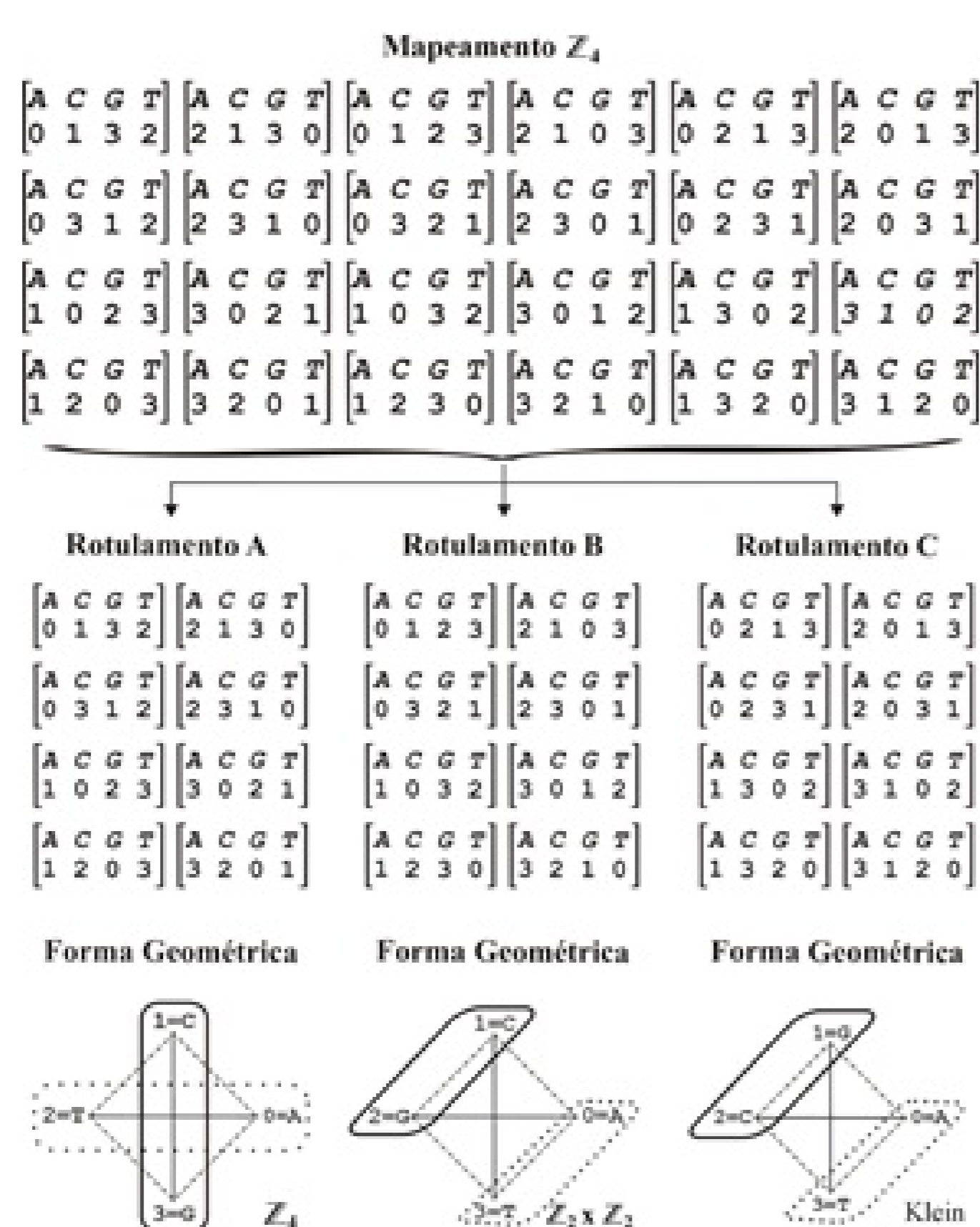


Figura 1. Rotulamentos.

