Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica
Departamento de Comunicações

Reticulados e suas Partições Aplicados à Codificação para Canais AWGN Limitados em Banda

Autor: Mauro Antonio Órrego da Costa e Silva
Orientador: Prof. Dr. Reginaldo Palazzo Jr.

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica.

Junho de 1991
Campinas - SP

À minha esposa e filhos, razão e motivação.

Aos meus pais, compreensão e apoio.

Agradecimentos

Agradeço em especial ao Prof. Dr. Reginaldo Palazzo Jr. pela oportunidade e pelas inestimáveis sugestões e discussões.

Agradeço também, pela valiosa participação na Banca Examinadora, aos Profs. Drs. Rege Romeu Scarabucci e Helio Waldman do DECOM/FEE/UNICAMP, e aos Profs. Drs. Aydano B. Carleial do INPE e José Carlos Khill do IMECC/UNICAMP.

Igualmente, agradeço à CAPES, à FEE/UNICAMP e ao CNPq pelo apoio financeiro em várias etapas deste trabalho, bem como à EESC/USP por possibilitar a sua conclusão.

Agradeço, ainda, à Serifa Editoração, pela presteza na editoração eletrônica e à EESC/USP pelo apoio gráfico.

Resumo

Neste trabalho são realizados estudos teóricos e aplicados dos reticulados e suas partições. Após uma breve revisão de conceitos algébricos e geométricos sobre os reticulados e suas partições, utilizando a terminologia correspondente para grupos abelianos, é desenvolvida uma descrição explícita da estrutura algébrica de partições arbitrárias de reticulados, incluindo a utilização de formas canônicas de matrizes inteiras. Em seguida, após uma análise da avaliação e da comparação de codificadores para o canal AWGN limitado em banda em termos de desempenho e complexidade, são revistas sumariamente as formas gerais dos esquemas de construção existentes desses codificadores utilizando reticulados e suas partições, evidenciando as características relevantes dos reticulados utilizados para a obtenção de codificadores de alto desempenho e baixa complexidade. É proposto, então, um esquema multinível de construção de reticulados, que possibilita o desenvolvimento de um algorítmo de decodificação por estágios de vários reticulados novos e conhecidos, para os quais são avaliados o desempenho e a complexidade. Verificou-se um substancial melhoramento do compromisso desempenho vs complexidade, no sentido de ter trazido os reticulados construídos para mais próximo da atual fronteira de eficiência de codificação, composta pelos melhores códigos conhecidos para o canal AWGN limitado em banda. Algumas extensões do estudo realizado são indicadas para pesquisas futuras.

Abstract

Theoretical and applied studies on lattices and their partitions are made in this work. After a brief review of algebraic and geometric concepts on lattices and their partitions, using the corresponding terminology for abelian groups, an explicit description of the algebraic structure of arbitrary lattice partitions is developed, including the use of canonical forms of integer matrices. Following this, after an analysis of the evaluation and comparison of encoders for the bandlimited AWGN channel in terms of performance and complexity, the general forms of the existing schemes using lattices and their partitions for the construction of these encoders are summarized, emphasizing the relevant characteristics of the used lattices to get encoders with high performance and low complexity. A multilevel scheme for lattice construction is then proposed, making possible the development of a multistage decoding algorithm for various known and new lattices, for which the performance and complexity are evaluated. A substantial improvement in the tradeoff between performance and complexity was reached, in the sense of taking the constructed lattices closer to the current efficiency frontier, which is set by the best codes for the bandlimited AWGN channel. Some extensions of these studies are pointed out for future investigations.

Sumário

Capítulo 1

Introdução

1.1 Histórico

Canais AWGN limitados em banda são canais operados com altas relações sinal/ruído e com número limitado de dimensões por unidade de tempo (Gallager [20], Wozencraft e Jacobs [38], Viterbi e Omura [35]), que podem servir de modelo para vários canais reais de importância comercial, em particular para o canal telefônico, historicamente o primeiro canal sobre o qual foram utilizadas técnicas de codificação para canais AWGN limitados em banda. É interessante notar que, embora o trabalho original de Shannon ([29]) tenha evidenciado um ganho potencial da ordem de 9 dB em relação à sinalização PAM neste canal, apenas na última década pode-se notar um esforço sério no desenvolvimento de técnicas de codificação construtivas que podem auferir frações consideráveis desse ganho teoricamente disponível.

O primeiro passo substancial nesta direção foi dado por Ungerboeck ([33]), combinando códigos convolucionais e uma técnica de rotulagem dos sinais utilizados denominada "mapeamento por partição de conjuntos", para obter ganhos de 3 dB a 6 dB, com uma larga faixa de compromissos entre desempenho e complexidade, utilizando sinais unidimensionais e bidimensionais. A combinação de constelações de sinais multidimensionais com códigos em treliça passou também a ser explorada, ampliando-se assim aquela faixa disponível de compromissos (Forney e outros [18], Calderbank e Sloane [5], Wei [36], Forney [13]). A combinação de códigos em bloco e códigos em treliça com constelações multidimensionais também tem sido explorada (Tanner [32], Calderbank [4]).

Em Forney e outros ([18]) foi indicado, pela primeira vez, que todos os esquemas então conhecidos poderiam ser descritos por uma mesma forma de construção, onde uma parte dos símbolos de informação é codificada por um codificador binário cuja saída codificada é usada para selecionar um subconjnto em uma partição da constelação utilizada e o restante dos símbolos de informação são usados para selecionar um sinal nesse subconjunto para transmissão pelo canal, evidenciando que essas duas fases poderiam ser ajustadas de forma bastante desacoplada. Calderbank e Sloane ([5]) observaram que muitos de tais esquemas de codificação poderiam ser obtidos utilizando constelações que fossem constituídas por pontos de reticulados multidimensionais, e cujas partições correspondessem a partições desses reticulados por correspondentes sub-reticulados. Forney ([13]), então, explicitou o desacoplamento acima para essas construções baseadas em reticulados, criando o conceito de código de classes laterais, que depende exclusivamente do codificador binário e da partição do reticulado utilizado, e que é responsável pela maior parte do desempenho e da complexidade dos esquemas de codificação baseados em reticulados.

Nesse trabalho são tratados aspectos teóricos e aplicados de reticulados e suas partições, em sua utilização nesses esquemas de codificação baseados em reticulados para o canal AWGN limitado em banda.

1.2 Descrição do trabalho

Inicialmente, no Capítulo 2, estabelecemos a terminologia e a notação adotadas neste trabalho, revisando conceitos sobre reticulados e suas partições com a utilização de conceitos correspondentes sobre grupos abelianos e suas partições.

No Capítulo 3 é tratado um aspecto teórico da partição de reticulados, especificamente a sua estrutura algébrica, que, ao lado de sua importância teórica, pode ter aplicações relacionadas à construção de reticulados e de códigos de classes laterais.

No Capítulo 4 são analisados os parâmetros para avaliação e comparação de codificadores para o canal AWGN limitado em banda, sendo então relacionados aos parâmetros dos reticulados e suas partições utilizados em esquemas de construção, baseados em reticulados, desses codificadores.

No Capítulo 5 são descritas as formas gerais de construção de reticulados, e uma construção multinível é então proposta, sendo dados alguns exemplos de reticulados, novos e conhecidos, que podem ser obtidos, sendo determinados seus parâmetros de desempenho.

No Capítulo 6 são descritos os métodos gerais de decodificação ótima de reticulados, e um algoritmo de decodificação sub-ótima dos reticulados obtidos no capítulo anterior é então proposto, sendo avaliadas a degradação decorrente em seus parâmetros de desempenho e a redução de complexidade obtida, em relação a algorítmos eficientes de decodificação ótimas conhecidos.

Finalmente, no Capítulo 7, são comentados os resultados obtidos, apontando-se as possíveis extensões desse trabalho para futuras investigações.

Capítulo 2

Reticulados e suas partições

2.1 Introdução

Este capítulo introduz os conceitos básicos sobre reticulados e suas partições necessários ao estudo dos próximos capítulos. Inicialmente são revistos, na Seção 2.2, alguns conceitos elementares sobre grupos abelianos, para então, nas Seções 2.3 e 2.4, serem introduzidas as caracterizações algébrica e geométrica dos reticulados e suas partições.

2.2 Conceitos elementares sobre grupos abelianos

Nesta breve revisão de conceitos elementares sobre grupos abelianos, são abordados aspectos algébricos e geométricos, estabelecendo-se a notação e a terminologia gerais adotadas neste trabalho.

Um **grupo abeliano** ([19],[31]) é um conjunto G sobre o qual é definida uma operação (denotada+) comutativa, associativa, com elemento identidade único (denotado 0) e com elemento simétrico único para cada elemento $g \in G$ (denotado (-g)). Um grupo abeliano é dito ser **finito** se possuir um número finito de elementos, denominado a **ordem** do grupo abeliano. Um grupo abeliano H é um **subgrupo** de um grupo abeliano G (denotando-se $H \leq G$) se $H \subseteq G$ e se a operação em H é a operação em G restrita a H. Todo subgrupo de um grupo abeliano finito é, obviamente, finito.

Um subgrupo H de um grupo abeliano G induz naturalmente uma **partição** de G: dois elementos $g,g' \in G$ são ditos congruentes módulo H (denotando-se $g \equiv g'$ (MOD H)) se $(g + (-g')) \in H$; esta relação é claramente uma relação de equivalência, ou seja, reflexiva, simétrica e transitiva, e particiona G em classes de equivalência disjuntas, chamadas **classes laterais** de H contidas em G; denota-se esta partição por G/H. Se tomarmos um elemento $g \in G$, qualquer elemento $g' \in G$, pertencente à mesma classe lateral de $H \leq G$ a qual g pertence, pode ser decomposto de forma única em $g' = g + h$, com $h \in H$ único, especificamente $h = g' + (-g)$. Assim, a classe lateral de $H \leq G$, a qual $g \in G$ pertence (denotada (g+H)), pode ser descrita como $(g+H) = \{g + h : h \in H\}$. Assim, todo subgrupo de um grupo abeliano finito, bem como suas classes laterais, tem ordem igual a um divisor da ordem deste grupo abeliano, ou seja, se G é finito então a ordem de G é igual à ordem de um subgrupo $H \leq G$ multiplicada pela ordem da partição G/H (isto é, o número de classes laterais de H na partição G/H, necessariamente finito para G finito). Note-se que H é a única classe lateral em G/H que possui $0 \in G$, o elemento identidade de G.

Define-se a **soma de subconjuntos** $R \subseteq G$ e $S \subseteq G$ de um grupo abeliano G como $R + S \triangleq \{x + s : s \in S, r \in R\}$, onde a soma r + s (com $r,s \in G$), é realizada em G. Essa operação é, obviamente, comutativa e associativa; Além disso, a soma de subgrupos de um grupo abeliano G é também um subgrupo de G, ou seja, $H_1 + \ldots + H_n \leq G$, para $H_i \leq G$, i=1,...,n.

Verifica-se facilmente que a partição G/H, de um grupo abeliano G em classes laterais de um subgrupo $H \leq G$, é um grupo abeliano, denominado **grupo quociente**, sob a operação de soma de subconjuntos de G definida acima, visto que G/H é fechado em relação a esta operação

(comutativa e associativa), com H como elemento identidade, e ((-g)+H) como o elemento simétrico de (g+H), para qualquer $g \in G$.

Um **isomorfismo** entre dois grupos abelianos é uma correspondência biunívoca entre eles que preserva somas. Por exemplo, o produto cartesiano $H_1 \times \ldots \times H_n$ de subgrupos de um grupo abeliano G, definido naturalmente como um grupo abeliano sob adição componente-por-componente, é isomórfico à soma $H_1 + \ldots + H_n$, sempre que a interseção entre cada H_i e a soma dos demais for o subgrupo trivial de G constituído apenas pelo elemento identidade; um isomorfismo pode ser definido, neste caso, simplesmente associando-se a cada $(h_1,\ldots,h_n) \in H_1 \times \ldots \times H_n$ o elemento $h \in H_1 + \ldots + H_n$ dado por $h = h_1 + \ldots + h_n$.

É conveniente, por vezes, rotular cada classe lateral em uma partição G/H por um único elemento, denominado **representante**, selecionado daquela classe, formando um **conjunto de representantes** para as classes laterais da partição G/H, denotado [G/H]. Exceto no caso trivial em que $H \leq G$ é constituído exclusivamente pelo elemento identidade de G, cada classe lateral terá mais que um elemento, tornando não-única a seleção de [G/H]. O elemento identidade é sempre selecionado como representante de H na partição G/H. Uma vez selecionado um conjunto de representantes [G/H], no entanto, cada classe lateral em G/H pode ser descrita como (c+H), para algum $c \in [G/H]$ único, de modo que podemos decompor G na soma $G = [G/H] + H$, chamada **decomposição em classes laterais**.

Como um conjunto de representantes [G/H] das classes laterais de uma partição G/H é um subconjunto de G, representantes podem ser somados em G, não produzindo necessariamente, no entanto, representantes como resultado desta soma. Todavia, a soma em G de representantes em [G/H] é um elemento de G, e logo pertence a uma e apenas uma classe lateral em G/H, que possui, por sua vez, um e apenas um representante em [G/H]; isto define uma operação em [G/H], que denominaremos de **soma de representantes módulo H**. Denotaremos a soma módulo H de dois representantes $c,c' \in [G/H]$ por $c \dot{+} c'$.

Uma vez selecionado um conjunto de representantes [G/H], denotaremos por $gMOD_{[G/H]}H$ o representante da classe à qual g pertence, para $g \in G$. Assim, tem-se que $gMOD_{[G/H]}H \dot{+} g'MOD_{[G/H]}H = (g + g')MOD_{[G/H]}H$, para $g,g' \in G$, e que $g \equiv g'$ (MOD H) se e somente se $gMOD_{[G/H]}H = g'MOD_{[G/H]}H$, para $g,g' \in G$. Em particular, para $c,c' \in [G/H]$, tem-se $cMOD_{[G/H]}H = c$ e $c'MOD_{[G/H]}H = c'$. E assim, $c \dot{+} c' = (c + c')MOD_{[G/H]}H$, precisamente a definição de $c \dot{+} c'$.

Verifica-se facilmente que um conjunto [G/H], de representantes das classes laterais de uma partição G/H de um grupo abeliano G por um subgrupo H, é um grupo abeliano, sob a operação de soma definida acima em [G/H], sendo claramente isomórfico ao grupo quociente G/H, com isomorfismo $\Phi : G/H \rightarrow [G/H]$ dado por $\Phi(g + H) = gMOD_{[G/H]}H$, para qualquer $g \in G$.

Em um grupo abeliano G, pode-se definir os **múltiplos** K.g, com $K \in Z$, de um elemento $g \in G$, recursivamente, por: (1) $K.g = 0$ se $K = 0$; (2) $K.g = g + (K - 1).g$ se $K > 0$; e (3) $K.g = (-g) + (K + 1).g$ se $K < 0$. Um elemento $g \in G$ distinto da identidade é dito ter **ordem** K se K for o menor inteiro positivo não nulo tal que K.g é igual à identidade; se K.g for distinto

da identidade para todo K positivo não-nulo, g é dito ter ordem infinita; a identidade é dita ter ordem zero, por convenção. Todos os elementos de um grupo abeliano finito tem ordem finita que divide a ordem do grupo.

Um tipo particularmente interessante de subgrupo de um grupo abeliano G, denominado subgrupo **finitamente gerado**, é aquele obtido pelas combinações lineares com coeficientes inteiros dos elementos de um subconjunto finito de G, denominado **conjunto gerador**. Por conveniência, todos os elementos de um conjunto gerador podem ser assumidos distintos da identidade, isto é, sendo $\{g_1,...,g_n\} \subseteq G$, tem-se que $K_1.g_1 + ... + K_n.g_n : K_i \in \mathbb{Z}$, $i = 1,...,n$, é um subgrupo de G, que constitui o subgrupo finitamente gerado pelo conjunto gerador $\{g_1,...,g_n\}$. Se G for finito, todo subgrupo de G é finitamente gerado. Para cada grupo abeliano G cujos elementos não-nulos tenham ordem infinita, é sempre possível e conveniente selecionar, para qualquer subgrupo finitamente gerado, um conjunto gerador $\{g_1,...,g_n\} \subseteq G$ que satisfaça: $K_1.g_1 + ... + K_n.g_n$ é igual à identidade, com $K_i \in \mathbb{Z}$, $i = 1,...,n$, se e somente se $K_i = 0$, $i = 1,...,n$.

Uma **medida de distância** entre elementos de um grupo abeliano G é uma função $d : G \times G \rightarrow \mathbb{R}_+$, definida em $G \times G$ com valores reais positivos, tal que, para quaisquer $g_1,g_2 \in G$, tem-se: (1) $d(g_1,g_2) = 0$ se $g_1 = g_2$; (2) $d(g_1,g_2) > 0$ se $g_1 \neq g_2$; e (3) $d(g_1,g_2) = d(g_2,g_1)$. Note-se que d não é necessariamente uma métrica, pois não é exigido que seja satisfeita a desigualdade triangular: $d(g_1,g_3) \leq d(g_1,g_2) + d(g_2,g_3)$, para quaisquer $g_1,g_2,g_3 \in G$. Define-se o **peso** de um elemento $g \in G$, com base em uma medida de distância d já estabelecida para G, como sendo a medida de distância entre g e o elemento identidade do grupo G.

Uma medida de distância sobre o produto cartesiano de grupos $G_1 \times ... \times G_n$ pode sempre ser definida naturalmente como a soma das medidas de distância das componentes; essa medida de distância é denominada **aditiva**. Uma medida de distância sobre um grupo G é dita ser **invariante por translação** se, para quaisquer $g_1,g_2,g_3 \in G$, tem-se $d(g_1,g_2) = d(g_1 + g_3,g_2 + g_3)$. Para uma medida de distância d que seja invariante por translação, a medida da distância entre dois elementos $g,g' \in G$ é igual ao peso do elemento $(g + (-g')) \in G$.

Dado um grupo G, munido de uma medida de distância d, podemos definir, para cada elemento $g \in G$, o **espectro de distâncias** a partir de g como sendo o conjunto $S_G^g \triangleq \{d(g',g) : g' \in G, g' \neq g\} \subseteq \mathbb{R}_+^*$, e a **distribuição de distâncias** a partir de g como sendo a função $N_G^g : S_G^g \rightarrow \mathbb{N}^*$, dada, para $\delta \in S_G^g$, por $N_G^g(\delta) \triangleq |K_G^g(\delta)|$, onde $K_G^g(\delta)$ é a **camada** de raio δ em torno de $g \in G$, definida por $K_G^g(\delta) \triangleq \{g' \in G : d(g',g) = \delta\} \subseteq G$, e $|.|$ indica cardinalidade de conjunto. Obviamente, se a medida de distância d do grupo G for invariante por translação, então S_G^g e N_G^g não dependem de $g \in G$, podendo ser denotados simplesmente por S_G e N_G e denominados o **espectro de pesos** e a **distribuição de pesos**, respectivamente; além disso, também neste caso, $K_G^g(\delta) = |g| + K_G^0(\delta)$, com $\delta \in S_G$, onde $K_G^0(\delta)$ é a camada de raio δ em torno da identidade; assim, visto de qualquer um de seus pontos, G tem sempre a mesma aparência.

Um grupo G, munido de uma medida de distância d, é dito ser um **grupo discreto** se existir $d_{MIN}(G) \triangleq MIN\{d(g,g') : g,g' \in G,\ g \neq g'\}$, denominado a **distância mínima** de G; neste caso, para todo $g \in G$, S_G^g também possui um mínimo, que denotaremos $d_{MIN}^g(G)$, e então $d_{MIN}(G) = MIN\{d_{MIN}^g(G) : g \in G\}$; denomina-se, neste caso, o valor de $N_G^g(d_{MIN}^g(G))$ por **número de vizinhos mais próximos** de $g \in G$. No caso de d ser invariante por translação, $d_{MIN}^g(G) = d_{MIN}(G)$, para todo $g \in G$, e $d_{MIN}(G)$ pode ser denominado o **peso mínimo** de G; denomina-se, neste caso, o valor de $N_G(d_{MIN}(G))$ por **número de vizinhos mais próximos** em G. Note-se que todo grupo finito é discreto.

Para uma classe lateral qualquer (c+H), com $c \in [G/H]$, de uma partição G/H, aplicam-se as definições e resultados acima, com a classe lateral (c+H) em lugar do grupo abeliano G, chegando-se aos conceitos de S_{c+H}^{c+h}, N_{c+H}^{c+h}, K_{c+H}^{c+h}, $d_{MIN}^{c+h}(c+H)$ e $d_{MIN}(c+H)$, para $h \in H$. Em particular, para uma medida de distância invariante por translação, S_{c+H}^{c+h}, N_{c+H}^{c+h}, $d_{MIN}^{c+h}(c+H)$ e $d_{MIN}(c+H)$ independem de $c \in [G/H]$ e $h \in H$, sendo iguais a S_H, N_H, $d_{MIN}(H)$ e $d_{MIN}(H)$, respectivamente; além disso, $K_{c+H}^{c+h}(\delta) = |c| + K_H^h(\delta) = |c+h| + K_H^0(\delta)$, com $\delta \in S_H$, para quaisquer $c \in [G/H]$ e $h \in H$. Em qualquer caso, no entanto, estes conceitos independem do particular [G/H] selecionado para a partição G/H. Denominaremos $d_{MIN}(c+H)$ por **distância intraclasse** de (c+H), e definimos a **distância intraclasse mínima** de uma partição G/H como sendo o mínimo das distâncias intraclasse das classes de G/H, denotando-a por $d_{MIN}[G/H]$. Novamente, para uma medida de distância d que seja invariante por translação, teremos $d_{MIN}[G/H] = d_{MIN}(H)$, o peso mínimo de H.

A **distância interclasses**, para duas classes laterais, (c+H) e (c'+H), distintas em uma partição G/H de um grupo abeliano G, é definida como o mínimo, caso exista, do conjunto $\{d(g,g') : g \in (c+H),\ g' \in (c'+H)\}$, sendo denotada por $d_{MIN}(c+H,c'+H)$, e assim definimos a **distância interclasses mínima** de uma partição G/H como sendo o mínimo das distâncias interclasses dos pares de classes laterais distintas de G/H, denotando-a por $d_{MIN}(G/H)$. Para uma medida de distância d invariante por translação, temos que $d_{MIN}(c+H, c'+H) = d_{MIN}((c - c')+H,H)$, denominado o **peso de classe** de $((c - c')+H)$, e assim $d_{MIN}(G/H)$ por ser denominado o **peso de classe mínimo** da partição G/H, neste caso.

A distância mínima de um grupo discreto G, munido de uma medida de distância, pode ser determinada, em termos de uma partição qualquer G/H, por $d_{MIN}(G) = MIN\{d_{MIN}[G/H], d_{MIN}(G/H)\}$, ou seja, como o mínimo entre a distância intraclasse mínima de G/H e a distância interclasses mínima de G/H; assim, se existem os mínimos expressos por $d_{MIN}[G/H]$ e $d_{MIN}(G/H)$ para algum $H \leq G$, então G é discreto; reciprocamente, se G é discreto, então, para todo $H \leq G$, existem os mínimos expressos por $d_{MIN}[G/H]$ e $d_{MIN}(G/H)$. Quando a medida de distância utilizada é invariante por translação, enunciamos esse relacionamento por $d_{MIN}(G) = MIN\{d_{MIN}(H), d_{MIN}(G/H)\}$, ou seja, o peso mínimo de G é igual ao mínimo entre o peso mínimo de H e o peso de classe mínimo de G/H. A Figura 2.2.1

ilustra esquematicamente as várias distâncias mínimas associadas com uma partição de grupo G/H.

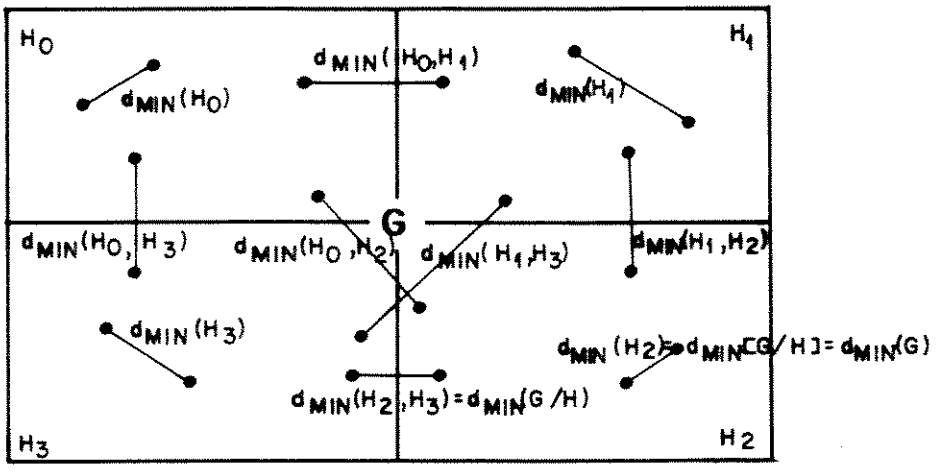

Fig. 2.2.1 - Distâncias mínimas em uma partição $G/H = \{H_0, H_1, H_2, H_3\}$ onde $H = H_0$ é um subgrupo do grupo abeliano G.

Um importante exemplo de grupos abelianos é apresentado a seguir com o propósito de ilustrar alguns dos conceitos e resultados elementares sobre grupos abelianos revisados acima.

O conjunto $GF(q)^n$, das n-uplas de elementos de um corpo finito GF(q), formam um espaço vetorial, sobre GF(q), sob adição componente-por-componente em GF(q) e multiplicação de componentes por escalar em GF(q). Um **código de bloco linear** ([24],[25]) de comprimento n, sobre GF(q), é um subespaço do espaço vetorial $GF(q)^n$. Um **código de bloco aditivo** ([3],[12]) de comprimento n, sobre GF(q), é um subgrupo do grupo aditivo do espaço vetorial $GF(q)^n$. Todo código de bloco linear sobre GF(q) é, obviamente, também código de bloco aditivo sobre GF(q). Além disso, se q for primo, então todo código de bloco aditivo sobre GF(q) é também código de bloco linear; neste caso costuma-se utilizar a denominação de **código de grupo** ([7]). Em capítulos posteriores, faremos uso de códigos de bloco aditivos; no entanto, os exemplos apresentados serão sobre GF(2), reduzindo-os aos códigos de bloco lineares, sobre os quais já se possui uma imensa quantidade de conhecimento acumulada na literatura.