3. Resultados e Discussão

As Figuras 2, 3 e 4 apresentam a construção das 24 permutações associadas ao mapeamento do código genético, divididas da seguinte maneira: 8 permutações associadas ao rotulamento A, nas quais cada nucleotídeo necessita caminhar duas arestas para encontrar o seu complementar, 8 permutações associadas ao rotulamento B e 8 ao rotulamento C, nas quais cada nucleotídeo necessita caminhar apenas uma aresta para encontrar o seu

complementar. Desta forma, foi estabelecida uma caracterização geométrica associada a cada uma das permutações dos rotulamentos do código genético.

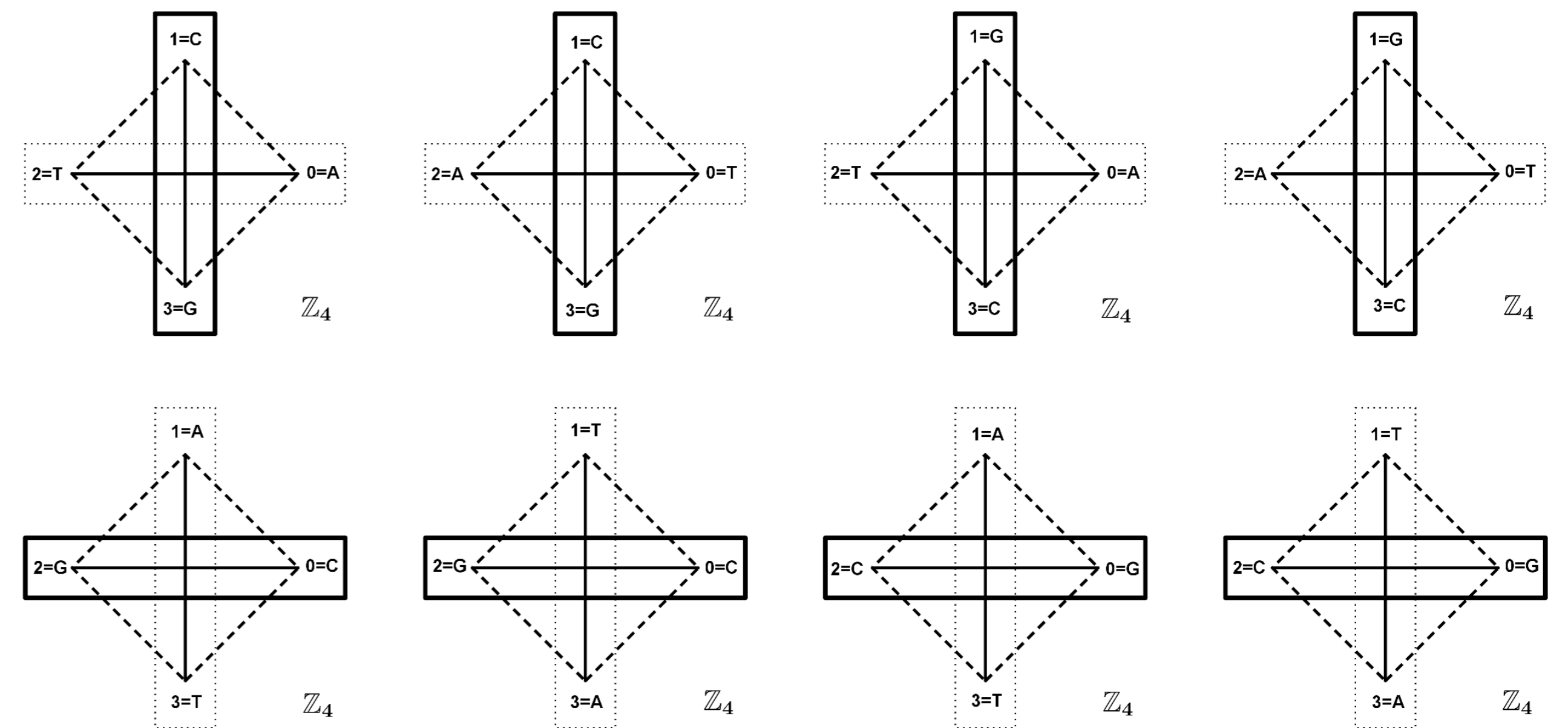


Figura 2. Permutações Associadas ao Rotulamento A.