Sendo $GF(q)^n$ um espaço vetorial finito sobre o corpo finito GF(q), todo código de bloco linear de comprimento n sobre GF(q) é também finito, assim como todo código de bloco aditivo de comprimento n sobre GF(q), ambos com número de elementos igual a um divisor de q^n. Mais especificamente, seja $q=p^b$, onde p é primo e $b > 1$ é inteiro. Seja C um subespaço de dimensão $K \leq n$ do espaço vetorial $GF(q)^n$, tendo então p^{bK} elementos; assim, C é um código linear de comprimento n e dimensão K sobre GF(q), e então possui uma base vetorial, com K vetores linearmente independentes cujas combinações lineares com coeficientes em GF(q) são todos os elementos de C. No entanto, C é também um subgrupo do grupo aditivo do espaço vetorial $GF(q)^n$, sendo assim um código de bloco aditivo de comprimento n sobre GF(q). As combinações lineares dos elementos de um conjunto gerador para C devem ser com coeficientes inteiros; porém, visto que GF(q) tem característica p, bastará restringi-los aos inteiros 0, 1,..., p-1; assim, um conjunto gerador para C deve ter pelo menos bK elementos. Um conjunto gerador

para C pode ser obtido a partir de uma base vetorial para C multiplicando-se cada vetor da base pelas potências $\alpha^0,\alpha^1,...,\alpha^{b-1}$ de uma raiz α em GF(q), de um polinômio irredutível de grau b sobre GF(p), obtendo-se os Kb elementos de um conjunto gerador para C. Isto decorre do fato de que qualquer elemento de $GF(p^b)$ pode ser expresso como um polinômio em α, sobre GF(p), de grau menor ou igual a b-1. No caso em que b=1 (ou seja, q é primo), toda base vetorial, para C como código de bloco linear, é, obviamente, também um conjunto gerador, para C como código de bloco aditivo, valendo a recíproca.

A medida de distância adotada em $GF(q)^n$ é a **distância de Hamming**, denotada d_H e definida em GF(q) por: (1) $d_H(c,c') = 0$ se $c = c'$, e (2) $d_H(c,c') = 1$ se $c \neq c'$, para $c,c' \in GF(q)$; em $GF(q)^n$, define-se a distância de Hamming aditivamente, ou seja: $d_H(c,c') = \sum_{i=1}^{n} d_H(c_i,c'_i)$, onde $c = (c_1,...,c_n)$, $c' = (c'_1,...,c'_n)$, para $c,c' \in GF(q)^n$. Essa medida de distância é, claramente, invariante por translação; assim, ficam definidos o espectro de pesos, a distribuição de pesos e o peso mínimo de um código de bloco aditivo, em geral, e de um código de bloco linear, em particular, já que $GF(q)^n$ é finito.

Encerramos, com este exemplo, essa breve revisão sobre grupos abelianos. As duas seções seguintes abordam as caracterizações algébrica e geométrica dos reticulados, em termos dos conceitos revisados nesta seção.

2.3 Caracterização algébrica dos reticulados e suas partições

Iniciamos, nesta seção, o estudo dos reticulados, abordando suas características algébricas, bem como de suas partições. Neste estudo, bem como ao longo deste trabalho, nos restringiremos aos reticulados reais de dimensão finita, que são os mais úteis para as aplicações descritas em capítulos posteriores.

Comecemos pelo universo de onde são extraídos estes reticulados: o espaço vetorial $\mathbb{R}^n$. O conjunto $\mathbb{R}^n$ das n-uplas reais formam um espaço vetorial, sob adição componente-por-componente em $\mathbb{R}$ e multiplicação de componentes por escalares reais. Um subespaço vetorial do $\mathbb{R}^n$ é um subconjunto do $\mathbb{R}^n$ que forma um espaço vetorial sob as mesmas operações de adição e multiplicação por escalar do $\mathbb{R}^n$ restritas aos seus elementos; o conjunto formado pelas combinações lineares com coeficientes reais de um conjunto $\{w_1,...,w_n\}$ de vetores do $\mathbb{R}^n$ é um subespaço do $\mathbb{R}^n$, sendo o conjunto $\{w_1,...,w_n\}$ chamado um conjunto gerador para o subespaço que gera. Um conjunto de vetores do $\mathbb{R}^n$ é dito linearmente independente se nenhuma combinação linear desses vetores, com coeficientes reais nem todos nulos, é o vetor nulo (a origem do $\mathbb{R}^n$). Uma base para um subespaço do $\mathbb{R}^n$ é um conjunto gerador para este subespaço, constituído de vetores linearmente independentes; duas bases para um mesmo subespaço $\mathbb{R}^n$ tem o mesmo número de vetores, chamado a dimensão deste subespaço; o espaço vetorial $\mathbb{R}^n$ tem dimensão n, e qualquer subespaço do $\mathbb{R}^n$ tem dimensão menor ou igual a n. Um operador linear não singular no $\mathbb{R}^n$ é uma função inversível, do $\mathbb{R}^n$ no $\mathbb{R}^n$, que preserva somas e multiplicação por escalares, e assim transforma bases em bases para o $\mathbb{R}^n$, ou seja, preserva

independência linear no $\mathbb{R}^n$. Estas considerações algébricas sobre o espaço vetorial $\mathbb{R}^n$ são suficientes para caracterizarmos algebricamente os reticulados do $\mathbb{R}^n$.

Um **reticulado** Λ no $\mathbb{R}^n$ ([27],[6],[10]) é um subgrupo do grupo aditivo do $\mathbb{R}^n$, finitamente gerado, com conjunto gerador constituído por vetores linearmente independentes do $\mathbb{R}^n$, ou seja, deve existir um conjunto de $m \leq n$ vetores $\{w_1,...,w_m\}$ linearmente independentes do $\mathbb{R}^n$, chamado uma **base** para Λ, de modo que

$$\Lambda = \{\lambda \in \mathbb{R}^n : \lambda = k_1 w_1 + \ldots + k_m w_m,\ k_i \in \mathbb{Z},\ i = 1,...,m\} \qquad (2.3.1)$$

Evidentemente, uma base para Λ é também uma base vetorial para um subespaço do $\mathbb{R}^n$; a dimensão $m \leq n$ deste subespaço é dita ser também a dimensão do reticulado Λ. Podemos, obviamente, restringir nossa atenção, sem perda essencial de generalidade, apenas aos reticulados do $\mathbb{R}^n$ cuja dimensão m seja igual a n; é o que faremos neste capítulo, bem como ao longo de todo este trabalho. Uma matriz M quadrada n x n, cujas colunas sejam os n vetores de uma base para um reticulado Λ do $\mathbb{R}^n$, é chamada uma **matriz geradora** de Λ, de modo que a equação (2.3.1) pode ser reescrita como:

$$\Lambda = M.\mathbb{Z}^n \triangleq \{M.z : z \in \mathbb{Z}^n\} \qquad (2.3.2)$$

Como qualquer subgrupo do grupo aditivo do $\mathbb{R}^n$, um reticulado Λ é um conjunto infinito de vetores do $\mathbb{R}^n$, tal que a soma de dois vetores de Λ também é um vetor de Λ, a origem é o vetor identidade de Λ, e se $\lambda \in \Lambda$ então $(-\lambda) \in \Lambda$; é claro, então, que todos os múltiplos inteiros de um vetor de Λ são vetores distintos de Λ, e que a diferença entre dois vetores de Λ também é um vetor de Λ. A restrição de existir, para um reticulado, um conjunto gerador de vetores linearmente independentes do $\mathbb{R}^n$ é exatamente o que o diferencia dos demais subgrupos do grupo aditivo do $\mathbb{R}^n$; este aspecto é tratado novamente na próxima seção.

Como um exemplo, o conjunto $\mathbb{Z}^n$, de todas as n-uplas de inteiros, é um reticulado do $\mathbb{R}^n$, tendo a base vetorial canônica $\{e_1,...,e_n\}$ do $\mathbb{R}^n$ (onde e_i, $i = 1,...,n$, é o vetor com 1 na i^a coordenada e 0 nas demais coordenadas) como um conjunto gerador ou base. O reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$ está ilustrado na Figura 2.3.1.

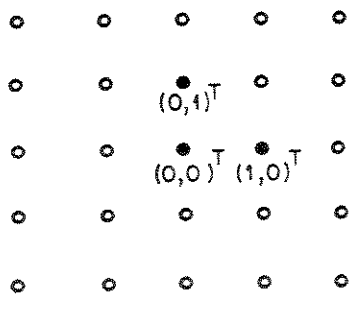

Fig. 2.3.1 - O reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$; foram ressaltadas a origem (0,0) e uma base para $\mathbb{Z}^2$, $\{(1,0)^T,(0,1)^T\}$

Um operador linear não-singular do $\mathbb{R}^n$, por preservar independência linear, transforma reticulados do $\mathbb{R}^n$ em reticulados do $\mathbb{R}^n$. Em particular, um operador linear não-singular T tal que $T\Lambda = \Lambda$, onde $T\Lambda \triangleq \{T(\lambda) : \lambda \in \Lambda\}$ é dito ser um **automorfismo** para Λ. Se Λ é um reticulado do $\mathbb{R}^n$, então Λ^m também é um reticulado do $\mathbb{R}^{nm}$.

Definido um reticulado Λ do $\mathbb{R}^n$, consideremos a partição $\mathbb{R}^n/\Lambda$. As classes laterais de Λ em $\mathbb{R}^n/\Lambda$ são translações de Λ, de forma que dois pontos do $\mathbb{R}^n$ pertencem à mesma classe lateral de Λ (ou seja, são congruentes módulo Λ) se sua diferença pertence a Λ. Assim, a classe lateral de Λ em $\mathbb{R}^n/\Lambda$ a qual $x \in \mathbb{R}^n$, ou ainda, a **versão transladada** de Λ determinada por x é dada por:

$$(x + \Lambda) = \{x + \lambda : \lambda \in \Lambda\} \tag{2.3.3}$$

Se selecionarmos um conjunto de representantes $[\mathbb{R}^n/\Lambda]$ das classes laterais de Λ em$\mathbb{R}^n/\Lambda$, teremos:

$$\mathbb{R}^n = [\mathbb{R}^n/\Lambda] + \Lambda \tag{2.3.4}$$

A Equação (2.3.4) tem duas interpretações alternativas igualmente importantes: (1) o $\mathbb{R}^n$ é a união de versões transladadas de Λ disjuntas, determinadas pelos vetores de $[\mathbb{R}^n/\Lambda]$; (2) o $\mathbb{R}^n$ é a união de versões transladadas de $[\mathbb{R}^n/\Lambda]$ disjuntas, determinadas pelos vetores de Λ. Claramente, existe um número infinito de classes laterais de Λ em $\mathbb{R}^n/\Lambda$, cuja união recobre o $\mathbb{R}^n$, com interseção vazia. Note-se que qualquer subconjunto de vetores do $\mathbb{R}^n$, que determinem versões transladadas de Λ disjuntas cuja união recobre o $\mathbb{R}^n$, é um possível conjunto de representantes $[\mathbb{R}^n/\Lambda]$ das classes laterais de Λ na partição $\mathbb{R}^n/\Lambda$, que denominaremos uma **região fundamental** de Λ e denotaremos por $R(\Lambda)$. Assim, com esta notação, a equação (2.3.4) pode ser reescrita como:

$$\mathbb{R}^n = R(\Lambda) + \Lambda \tag{2.3.5}$$

Diremos que Λ é um **sub-reticulado** de um reticulado Γ do $\mathbb{R}^n$, se Λ for um subgrupo de Γ; isto equivale a dizer que Λ é um reticulado do $\mathbb{R}^n$ e $\Lambda \subseteq \Gamma$. Indicaremos que Λ é sub-reticulado do reticulado Γ pela notação $\Lambda \leq \Gamma$, a mesma utilizada para subgrupos. Como já mencionado nesta seção, consideraremos apenas reticulados n-dimensionais do $\mathbb{R}^n$; embora um reticulado possua sub-reticulados em todas as dimensões menores ou iguais à sua, apenas consideraremos aqueles cuja dimensão seja igual à sua.

Um subreticulado Λ de um reticulado Γ do $\mathbb{R}^n$ dá origem a uma partição Γ/Λ de Γ em classes laterais de Λ, tal que se $[\Gamma/\Lambda]$ é um conjunto de representantes das classes laterais de Γ/Λ, então:

$$\Gamma = [\Gamma/\Lambda] + \Lambda \tag{2.3.6}$$

As classes laterais de Λ em Γ/Λ são, então, as classes laterais de Λ em $\mathbb{R}^n/\Lambda$ contidas em Γ. Na próxima seção é demonstrado que sempre o número de classes laterais de Λ em Γ/Λ é finito, desde que Λ e Γ tenham a mesma dimensão, como já assumido; este número é a

ordem do grupo quociente Γ/Λ, também chamado **índice** de Λ em Γ, e denotado $|\Gamma/\Lambda|$. Note que a equação (2.3.6) sugere uma maneira bastante geral de se construir novos reticulados, a partir de reticulados já conhecidos: a união de classes laterais de Λ em $\mathbb{R}^n/\Lambda$, criteriosamente selecionadas, resulta em outro reticulado Γ; esta é a técnica de construção mais comumente utilizada na busca de novos reticulados.