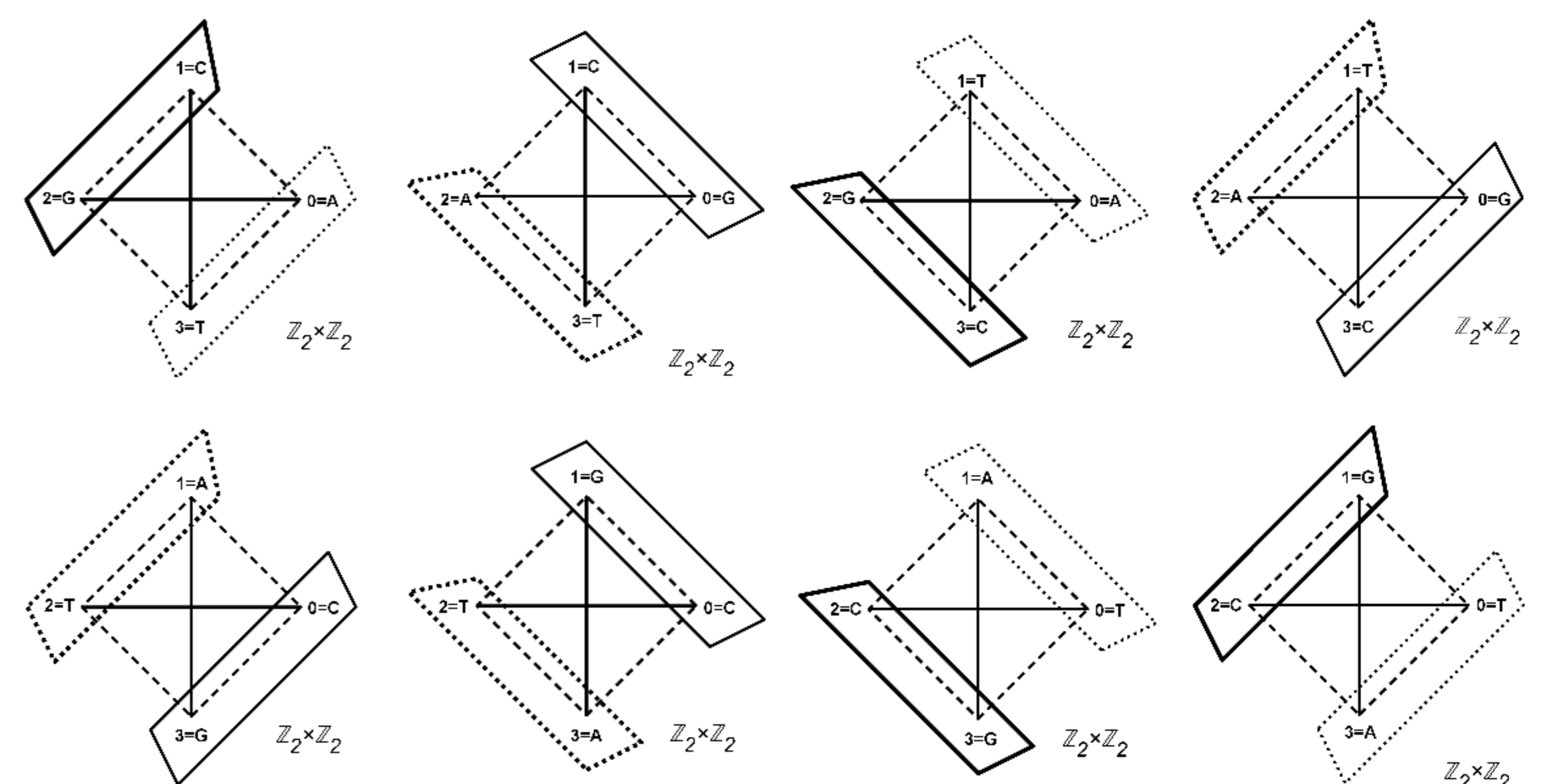


Figura 3. Permutações Associadas ao Rotulamento B.

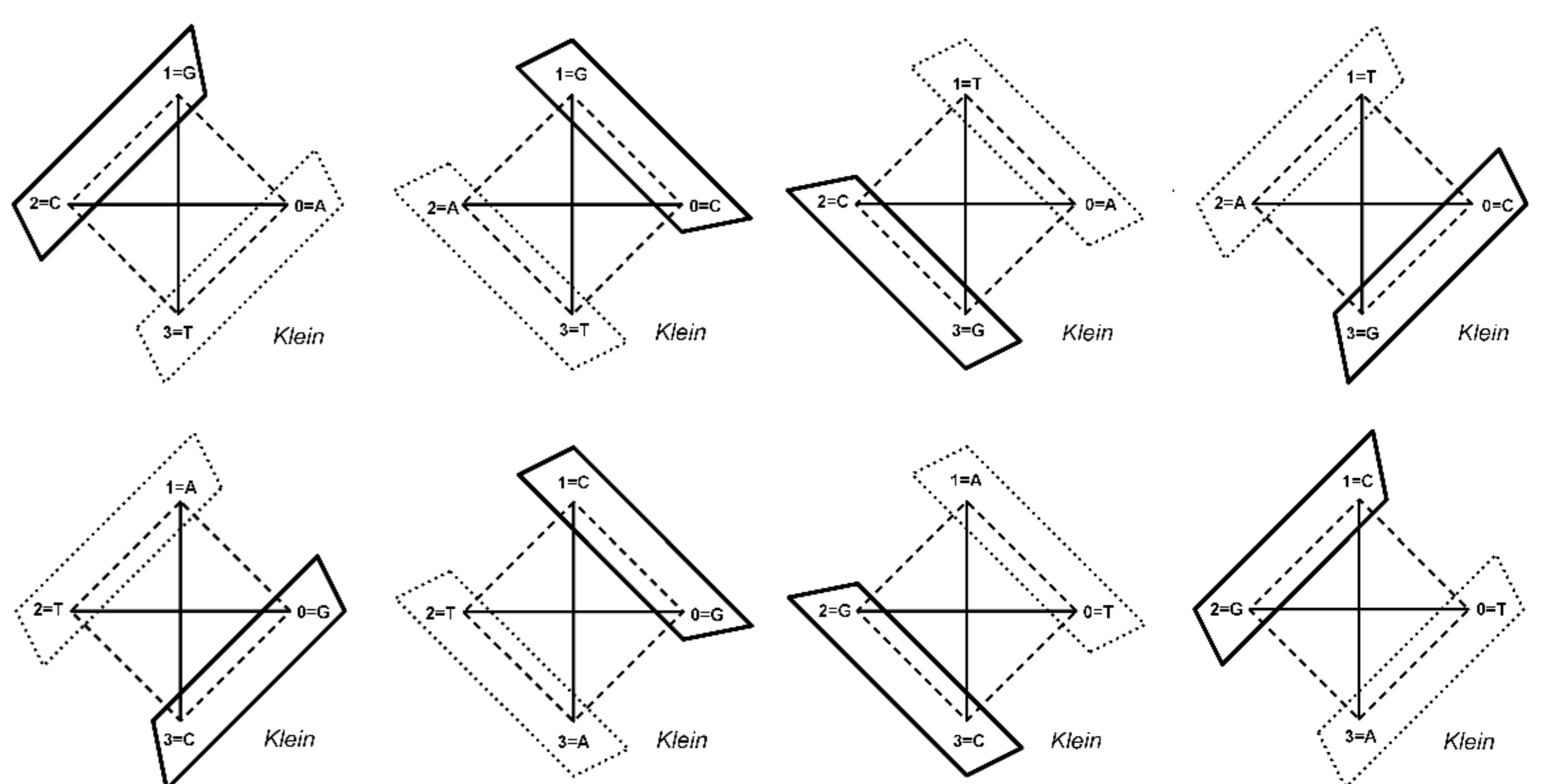


Figura 4. Permutações Associadas ao Rotulamento C.

Referências

- [1] A. S. L. Rocha, R. Palazzo Jr., M.C. Silva-Filho. Modelo de sistema de comunicações digital para o mecanismo de importação de proteínas mitocondriais através de códigos corretores de erros, Tese de Doutorado, FEEC-Unicamp, 2010.
- [2] L. C. B Faria, R. Palazzo Jr. Existências de códigos corretores de erros e protocolos de comunicação em sequências de DNA, Tese de Doutorado, FEEC-Unicamp, 2011.

4. Agradecimentos