Como um exemplo, o conjunto D_n de todas as n-uplas de inteiros com um número par de coordenadas ímpares, é um sub-reticulado do reticulado $\mathbb{Z}^n$, gerando uma partição $\mathbb{Z}^n/D_n$ de ordem 2, cujos elementos são as classes laterais D_n e $((1,0,...,0)^T + D_n)$; assim uma possível seleção para o conjunto de representantes das classes laterais de D_n em Z^n/D_n seria $[Z^n/D_n] = \{(0,0,...,0)^T,(1,0,...,0)^T\}$. O reticulado D_2 do $\mathbb{R}^n$ está ilustrado na Figura 2.3.2, juntamente com o reticulado $\mathbb{Z}^2$, já ilustrado anteriormente.

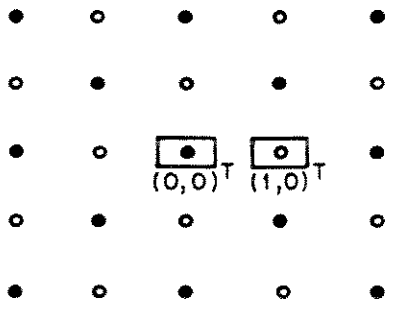

Fig. 2.3.2 - O reticulado D_2 do $\mathbb{R}^2$ (círculos cheios), subreticulado do $\mathbb{Z}^2$ (círculos cheios e vazios); dois possíveis representantes das classes laterais em $\mathbb{Z}^2/D_2$ estão indicados (quadrados).

Pode-se utilizar operadores lineares não-singulares para obter-se sub-reticulados de um dado reticulado. Um operador linear não-singular T do $\mathbb{R}^n$ tal que $T\Lambda \leq \Lambda$ é dito ser um **endomorfismo** para Λ.

Se N é uma matriz geradora de um reticulado Γ e M é uma matriz geradora de um reticulado $\Lambda \leq \Gamma$, então obviamente M = N.P, onde P é uma matriz quadrada de números inteiros, que denominamos uma **matriz particionadora** para a partição Γ/Λ.

Se Γ/Λ é uma partição de reticulados do $\mathbb{R}^n$, então Γ^m/Λ^m é uma partição de reticulados do $\mathbb{R}^{mn}$, e $\Gamma^m/\Lambda^m = (\Gamma/\Lambda)^m$. (Rigorosamente, Γ^m/Λ^m e $(\Gamma/\Lambda)^m$ são isomórficos, e não iguais).

Encerramos, assim, esta breve caracterização algébrica dos reticulados e suas partições. Passemos, então, a uma caracterização geométrica dos reticulados e suas partições.

2.4 Caracterização geométrica dos reticulados e suas partições

Novamente, iniciaremos pelo $\mathbb{R}^n$. A medida de distância que adotaremos para o $\mathbb{R}^n$ é a **distância Euclideana quadrática**, denotada d^2 e definida em $\mathbb{R}$ por $d^2(x,y) = (x-y)^2$, para

$x, y \in \mathbb{R}^n$; em $\mathbb{R}^n$ defini-se a distância Euclideana quadrática aditivamente, ou seja: $d^2(x,y) = \sum_{i=1}^{n} d^2(x_i, y_i)$, onde $x = (x_1,...,x_n)$, $y = (y_1,...,y_n)$, para $x, y \in \mathbb{R}^n$; essa medida de distância é, obviamente, invariante por translação. O **peso Euclideano quadrático** de um vetor $x \in \mathbb{R}^n$, i.e., sua distância quadrática à origem, será denotado $\|x\|^2$; assim, para $x, y \in \mathbb{R}^n$, $d^2(x,y) = \|x-y\|^2$. Embora d^2 não satisfaça a desigualdade triangular, $d_E \triangleq (d^2)^{1/2}$ satisfaz: $d_E(x,y) \leq d_E(x,z) + d_E(z,y)$, para $x, y, z \in \mathbb{R}^n$; d_E é a **distância Euclideana**, naturalmente; não utilizamos d_E como medida de distância, pois, embora seja inclusive uma métrica, não é aditiva; quando necessário, utilizaremos a notação $\|x\|$, para $x \in \mathbb{R}^n$, como a distância Euclideana de x à origem, significando, obviamente, $\|x\| \triangleq (\|x\|^2)^{1/2}$, ou seja, seu **peso Euclideano**. Um operador linear não-singular do $\mathbb{R}^n$ é dito uma **isometria** se preserva distâncias; operadores ortogonais são isometrias. Define-se o **produto escalar Euclideano** no $\mathbb{R}^n$ por $\langle x,y \rangle \triangleq \sum_{i=1}^{n} x_i.y_i$, onde $x = (x_1,...,x_n)$; $y = (y_1,...,y_n)$, para $x, y \in \mathbb{R}^n$; o **ângulo** entre dois vetores $x,y \in \mathbb{R}^n$ não-nulos é definido por $\hat{xy} \triangleq \cos^{-1}(\langle x,y \rangle / \|x\|.\|y\|)$. Um operador linear não-singular do $\mathbb{R}^n$ é dito uma **similaridade** se preserva ângulos; operadores ortogonais multiplicados por escalares reais não-nulos são similaridades. Toda isometria é uma similaridade, obviamente; além disso, a similaridade αT, onde T é um operador ortogonal e α é um escalar real não-nulo, multiplica distâncias quadráticas por $|\alpha|^2$. Duas regiões são **isométricas** (resp. **semelhantes**) se uma pode ser obtida a partir da outra pela aplicação de uma isometria (resp. similaridade) seguida por uma translação. No $\mathbb{R}^n$ define-se **volume** de uma região em relação ao hipercubo cuja aresta tem comprimento unitário; o volume da união de regiões disjuntas no $\mathbb{R}^n$ é igual a soma dos volumes dessas regiões; duas regiões isométricas tem o mesmo volume; duas regiões semelhantes R_1 e R_2, onde R_1 foi obtida de R_2 pela aplicação da similaridade αT (onde α é um escalar real não-nulo e T é um operador ortogonal) seguida por uma translação, tem volumes relacionados por $V(R_1) = |\alpha|^n.V(R_2)$, onde n é a dimensão do espaço. O produto cartesiano de uma região R por ela mesma m vezes tem volume $V(R^m) = V(R)^m$. Estas considerações geométricas sobre o espaço vetorial $\mathbb{R}^n$, agora então **espaço Euclideano** $\mathbb{R}^n$, é suficiente para caracterizarmos geometricamente os reticulados do $\mathbb{R}^n$.

Uma definição de reticulados do $\mathbb{R}^n$ frequentemente adotado na literatura é a seguinte: um reticulado do $\mathbb{R}^n$ é um subgrupo discreto do $\mathbb{R}^n$, segundo a distância Euclideana quadrática. É interessante, então, salientar a equivalência entre esta definição e aquela da seção anterior; ou seja, sendo Λ um subgrupo do grupo aditivo do $\mathbb{R}^n$, Λ é discreto se e somente se Λ possui um conjunto gerador constituído de vetores linearmente independentes do $\mathbb{R}^n$; essa equivalência de fato existe, podendo ser encontrada na literatura ([6]).

Sendo um reticulado Λ do $\mathbb{R}^n$ um subgrupo discreto do grupo aditivo do $\mathbb{R}^n$, sob a medida de distância Euclideana quadrática, e sendo essa medida invariante por translação, ficam definidos a **distância Euclideana quadrática mínima** de Λ (ou ainda, o **peso Euclideano quadrático mínimo não-nulo** de Λ), denotada $d^2_{MIN}(\Lambda)$, seu espectro de pesos S_Λ e sua distribuição de pesos N_Λ. As camadas em torno de um vetor $\lambda \in \Lambda$ são as camadas em torno da origem transladadas pelo vetor λ. Denotaremos o **número de vizinhos mais próximos** $N_\Lambda(d^2_{MIN}(\Lambda))$ em Λ por $M_0(\Lambda)$.

Como um exemplo, o reticulado $\mathbb{Z}^n$ do $\mathbb{R}^n$ tem distância quadrática mínima $d^2_{MIN}(\mathbb{Z}^n) = 1$, para todo n; seu espectro de pesos $S_{\mathbb{Z}^n}$ é constituído pelos inteiros δ que podem ser representados como a soma de n quadrados de inteiros, e, para cada tal inteiro $\delta \in S_{\mathbb{Z}^n}$, $N_{\mathbb{Z}^n}(\delta)$ é o número de maneiras distintas de se representar δ desta forma; o número de vizinhos mais próximos em $\mathbb{Z}^n$ é $M_0(\mathbb{Z}^n) = 2n$. As primeiras camadas do reticulado $\mathbb{Z}^2$ estão ilustradas na Figura 2.4.1.

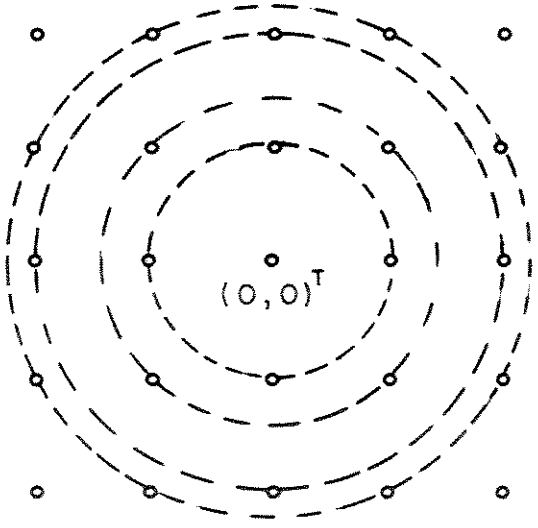

Fig. 2.4.1 - As quatro primeiras camadas do reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$, onde se vê $d^2_{MIN}(\mathbb{Z}^2) = 1$ e $M_0(\mathbb{Z}^2) = 4$.

O conjunto dos pontos do $\mathbb{R}^n$ mais próximos de um vetor λ de um reticulado Λ do $\mathbb{R}^n$ é chamada a **região de Voronói ou Dirichlet** de Λ em torno do ponto $\lambda \in \Lambda$, denotada $V_\Lambda(\lambda)$. Devido à invariância por translação da distância Euclideana quadrática, $V_\Lambda(\lambda) = \lambda + V_\Lambda(0) \triangleq \{\lambda + x : x \in V_\Lambda(0)\}$; assim, as regiões de Voronói de Λ são isométricas entre si por simples translação. As regiões de Voronói em todos os pontos de um reticulado do $\mathbb{R}^n$ são disjuntas (exceto, naturalmente, pelas fronteiras) cuja união recobre o $\mathbb{R}^n$; diz-se então que as regiões de Voronói de um reticulado do $\mathbb{R}^n$ tesselam o $\mathbb{R}^n$. A região de Voronói na origem de um reticulado Λ do $\mathbb{R}^n$ pode ser expressa como a interseção entre todos os semi-espaços $\aleph(\lambda) \triangleq \{x \in \mathbb{R}^n : \langle x,\lambda \rangle \leq \frac{1}{2} \cdot \|\lambda\|^2\}$, onde $\lambda \in \Lambda$ e $\lambda \neq 0$; no entanto, nem todos os semi-espaços $\aleph(\lambda)$, $\lambda \in \Lambda$, são necessários para definir-se $V_\Lambda(0)$, mas apenas um número finito destes; assim,

$V_\Lambda(0) = \bigcap_{\lambda \in \Re_\Lambda} \aleph(\lambda)$, onde $\Re_\Lambda \subseteq \Lambda$ é finito, sendo seus elementos, então, chamados os **vetores relevantes** de Λ (na origem); obviamente, todos os vizinhos mais próximos da origem em Λ (elementos da 1ª camada $K_\Lambda(d^2_{MIN}(\Lambda))$ de Λ), chamados **vetores mínimos** de Λ, são relevantes, mas podem existir outros vetores relevantes de Λ, não mínimos. Cada face de $V_\Lambda(0)$ está, então, contida em um dos hiperplanos $\Im(\lambda) \triangleq \{x \in \mathbb{R}^n : \langle x,\lambda \rangle = \frac{1}{2}.\|\lambda\|^2\}$, onde $\lambda \in \Re_\Lambda$. Os vértices de $V_\Lambda(\lambda)$, com $\lambda \in \Lambda$, são denominados os **buracos** de Λ em torno de λ; os buracos de Λ, em torno de $\lambda \in \Lambda$, mais distantes de λ são denominados **buracos profundos**, e os demais **buracos rasos**.

Os volumes das regiões de Voronói, por estas serem congruentes, são iguais. Vimos, na seção anterior, que qualquer região do $\mathbb{R}^n$, cujas versões transladadas determinadas pelos vetores de um reticulado Λ do $\mathbb{R}^n$ fossem disjuntas e reunidas recobrissem o $\mathbb{R}^n$, poderia ser tomada como um conjunto de representantes $[\mathbb{R}^n/\Lambda]$ para a partição de grupo $\mathbb{R}^n/\Lambda$; concluímos, então, que $V_\Lambda(0)$ é uma possível região fundamental de Λ, chamada um **conjunto de representantes padrão** para $\mathbb{R}^n/\Lambda$, por possuir o elemento de peso Euclideano quadrático mínimo de cada classe de Λ em $\mathbb{R}^n/\Lambda$. Qualquer outra região fundamental de Λ teria, pelo exposto acima, o mesmo volume de $V_\Lambda(0)$; esse volume é chamado o **volume fundamental** de Λ. O cálculo deste volume pode ser realizado, porém, utilizando-se uma região fundamental mais simples do que a região de Voronói, como será visto posteriormente.

Como exemplo, o reticulado $\mathbb{Z}^n$ do $\mathbb{R}^n$ tem $V_{\mathbb{Z}^n}(0) = [-\frac{1}{2}, \frac{1}{2})^n$, o hipercubo de lados unitários paralelos aos eixos coordenados do $\mathbb{R}^n$ e centrado na origem, com volume fundamental $V(\mathbb{Z}^n) = 1$. A região de Voronói na origem do reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$ está ilustrada na Figura 2.4.2.

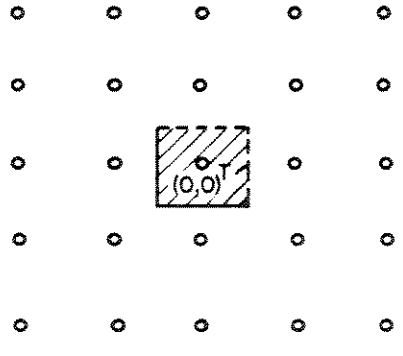

Fig. 2.4.2 - A região de Voronói na origem $V_{\mathbb{Z}^2}(0)$ do reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$ (região hachurada)

Pode-se inscrever, em cada região de Voronói de um reticulado Λ do $\mathbb{R}^n$, uma hiperesfera de raio $\rho(\Lambda) = \frac{1}{2}.(d^2_{MIN}(\Lambda))^{1/2}$, formando assim um **empacotamento esférico reticulado**, constituído de hiperesferas disjuntas de raio máximo com centros nos vetores de Λ. Portanto, a **densidade** deste empacotamento será a razão entre o volume de uma hiperesfera e o volume da

região de Voronói onde esta foi inscrita; sendo $V_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}-1\right)}$ o volume de uma hiperesfera do $\mathbb{R}^n$ de raio unitário, tem-se que o volume de uma hiperesfera de raio $\rho(\Lambda)$ será $V_n \cdot \rho(\Lambda)$; assim a densidade de Λ (i.e., do empacotamento esférico associado a Λ), será $\Delta(\Lambda) \triangleq V_n \cdot \frac{\rho(\Lambda)^n}{V(\Lambda)}$; o fator $\delta(\Lambda) \triangleq \frac{\rho(\Lambda)^n}{V(\Lambda)}$ é conhecido como **densidade de centros** de Λ. Tanto $\Delta(\Lambda)$ como $\delta(\Lambda)$ são invariantes por transformação de similaridade, sendo assim próprias para comparar a eficiência de empacotamento de reticulados de mesma dimensão; no entanto, não são invariantes por produto cartesiano (ou seja, $\Delta(\Lambda^m) \neq \Delta(\Lambda)$ e $\delta(\Lambda^m) \neq \delta(\Lambda)$, em geral para $m > 1$), não sendo, portanto, adequados para comparar a eficiência de empacotamento de reticulados de dimensões diferentes. Um fator adequado para isto é o chamado **parâmetro de Hermite** de Λ, definido por $\gamma(\Lambda) \triangleq 4 \cdot \delta(\Lambda)^{2/n} = \frac{d^2_{MIN}(\Lambda)}{V(\Lambda)^{2/n}}$ onde n é a dimensão de Λ; como $d^2_{MIN}(\Lambda^m) = d_{MIN}(\Lambda)$ (note-se que se $\lambda \in \Lambda$ e $\|\lambda\|^2 = d^2_{MIN}(\Lambda)$ então $(\lambda,0,...,0) \in \Lambda^m$ e $\|(\lambda,0,...,0)\|^2 = d^2_{MIN}(\Lambda)$; além disso, $d^2_{MIN}(\Lambda^m) \geq d^2_{MIN}(\Lambda))$ e $V(\Lambda^m) = V(\Lambda)^m$, tem-se $\gamma(\Lambda^m) = \gamma(\Lambda)$, e assim o parâmetro de Hermite é invariante por produto cartesiano, além de ser invariante por similaridade, razão pela qual será o parâmetro adotado, neste trabalho, para se medir a eficiência de empacotamento de um reticulado.

Os parâmetros geométricos principais de um reticulado $\Gamma(d^2_{MIN}(\Gamma), M_0(\Gamma), V(\Gamma)$ e $\gamma(\Gamma))$ podem ser relacionados aos parâmetros geométricos de um subreticulado Λ e da correspondente partição Γ/Λ, como segue.

Como para qualquer partição de grupo, adotando uma medida de distância invariante por translação, temos para $\Lambda \leq \Gamma$:

$$d^2_{MIN}(\Gamma) = MIN\left\{d^2_{MIN}(\Gamma/\Lambda);\ d^2_{MIN}(\Lambda)\right\} \qquad (2.4.1)$$

onde $d^2_{MIN}(\Gamma)$ e $d^2_{MIN}(\Lambda)$ são as distâncias Euclideanas quadráticas mínimas de Γ e Λ, respectivamente, e $d^2_{MIN}(\Gamma/\Lambda)$ é o mínimo entre os pesos Euclideanos quadráticos mínimos das classes de Γ/Λ distintas de Λ, ou seja:

$$d^2_{MIN}(\Gamma) = MIN\left\{d^2_{MIN}((c+\Lambda),\Lambda) : (c+\Lambda) \in \Gamma/\Lambda,\ (c+\Lambda) \neq \Lambda\right\} \qquad (2.4.2)$$

onde, para $(c+\Lambda) \in \Gamma/\Lambda$, $(c+\Lambda) \neq \Lambda$, temos:

$$d^2_{MIN}((c+\Lambda),\Lambda) = MIN\left\{\|\lambda\|^2 : \lambda \in (c+\Lambda)\right\} \qquad (2.4.3)$$

O número de vizinhos mais próximos em Γ pode ser obtido da seguinte forma: seja $M_0(c+\Lambda)$, com $(c+\Lambda) \in \Gamma/\Lambda$ e $(c+\Lambda) \neq \Lambda$, o número de vetores de $(c+\Lambda)$ cujo peso Euclideano quadrático é igual a $d^2_{MIN}((c+\Lambda),\Lambda)$; seja $M_0(\Gamma/\Lambda)$ a soma $\sum M_0(c+\Lambda)$ estendida às classes laterais em Γ/Λ cujo peso Euclideano quadrático mínimo $d^2_{MIN}((c+\Lambda),\Lambda)$ é igual a $d^2_{MIN}(\Gamma/\Lambda)$; assim:

$$M_0(\Gamma) = \begin{cases} M_0(\Lambda) & , \text{se } d^2_{MIN}(\Lambda) < d^2_{MIN}(\Gamma/\Lambda) \\ M_0(\Lambda) + M_0(\Gamma/\Lambda) & , \text{se } d^2_{MIN}(\Lambda) = d^2_{MIN}(\Gamma/\Lambda) \\ M_0(\Gamma/\Lambda) & , \text{se } d^2_{MIN}(\Lambda) > d^2_{MIN}(\Gamma/\Lambda) \end{cases} \tag{2.4.4}$$

O relacionamento entre os volumes fundamentais, $V(\Gamma)$ e $V(\Lambda)$ é obtido com as Equações (2.3.5), para Γ e Λ, e (2.3.6) para Γ/Λ:

$$\mathbb{R}^n = R(\Gamma) + \Gamma \tag{2.4.5}$$

$$\mathbb{R}^n = R(\Lambda) + \Lambda \tag{2.4.6}$$

$$\Gamma = [\Gamma/\Lambda] + \Lambda \tag{2.4.7}$$

Substituindo (2.4.7) em (2.4.5), obtém-se:

$$\mathbb{R}^n = R(\Gamma) + [\Gamma/\Lambda] + \Lambda \tag{2.4.8}$$

Comparando-se (2.4.6) com (2.4.8), vem:

$$R(\Lambda) = R(\Gamma) + [\Gamma/\Lambda] \tag{2.4.9}$$

Observa-se que nas equações (2.4.5)-(2.4.9), um elemento do conjunto do membro esquerdo é expresso de forma única como a soma de elementos dos conjuntos do membro direito; em particular, podemos interpretar a Equação (2.4.9) da seguinte forma: uma região fundamental de Λ pode ser obtida pela união de $|\Gamma/\Lambda|$ versões transladadas disjuntas de uma região fundamental de Γ determinadas pelos vetores de um conjunto de representantes $[\Gamma/\Lambda]$ selecionado para a partição Γ/Λ. Em conseqüência, obtém-se:

$$V(\Lambda) = |\Gamma/\Lambda| \, . \, V(\Gamma) \tag{2.4.10}$$

ou seja, o volume fundamental de $\Lambda \leq \Gamma$ é igual ao volume fundamental de Γ multiplicado pelo número de classes laterais na partição Γ/Λ. Assim, verifica-se também que, como volumes fundamentais de reticulados são sempre finitos e não nulos, o número de classes laterais em uma partição de reticulados de mesma dimensão é sempre finito, como mencionado na seção anterior.

Uma vez relacionadas as distâncias quadráticas mínimas, $d^2_{MIN}(\Gamma)$ e $d^2_{MIN}(\Lambda)$, e os volumes, $V(\Gamma)$ e $V(\Lambda)$, de dois reticulados Γ e $\Lambda \leq \Gamma$ do $\mathbb{R}^n$, torna-se imediato relacionar seus parâmetros de Hermite, $\gamma(\Gamma)$ e $\gamma(\Lambda)$:

$$\gamma(\Gamma) = \begin{cases} |\Gamma/\Lambda|^{2/n} \cdot \gamma(\Lambda), & \text{se } d^2_{MIN}(\Lambda) \leq d^2_{MIN}(\Gamma/\Lambda) \\ |\Gamma/\Lambda|^{2/n} \cdot \dfrac{d^2_{MIN}(\Gamma/\Lambda)}{d^2_{MIN}(\Lambda)} \cdot \gamma(\Lambda), & \text{se } d^2_{MIN}(\Lambda) > d^2_{MIN}(\Gamma/\Lambda) \end{cases} \qquad (2.4.11)$$

Se $\Lambda \leq \Gamma$ é semelhante a Γ, sendo Λ obtida a partir de Γ pela aplicação de uma similaridade endomórfica da forma αT, onde T é um operador ortogonal e α é um escalar real não-nulo, então os resultados anteriores se simplificam: $d^2_{MIN}(\Gamma) = d^2_{MIN}(\Lambda)/|\alpha|^2$, $M_0(\Gamma) = M_0(\Lambda)$, $V(\Gamma) = V(\Lambda)/|\alpha|^n$ e $\gamma(\Gamma) = \gamma(\Lambda)$; assim, necessariamente, teremos $d^2_{MIN}(\Gamma) \leq d^2_{MIN}(\Lambda)$ e $V(\Gamma) \leq V(\Lambda)$, pois $|\alpha| \geq 1$ para que $\alpha.T$ seja um endomorfismo.

Numa classe $(c+\Lambda)$, onde $c \in [\Gamma/\Lambda]$ e $c \neq 0$, um vetor com norma igual a $d^2_{MIN}((c+\Lambda),\Lambda)$ é denominado um **líder** ou **representante padrão** de $(c+\Lambda)$ (a origem é considerada o líder da classe Λ); um conjunto de representantes $[\Gamma/\Lambda]$ constituído exclusivamente de líderes é denominado um **conjunto padrão de representantes** para Γ/Λ.

Encerramos esta breve introdução aos reticulados do $\mathbb{R}^n$ com uma observação importante. Como mencionamos na Seção 2.3, uma união finita de classes laterais de um reticulado Λ do $\mathbb{R}^n$, na partição $\mathbb{R}^n/\Lambda$, pode constituir um outro reticulado do $\mathbb{R}^n$. Mesmo quando esta união não é um reticulado (subgrupo discreto do $\mathbb{R}^n$) ainda assim é um subconjunto discreto (infinito) do $\mathbb{R}^n$, de modo que podemos colocar esferas disjuntas de mesmo raio não-nulo máximo, com centros em cada um de seus pontos, formando um empacotamento esférico dito periódico. A principal desvantagem desses empacotamentos em relação aos reticulados é que, normalmente, esses empacotamentos periódicos apresentam regiões de Voronói (definidas como para os reticulados) não isométricas, e até mesmo com volumes distintos.

Capítulo 3

Estrutura algébrica da partição de reticulados

3.1 Introdução

Este capítulo trata da determinação de um conjunto de representantes de classes laterais $[\Gamma/\Lambda]$ para uma partição de reticulados Γ/Λ, e da soma destes representantes módulo Λ, quando Γ e Λ são especificados por matrizes geradoras. Após este desenvolvimento na Seção 3.2, estudamos na Seção 3.3 as matrizes geradoras (resp., particionadoras) para um mesmo reticulado (resp., uma mesma partição), concluindo na Seção 3.4 com formas especiais das matrizes particionadoras.

3.2 Conjunto básico de representantes e sua estrutura algébrica

Vimos na Seção 2.3 que um reticulado Λ do $\mathbb{R}^n$ é um subgrupo do grupo aditivo do espaço vetorial $\mathbb{R}^n$, constituído pelas combinações lineares com coeficientes inteiros de n vetores linearmente independentes do $\mathbb{R}^n$, $\{w_1,...,w_n\}$, ou seja, pela Equação (2.3.2):

$$\Lambda = M.\mathbb{Z}^n$$

onde M é uma matriz geradora de Λ, cujas colunas são os vetores w_j, $j = 1,\ldots,n$. Forma-se uma partição de grupo $\mathbb{R}^n/\Lambda$, e uma região fundamental $R(\Lambda)$ é qualquer conjunto de representantes $[\mathbb{R}^n/\Lambda]$ para as classes laterais de Λ em $\mathbb{R}^n/\Lambda$, de forma que, pela equação (2.3.5):

$$\mathbb{R}^n = R(\Lambda) + \Lambda$$

Uma possível e adequada região fundamental para Λ é o paralelotopo $R_M(\Lambda)$, associado a uma matriz geradora M de Λ, dado por:

$$R_M(\Lambda) \triangleq M.[0;1)^n = \{M.\delta : \delta \in [0;1)^n\} \qquad (3.2.1)$$

e que denominaremos **região fundamental básica** associada à base de Λ constituída pelas colunas de M. Para verificar que $R_M(\Lambda)$ constitui realmente um conjunto de representantes $[\mathbb{R}^n/\Lambda]$ das classes laterais de Λ na partição $\mathbb{R}^n/\Lambda$, introduzimos a seguinte função, definida por uma matriz geradora M de Λ (note-se que qualquer matriz geradora é não-singular):

$$x\mathrm{MOD}_M\Lambda \triangleq M.\mathrm{REST}(M^{-1}.x),\ x \in \mathbb{R}^n \qquad (3.2.2)$$

onde REST(y) é a parte não-inteira de y, definida para $y = [y_1,...,y_n] \in \mathbb{R}^n$ por:

$$\mathrm{REST}(y) = \begin{bmatrix} \mathrm{REST}(y_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{REST}(y_n) \end{bmatrix} \triangleq \begin{bmatrix} y_1-[y_1] \\ \cdot \\ \cdot \\ \cdot \\ y_n-[y_n] \end{bmatrix}$$

sendo $[y_i]$, para $y_i \in \mathbb{R}$, o maior inteiro menor ou igual a y_i. Assim, temos que $REST(y) \in [0;1)^n$, para $y \in \mathbb{R}^n$, de forma que $\mathbb{R}^n MOD_M \Lambda \subseteq R_M(\Lambda)$; como $R_M(\Lambda) MOD_M \Lambda = R_M(\Lambda)$, temos que:

$$\mathbb{R}^n MOD_M \Lambda = R_M(\Lambda) \tag{3.2.3}$$

É evidente que $REST(y) = REST(y')$, com $y,y' \in \mathbb{R}^n$, se e somente se $y-y' \in \mathbb{Z}^n$; como $xMOD_M\Lambda - x'MOD_M\Lambda = M.(REST(M^{-1}.x) - REST(M^{-1}.x'))$, com $x,x' \in \mathbb{R}^n$, podemos afirmar, em vista da não-singularidade de M, que:

$$xMOD_M\Lambda = x'MOD_M\Lambda \text{ se e somente se } x-x' \in \Lambda \tag{3.2.4}$$

para quaisquer $x,x' e \mathbb{R}^n$. Assim, de (3.2.3) e (3.2.4), tem-se que, para qualquer matriz geradora M de Λ, $R_M(\Lambda)$ é um conjunto de representantes $[\mathbb{R}^n/\Lambda]$ das classes laterais de Λ em $\mathbb{R}^n/\Lambda$, e $xMOD_M\Lambda$, para $x \in \mathbb{R}^n$, é o representante, em $R_M(\Lambda)$, da classe lateral de Λ à qual x pertence. Também com a ajuda da função $(.)MOD_M\Lambda$, podemos somar representantes em $R_M(\Lambda)$ módulo Λ, pois para $c,c' \in R_M(\Lambda)$ tem-se

$$c \dot{+} c' = (c + c')MOD_M\Lambda \tag{3.2.5}$$

onde a soma $c + c'$ no argumento da função $(.)MOD_M\Lambda$ é a soma vetorial em $\mathbb{R}^n$ e a soma $c \dot{+} c'$ no lado esquerdo é a soma em $R_M(\Lambda)$ módulo Λ, como convencionado no capítulo anterior. A Figura 3.2.1 ilustra a região fundamental básica $R_M(D_2)$ do reticulado D_2 do $\mathbb{R}^n$, associada à matriz geradora de D_2 dada por:

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

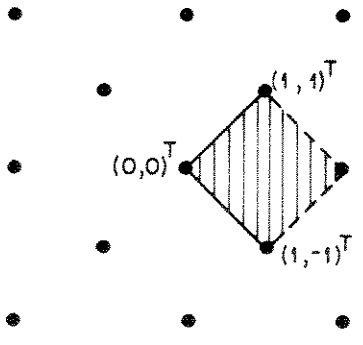

Fig. 3.2.1 - A região fundamental básica $R_M(D_2)$ associada à matriz $M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Para qualquer reticulado Λ do $\mathbb{R}^n$, o número de representantes de classes laterais em qualquer $[\mathbb{R}^n/\Lambda]$ é infinito, pois infinito também é o número dessas classes; assim, $R_M(\Lambda)$ tem infinitos elementos. No entanto, o volume de $R_M(\Lambda)$ é finito, e igual ao volume de qualquer outra região fundamental, como já visto na Seção 2.4; em particular, é igual ao volume da região

de Voronói. O volume fundamental $V(\Lambda)$ é, entretanto, extremamente simples de ser determinado utilizando uma região fundamental básica $R_M(\Lambda)$, com alguma matriz geradora M de Λ:

$$V(\Lambda) = |\det(M)| \tag{3.2.6}$$

Como qualquer matriz geradora M é não-singular, $V(\Lambda)$ é finito não-nulo. Vimos também na Seção 2.4, que o número $|\Gamma/\Lambda|$ de classes laterais de $\Lambda \leq \Gamma$ na partição Γ/Λ está relacionado aos volumes fundamentais $V(\Gamma)$ e $V(\Lambda)$ pela Equação (2.4.10):

$$V(\Lambda) = |\Gamma/\Lambda|.V(\Gamma)$$

de onde concluímos que, como volumes fundamentais são finitos não-nulos, $|\Gamma/\Lambda|$ é finito e não-nulo, para Λ e Γ de mesma dimensão, como convencionado. Assim, diferentemente da partição $\mathbb{R}^n/\Lambda$, uma partição Γ/Λ é sempre finita e não-vazia, como já visto.

É interessante, para fins práticos de manipulação das classes laterais de uma partição Γ/Λ, determinarmos um conjunto de representantes $[\Gamma/\Lambda]$ dessas classes laterais, pois como já vimos a soma de representantes módulo Λ corresponde à soma das respectivas classes laterais, sendo esta correspondência uma das principais aplicações do conceito de conjunto de representantes. Vimos, nesta seção, que $R_M(\Lambda)$, com M sendo uma matriz geradora de Λ, é uma possível região fundamental de Λ; em particular, $R_M(\Lambda)$ possui representantes para todas as classes laterais de Λ na partição $\mathbb{R}^n/\Lambda$, ou seja, contidas em $\mathbb{R}^n$. Se $\Lambda \leq \Gamma$ então Γ é a união de $|\Gamma/\Lambda|$ classes laterais de Λ; como $R_M(\Lambda)$ possui representantes para todas as classes laterais de Λ contidas no $\mathbb{R}^n$, possui então representantes para as classes laterais de Λ contidas em $\Gamma \leq \mathbb{R}^n$. Assim, é natural que se analise a conveniência de utilizarmos estes representantes como um possível conjunto de representantes $[\Gamma/\Lambda]$ das classes laterais de Λ na partição Γ/Λ. É o que faremos a seguir.

Seja, então, M uma matriz geradora qualquer de Λ. A região fundamental básica $R_M(\Lambda)$ de Λ, associada à matriz geradora M de Λ, por ser um conjunto de representantes de uma partição de grupo, especificamente $\mathbb{R}^n/\Lambda$, não pode possuir dois vetores de uma mesma classe lateral como elementos, embora possua um ponto de cada classe lateral (de Λ contida no $\mathbb{R}^n$). Assim, se $\Lambda \leq \Gamma$, ao tomarmos a interseção de $R_M(\Lambda)$ com Γ, obteremos um ponto, e apenas um ponto, de cada classe lateral de Λ contida em Γ; desta forma, o conjunto de representantes $[\Gamma/\Lambda]$ selecionado em $\mathbb{R}_M(\Lambda)$ seria:

$$[\Gamma/\Lambda]_M \triangleq \Gamma \bigcap R_M(\Lambda) \tag{3.2.7}$$

A Figura 3.2.2 ilustra a interseção $[\mathbb{Z}^2/D_2]_M = \left\{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}\right\}$ entre o reticulado $\mathbb{Z}^2$ e a região fundamental básica $R_M(D_2)$ do reticulado D_2, dada na Figura 3.2.1.

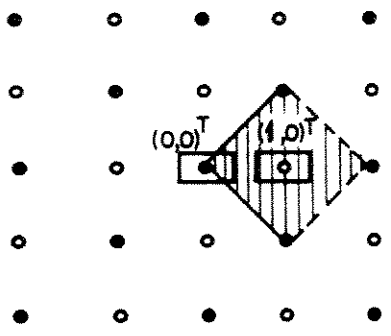

Fig. 3.2.2 - O conjunto básico de representantes $[\mathbb{Z}^2/D_2]$, utilizando-se a região fundamental básica $R_M(D_2)$ dada na Figura 3.2.1

Embora a determinação dos pontos de um reticulado que pertença a uma dada região arbitrária do espaço seja um problema, em geral, complexo, ([6],[10]), neste caso a solução é relativamente simples. Esta se baseia numa determinação indireta da interseção indicada em (3.2.7) utilizando a função $(.)MOD_M\Lambda$ definida em (3.2.2), da seguinte forma:

$$[\Gamma/\Lambda]_M = \Gamma MOD_M\Lambda \triangleq \{\gamma MOD_M\Lambda : \gamma \in \Gamma\} \tag{3.2.8}$$

É claro que, como $\Gamma \cap R_M(\Lambda) \subseteq \Gamma$, temos $(\Gamma \cap R_M(\Lambda))MOD_M\Lambda \subseteq \Gamma MOD_M$; porém qualquer elemento de $R_M(\Lambda)$ é um ponto fixo de $(.)MOD_M\Lambda$, de modo que $(\Gamma \cap R_M(\Lambda))MOD_M\Lambda = \Gamma \cap R_M(\Lambda)$; assim $\Gamma \cap R_M(\Lambda) \subseteq \Gamma MOD_M\Lambda$. Além disso, se $\gamma \in \Gamma$ então, como $\Gamma \cap R_M$ é um conjunto de representantes para as classes laterais de Λ em Γ/Λ, existe $c \in \Gamma \cap R_M$ tal que $\gamma - c \in \Lambda$; logo, por (3.2.4), $\gamma MOD_M\Lambda = c$; assim, $\Gamma MOD_M\Lambda \subseteq \Gamma \cap R_M$. Concluímos, então, que $\Gamma \cap R_M(\Lambda) = \Gamma MOD_M\Lambda$, validando a Equação (3.2.8), ou seja, $[\Gamma/\Lambda]_M = \Gamma MOD_M\Lambda$.

A determinação de $\Gamma MOD_M\Lambda$ é desenvolvida como segue. Sejam N e M matrizes geradoras de Γ e $\Lambda \leq \Gamma$, respectivamente; assim M = N.P, onde P é uma matriz de números inteiros não-singular, denominada matriz particionadora para Γ/Λ, como definida na Seção 2.3. Seja $\gamma \in \Gamma$; assim, $\gamma = N.\mu$, onde $\mu \in \mathbb{Z}^n$. Pela definição (3.2.2), temos:

$$\begin{aligned} \gamma MOD_M\Lambda &= M.REST(M^{-1}.\gamma) = \\ &= M.REST(M^{-1}.N.\mu) = \\ &= M.REST(P^{-1}.\mu) \end{aligned} \tag{3.2.9}$$

Denotemos por D(y), com $y = [y_1,...,y_n] \in \mathbb{R}^n$, a matriz diagonal com y_i como o elemento da diagonal na i^a linha, $1 \leq i \leq n$. Além disso, para um vetor $q = [q_1,...,q_n]$ com componentes inteiras positivas não-nulas, definamos:

$$zMODq \triangleq \begin{bmatrix} z_1 MOD q_1 \\ \cdot \\ \cdot \\ \cdot \\ z_n MOD q_n \end{bmatrix} , \quad z = \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{bmatrix} \in \mathbb{Z}^n \tag{3.2.10}$$

onde $z_i MOD q_i \triangleq REST(z_i/q_i)$, $1 \leq i \leq n$. Nestes termos, vale a seguinte propriedade, facilmente verificável:

$$REST(D(q)^{-1}.z) = D(q)^{-1}.(zMODq),\ z \in \mathbb{Z}^n \tag{3.2.11}$$

onde q é um vetor com componentes inteiras positivas não-nulas. Nosso objetivo será determinar vetores q e z, em função de P e μ, tais que $P^{-1}.\mu = D(q)^{-1}.z$, de maneira que se possa aplicar a propriedade (3.2.11) para simplificar a expressão (3.2.9).

Observando que $|det(P)|.P^{-1}$ é uma matriz não-singular de números inteiros, definamos o vetor d de componentes inteiras positivas não-nulas como tendo sua i^a componente igual ao máximo divisor comum dos elementos da i^a linha de $|det(P)|.P^{-1}$, $1 \leq i \leq n$; ou, em símbolos:

$$d = MDC(|det(P)|.P^{-1}) \tag{3.2.12}$$

É evidente que $D(d)^{-1}.(|det(P)|.P^{-1})$ é matriz de números inteiros; assim podemos definir, para $\mu \in \mathbb{Z}^n$, o seguinte vetor $z \in \mathbb{Z}^n$:

$$z = D(d)^{-1}.|det(P)|.P^{-1}.\mu \tag{3.2.13}$$

Além disso, como $(D(d)^{-1}.|det(P)|).P^{-1}$ e P são matrizes de números inteiros, podemos definir o seguinte vetor q com componentes inteiras positivas não-nulas:

$$D(q) = D(d)^{-1}.|det(P)| \tag{3.2.14}$$

Visto que para vetores z e q, assim definidos, temos:

$$D(q)^{-1}.z = (D(d)^{-1}.|det(P)|)^{-1}.D(d).|det(P)|.P^{-1}.\mu = P^{-1}.\mu$$

podemos utilizar a propriedade (3.2.11), com estes vetores q e z, para simplificar a expressão (3.2.9); assim teremos:

$$\gamma MOD_M \Lambda = M.D(q)^{-1}.(zMODq) =$$

$$= M.D(q)^{-1}.((D(q).P^{-1}.\mu)MODq) =$$

$$= M.D(q)^{-1}.((D(q).M^{-1}.N.\mu)MODq) \tag{2.2.15}$$

Denotando:

$$M_q \triangleq M.D(q)^{-1},\ P_q = P.D(q)^{-1} \tag{3.2.16}$$

tem-se:

$$\gamma MOD_M\Lambda = M_q.((M_q^{-1}.\gamma)MODq) = M_q.((P_q^{-1}.\mu)MODq) \tag{3.2.17}$$

É interessante notar que o vetor q, dado por (3.2.14), tem componentes mínimas.

Assim, $[\Gamma/\Lambda]_M = \Gamma MOD_M\Lambda$ pode ser expresso por:

$$[\Gamma/\Lambda]_M = \left\{M_q.((P_q^{-1}.\mu)MODq) : \mu \in \mathbb{Z}^n\right\} \tag{3.2.18}$$

e a soma de representantes $c,c' \in [\Gamma/\Lambda]_M$, módulo Λ, pode ser expressa por:

$$c \dotplus c' = M_q((M_q^{-1}.c + M_q^{-1}.c')MODq) \tag{3.2.19}$$

em lugar de (3.2.5), concluindo-se assim que os elementos de $M_q^{-1}.[\Gamma/\Lambda]_M$ formam um grupo sob adição módulo q, ou seja, formam um subgrupo do grupo abeliano $\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn}$, onde $\mathbb{Z}_{qi}$ é o grupo abeliano dos resíduos inteiros módulo q_i (isomórfico à partição unidimensional $\mathbb{Z}/q_i\mathbb{Z}$), $1 \le i \le n$.

A determinação de $[\Gamma/\Lambda]_M$ por (3.2.18) pode ser simplificada ainda mais: extendendo a definição (3.2.10) para matrizes (quando, então, para uma matriz de números inteiros Z e um vetor q com componentes inteiras positivas não-nulas, ZMODq denota a matriz cuja j^a coluna é Z_jMODq, onde Z_j é a j^a coluna de Z, $1 \le j^a \le n$), verifica-se facilmente que:

$$(P_q^{-1}.\mu)MODq = ((P_q^{-1}\ MODq).\mu)MODq \tag{3.2.20}$$

para $\mu \in \mathbb{Z}^n$. Note-se que, sendo $m = MMC[q_1,...,q_n]$, é suficiente que, em (3.2.20), μ assuma valores em $\mathbb{Z}_m$ x...x $\mathbb{Z}_m$ para que, após multiplicação por M_q, sejam obtidos todos os vetores de $[\Gamma/\Lambda]_M$ (nesta enumeração cada vetor de $[\Gamma/\Lambda]_M$ será, obviamente, obtido $m^n/|\Gamma/\Lambda|$ vezes como resultado, pois existem $\prod_{i=1}^{n} q_i$ vetores distintos em $\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn}$ que μ pode assumir; na Seção 3.4 veremos como contornar este problema, utilizando formas canônicas da matriz de partição, evitando assim estes cálculos redundantes de um mesmo representante, ou seja, de modo que $m^n = |\Gamma/\Lambda|$). Assim, podemos reescrever (3.2.18) como:

$$[\Gamma/\Lambda]_M = \left\{M_q.(((P_q^{-1}MODq).\mu)MODq) : \mu \in \mathbb{Z}_m\ x...x\ \mathbb{Z}_m\right\} \tag{3.2.21}$$

Note-se, portanto, que as n colunas de $P_q^{-1}MODq$ formam um conjunto gerador para o subgrupo $M_q^{-1}.[\Gamma/\Lambda]_M$ de $\mathbb{Z}_m$ x...x $\mathbb{Z}_m$ (um conjunto gerador mínimo para $M_q^{-1}.[\Gamma/\Lambda]_M$ pode ser obtido, obviamente, tomando-se um conjunto mínimo de colunas de $P_q^{-1}MODq$ tal que as restantes sejam combinações lineares, em $\mathbb{Z}_m$ x...x $\mathbb{Z}_m$, das colunas desse conjunto). Tomemos, como um exemplo elementar, o caso da partição $\mathbb{Z}^2/D_2$ ilustrada acima, com $M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ e $N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ como as matrizes geradoras de D_2 e $\mathbb{Z}^2$, respectivamente. Assim, teríamos:

$$P = N^{-1}.M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

$$d = MDC(|\det(P)|.P^{-1}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad D(d) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$D(q) = D(d)^{-1}.|\det(P)| = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad q = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \qquad m = 2$$

$$M_q = M.D(q)^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \qquad P_q = P.D(q)^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

$$P_q^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad P_q^{-1}MODq = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Logo: $[\mathbb{Z}^2/D_2] = \left\{ \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} . \left(\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} . \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right) MOD \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right) : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \in \mathbb{Z}_2 \text{ x } \mathbb{Z}_2 \right\} =$

$$= \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

onde:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \dot{+} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \dot{+} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix} \dot{+} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \dot{+} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Observe-se que $M_q^{-1}.[\mathbb{Z}^2/D_2]$ tem um conjunto gerador unitário $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

Assim, fica resolvido, de forma fechada, o problema da determinação de um conjunto de representantes $[\Gamma/\Lambda]$ para uma partição Γ/Λ de reticulados (de mesma dimensão) especificados por matrizes geradoras; a solução deste problema pode ser usada, por exemplo, na busca exaustiva de novos reticulados. Trataremos, agora, do problema recíproco, ou seja, da determinação de uma matriz geradora para um reticulado Γ, quando se conhece uma matriz geradora para um subreticulado $\Lambda \leq \Gamma$ e um conjunto de representantes $[\Gamma/\Lambda]$ para a partição; em particular, a solução deste problema fornece uma matriz geradora para reticulados contruídos

por união finita de versões transladadas de um reticulado para o qual já se conhece uma matriz geradora (este é o método de construção de reticulados mais difundido na literatura, como já mencionado na Seção 2.3), possibilitando o estudo dessas construções em termos de formas quadráticas.

Assim, como no problema anterior, formulamos a determinação de uma solução particular $[\Gamma/\Lambda]_M$ para a seleção de um conjunto de representantes $[\Gamma/\Lambda]$, delineamos a seguir um procedimento para a determinação de uma possível matriz geradora N para um reticulado Γ, conhecido $[\Gamma/\Lambda]$ e uma matriz geradora M de Λ.

Inicialmente, determinamos $[\Gamma/\Lambda]_M = [\Gamma/\Lambda]MOD_M\Lambda$; note-se que $[\Gamma/\Lambda]$ e $[\Gamma/\Lambda]_M$ possuem o mesmo número $|\Gamma/\Lambda|$ finito de elementos. Do exposto no problema anterior $M^{-1}.[\Gamma/\Lambda]_M$ é um conjunto de vetores com componentes fracionárias; supondo que estas encontram-se sob forma irredutível (numerador e denominador primos-entre-si), fazemos q_i, $1 \le i \le n$, igual ao mínimo múltiplo comum dos denominadores (componentes nulas são consideradas como tendo numerador nulo e denominador unitário). Multiplicando cada vetor de $M^{-1}.[\Gamma/\Lambda]_M$ por D(q), obtemos $M_q^{-1}.[\Gamma/\Lambda]_M$ que, como vimos, forma um subgrupo de $\mathbb{Z}_m \times ... \times \mathbb{Z}_m$. Determina-se, então, um conjunto gerador mínimo para $M_q^{-1}.[\Gamma/\Lambda]_M$, e forma-se então uma matriz n x n inteira, tomando-se os elementos desse conjunto gerador como colunas desta matriz, completando-a com colunas nulas se o número de geradores for menor do que n; esta é a matriz $P_q^{-1}MODq$, para alguma matriz de partição $P = M^{-1}.N$. Se determinarmos P_q^{-1}, então $P = P_q.D(q)$ e $N = M.P^{-1}$ ficam determinadas, resolvendo o problema. Esta, no entanto, não é uma tarefa trivial, pois equivale a determinar uma matriz inteira K de modo que $P_q^{-1} = P_q^{-1}MODq + K.D(q)$ tenha determinante igual, em valor absoluto, a $|\det(P^{-1}.D(q))| = \left(\prod_{i=1}^{n} q_i\right)/|\Gamma/\Lambda|$; por outro lado, é evidente que a determinação desta matriz inteira K pode ser simplificada por uma escolha adequada do conjunto gerador mínimo para $M_q^{-1}.[\Gamma/\Lambda]_M$ que constitui as colunas não-nulas de $P_q^{-1}MODq$. Na Seção 3.4, onde são analisadas as formas canônicas da matriz particionadora, voltaremos a este problema. Antes, no entanto, devemos analisar como se relacionam matrizes particionadoras para uma mesma partição de reticulados; isto é realizado na Seção 3.3 a seguir.

3.3 Matrizes geradoras e particionadoras

Como vimos, um reticulado Λ é especificado, de modo geral, por uma matriz geradora M real quadrada não-singular, de modo que Λ é o conjunto de todas as combinações lineares com coeficientes inteiros das colunas de M; as colunas de M formam assim um conjunto gerador, ou base, de Λ. Cada reticulado Λ, no entanto, possui várias matrizes geradoras (ou bases). Passamos a analisar, então, o relacionamento entre duas matrizes geradoras M e M' para um mesmo reticulado Λ.

Sejam, então, M e M' duas matrizes geradoras do reticulado Λ. Como as colunas de cada uma dessas matrizes são combinações lineares com coeficientes inteiros das colunas da outra, temos:

$$M = M'.U \tag{3.3.1}$$

e

$$M' = M.V \tag{3.3.2}$$

onde U e V são matrizes quadradas de números. Substituindo (3.3.2) em (3.3.1), vem:

$$M = M.V.U \tag{3.3.3}$$

Já que M é inversível, (3.3.2) fica:

$$V.U = I$$

onde I é a matriz identidade de ordem n. Como U e V são matrizes quadradas, temos:

$$V = U^{-1}$$

Como U e $V = U^{-1}$ são matrizes de números inteiros, det(U) e det(V) = 1/det(U) são inteiros; logo:

$$\det(U) = \pm 1 \tag{3.3.4}$$

Além disso, de (3.3.1) e (3.3.4), tiramos:

$$|\det(M)| = |\det(M')|;$$

ou seja, confirma-se que o volume de Λ, dado por $V(\Lambda) = |\det(M)|$ onde M é uma matriz geradora de Λ, independe da matriz geradora (ou base) de Λ usada, como já era esperado.

Assim, toda mudança de base corresponde a uma multiplicação, à direita da matriz geradora associada, por uma matriz de números inteiros com determinante igual a ± 1.

Reciprocamente, toda multiplicação, à direita de uma matriz geradora, por uma matriz de números inteiros com determinante igual a ± 1 correspondente a mudar a base associada, obtendo-se assim outra matriz geradora. Isto é, se Λ é o reticulado gerado por M, e Λ' é o reticulado gerado por M' = M.V, onde V é uma matriz de números inteiros com determinante igual a ± 1, então $\Lambda = \Lambda'$; senão vejamos: já que V é uma matriz de números inteiros, temos:

$$V.\mathbb{Z}^n \leq \mathbb{Z}^n \tag{3.3.5}$$

Como V tem determinante igual a ± 1, V^{-1} também é uma matriz de números inteiros, e então temos:

$$\mathbb{Z}^n = V^{-1}.(V.\mathbb{Z}^n) \leq V.\mathbb{Z} \tag{3.3.6}$$

Assim, de (3.3.5) e (3.3.6), temos:

$$V.\mathbb{Z}^n = \mathbb{Z}^n$$

Conseqüentemente:

$$\Lambda' = M'.\mathbb{Z}^n = M.V.\mathbb{Z}^n = M.\mathbb{Z}^n = \Lambda$$

como afirmado.

Deste relacionamento entre matrizes geradoras, podemos obter a forma geral de um automorfismo para um reticulado. Como definido na Seção 2.3, um automorfismo T para um reticulado Λ é um operador não-singular do $\mathbb{R}^n$ tal que $T\Lambda = \Lambda$. Sabemos que para qualquer operador não-singular T, o conjunto $T\Lambda$ é um reticulado do $\mathbb{R}^n$, tendo TM como matriz geradora, onde M é uma matriz geradora de Λ. Se $T\Lambda = \Lambda$, então TM é na realidade uma outra matriz geradora de Λ; assim deve existir uma matriz V de números inteiros com determinante igual a ± 1 tal que T.M = M.V; ou seja, todo automorfismo de Λ tem a forma:

$$T = M.V.M^{-1}$$

para qualquer matriz geradora M de Λ e alguma matriz V de números inteiros com determinante igual a ± 1.

Vimos também que se $\Lambda \leq \Gamma$ e M(resp. N) é matriz geradora de Λ(resp. Γ), então M = N.P, onde P é uma matriz de números inteiros não-singular, denominada uma matriz particionadora para a partição Γ/Λ. Cada partição Γ/Λ, no entanto, possui várias matrizes particionadoras, uma para cada seleção das matrizes geradoras (ou bases) de Λ e Γ. Passamos a analisar, então, o relacionamento entre duas matrizes particionadoras P e P' para uma mesma partição Γ/Λ.

Sejam, então, P e P' duas matrizes particionadoras para a partição Γ/Λ. Logo, existem matrizes geradoras M e M' de Λ, e N e N' de Γ tais que:

$$M = N.P \tag{3.3.7}$$

$$M' = N'.P' \tag{3.3.8}$$

No entanto, temos também:

$$M' = M.U \tag{3.3.9}$$

e

$$N' = N.V \tag{3.3.10}$$

onde U e V são matrizes de números inteiros com $\det(U) = \pm 1$ e $\det(V) = \pm 1$. Logo, (3.3.7 - 10) implicam em:

$$P' = N'^{-1}.M' = V^{-1}.N^{-1}.M.U. = V^{-1}.P.U$$

Além disso, temos:

$$|\det(P')| = |\det(P)|;$$

ou seja, confirma-se que o número de classes laterais de Λ em Γ, dado por $|\Gamma/\Lambda| = V(\Lambda)/V(\Gamma) = |\det(M)|/|\det(N)| = |\det(P)|$ independe da matriz particionadora P usada para a partição, ou seja, das matrizes geradoras (ou bases) de Λ e Γ usadas, como já era esperado.

Assim, toda mudança de bases, em uma partição, corresponde a multiplicar, à direita e à esquerda da matriz particionadora associada, por matrizes de números inteiros com determinantes iguais a $\pm$ 1.

Reciprocamente, toda multiplicação, à direita e à esquerda de uma matriz particionadora, por matrizes de números inteiros com determinantes iguais a $\pm$ 1 corresponde a mudar as bases associadas, obtendo-se assim outra matriz particionadora. Isto é, se P é a matriz particionadora para Γ/Λ, então $P' = V^{-1}.P.U$, onde V e U são matrizes de números inteiros com determinantes iguais a $\pm$ 1, é também matriz particionadora para Γ/Λ; senão vejamos: como P é matriz particionadora para Γ/Λ, existem matrizes geradoras M e N dos reticulados Λ e Γ, respectivamente, tais que M = N.P; assim, M' = M.U e N' = N.V são também matrizes geradoras de Λ e Γ, respectivamente, e:

$$N'.P' = (N.V).(V^{-1}.P.U) = (N.P).U = M.U = M'$$

como afirmado.

Deste relacionamento entre matrizes particionadoras podemos obter a forma geral de um endomorfismo para um reticulado. Como definido na Seção 2.3, um endomorfismo T para um reticulado Γ é um operador não singular do $\mathbb{R}^n$ tal que $T\Gamma \leq \Gamma$. Novamente lembramos que, para qualquer operador linear não-singular T, o conjunto $T\Gamma$ é um reticulado do $\mathbb{R}^n$, tendo T.N como matriz geradora, onde N é uma matriz geradora de Γ. Se $T\Gamma \leq \Gamma$, então T.N = N.P, onde P é uma matriz de números inteiros com $|\det(P)| = |\det(T)|$; ou seja, todo endomorfismo de Γ tem a forma:

$$T = N.P.N^{-1}$$

para qualquer matriz geradora N de Γ e alguma matriz P de números inteiros com $|\det(P)| = |\det(T)|$.

Visto como se relacionam as matrizes particionadoras para uma mesma partição de reticulados, a questão que surge naturalmente é: existiriam formas computacionalmente mais adequadas da matriz de partição para uma dada partição de reticulados? Esta questão é analisada na Seção 3.4 a seguir, ilustrando suas aplicações aos problemas direto e recíproco da Seção 3.2.

3.4 Formas canônicas da matriz particionadora

Pelo exposto na Seção 3.3, observa-se que as matrizes inteiras com determinante igual a±1 desempenham um papel importante no relacionamento entre matrizes particionadoras para uma mesma partição. Assim, na busca de formas canônicas de matrizes particionadoras, seria útil termos conhecimento de propriedades destas matrizes de números inteiros cujos determinantes sejam iguais a ± 1, que denominaremos **Matrizes Unitárias** ([26]).

Como vimos, uma matriz unitária é, obviamente, não-singular, tendo como inversa uma matriz de números inteiros, também unitária. Além disso, as matrizes unitárias são as únicas matrizes de números inteiros não-singulares, cuja inversa é também uma matriz de números inteiros.

Uma característica interessante é que qualquer linha ou coluna de uma matriz unitária é constituída de números inteiros cujo máximo divisor comum é igual a 1, ou seja, são primos entre si. Para verificar, basta observar que o determinante de uma matriz de números inteiros é um número inteiro que, pelo desenvolvimento por menores de Laplace, pode ser expresso como uma combinação linear dos elementos de uma linha ou coluna arbitrária desta matriz, com coeficientes inteiros; como, pela teoria elementar dos números, qualquer tal combinação é um múltiplo do máximo divisor comum daqueles elementos, concluímos que o determinante de uma matriz de números inteiros é um múltiplo do máximo divisor comum dos elementos de qualquer de suas linhas ou colunas. Para o caso de uma matriz unitária, onde o determinante é igual a ± 1, conclui-se então que o máximo divisor comum dos elementos de qualquer de suas linhas ou colunas é necessariamente igual a 1, como afirmado.

Porém, importante mesmo é o fato de que dado um vetor de n números inteiros primos entre si, sempre é possível construir uma matriz unitária com aquele vetor como uma linha ou coluna previamente estabelecida.

Estas considerações sobre matrizes unitárias são suficientes para desenvolvermos as formas canônicas da matriz particionadora para uma partição de reticulados.

Vimos na Seção 3.3 que, dada uma matriz particionadora P, para uma partição Γ/Λ (existindo, portanto, matrizes geradoras M e N dos reticulados Λ e Γ, respectivamente, tais que $P = N^{-1}.M$), uma multiplicação à direita de P por uma matriz unitária U, corresponde a uma

multiplicação à direita de M pela mesma matriz unitária U, mudando assim a matriz geradora de para M' = M.U., preservando a matriz geradora N de Γ, e mudando portanto a matriz particionadora para P' = N^{-1}.(M.U) = P.U. Ocorre que existe sempre uma matriz unitária U de modo que P' seja triangular inferior, produzindo então uma primeira forma canônica; esta forma canônica será útil para a solução do problema (recíproco) da Seção 3.2, deixado provisoriamente em aberto. Por raciocínio idêntico, existem matrizes unitárias U e V de modo que P' = V^{-1}.P.U seja diagonal, produzindo então uma segunda forma canônica; esta forma canônica será útil para a simplificação da solução do problema (direto) dada na Seção 3.2.

3.4.1 Forma canônica triangular

Qualquer matriz quadrada A não-singular de números inteiros pode ser reduzida por **operações elementares sobre colunas**, definidas como multiplicações à direita de A por matrizes unitárias, a uma matriz triangular inferior única $\hat{A}$, onde todos os elementos não-nulos de $\hat{A}$ são inteiros positivos, com cada elemento à esquerda da diagonal principal sendo menor do que o elemento diagonal de sua linha. A obtenção desta forma canônica triangular, conhecida na literatura como **forma normal de Hermite** ([26]), não será reproduzida aqui, pois a sua existência e unicidade são suficientes para os nossos propósitos, como visto abaixo.

Retomemos, então, o problema da determinação de uma matriz geradora N de um reticulado Γ do $\mathbb{R}^n$, conhecidos uma matriz geradora M de um subreticulado $\Lambda \leq \Gamma$ e um conjunto de representantes $[\Gamma/\Lambda]$ para a partição de reticulados Γ/Λ, estabelecido ao final da Seção 2.2.

Analisemos, inicialmente, de que modo podemos reduzir a matriz P_q^{-1} à forma canônica triangular. A matriz P_q^{-1} pode ser expressa como:

$$P_q^{-1} = D(d)^{-1}.|\det(P)|.P^{-1}$$

onde $P = N^{-1}.M$, e:

$$D(d) = MDC|\det(d)|.P^{-1}$$

Assim, se a matriz de números inteiros $|\det(P)|.P^{-1}$ for reduzida à forma canônica triangular, então P_q^{-1} também será reduzida à forma canônica triangular, pois (1) D(d) é invariante por operações elementares sobre as colunas de $|\det(P)|.P^{-1}$, como facilmente verificável; (2) a multiplicação à esquerda de $|\det(P)|.P^{-1}$ por D(d) apenas multiplica cada linha de $|\det(P)|.P^{-1}$ pela correspondente componente de d, resultando em uma matriz P_q^{-1} em forma canônica triangular se $|\det(P)|.P^{-1}$ for reduzida à forma canônica triangular. A redução de $|\det(P)|.P^{-1}$ à forma canônica triangular, por sua vez, correspondendo a multiplicações à direita de $|\det(P)|.P^{-1}$ por matrizes unitárias, equivale a mudanças da matriz geradora N de Γ. Em vista da existência e unicidade da forma canônica triangular, concluímos que P_q^{-1} pode sempre ser colocada em forma canônica triangular, qualquer que seja a matriz geradora M de Λ utilizada, adotando-se uma certa (e única) matriz geradora N de Γ.

Vejamos, então, as simplificações, na solução do problema (direto) dada na Seção 3.2, decorrentes da conclusão acima. Vimos que, dados M e $[\Gamma/\Lambda]$, podemos obter $[\Gamma/\Lambda]_M$; assim o subgrupo $D(q).M^{-1}.[\Gamma/\Lambda]_M$ de $\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn}$ pode ser determinado. Vimos, também, que $P_q^{-1}MODq$ é uma matriz cujas colunas formam um conjunto gerador para $D(q).M^{-1}.[\Gamma/\Lambda]_M$. Portanto, a solução, proposta na Seção 3.2, consistia de um conjunto gerador arbitário para $D(q).M^{-1}.[\Gamma/\Lambda]_M$ com n elementos que, dispostos em ordem arbitrária como colunas, resulta em uma matriz $P_q^{-1}MODq$ para alguma matriz P_q^{-1} dada por:

$$P_q^{-1} = P_qMODq + D(q).K$$

onde K é uma matriz de números inteiros a ser determinada de modo que os elementos de cada linha de P_q^{-1} sejam primos-entre-si e que $|\det(P_q^{-1})| = \left(\prod_{i=1}^{n} q_i\right)/|\Gamma/\Lambda|$. Como mencionado naquela seção, a determinação, para uma matriz $P_q^{-1}MODq$ selecionada, da matriz de números inteiros K é bastante complexa, em geral. No entanto, se P_q^{-1} está na forma triangular, então $P_q^{-1}MODq$ e K são matrizes triangulares inferiores (não necessariamente em forma canônica); assim, da existência e unicidade desta forma canônica para P_q^{-1}, concluímos que existem matrizes $P_q^{-1}MODq$ e K triangulares inferiores e únicas tais que a matriz P_q^{-1} resultante (com $|\det(P_q^{-1})| = \left(\prod_{i=1}^{n} q_i\right)/|\Gamma/\Lambda|$ e cujos elementos de cada linha são primos-entre-si) está na forma canônica triangular. A solução, portanto, torna-se computacionalmente mais simples, pois agora: (1) existem menos elementos da matriz K a determinar; (2) o determinante P_q^{-1} é simplesmente o produto dos seus elementos na diagonal principal; (3) todos os elementos não-nulos de P_q^{-1} são positivos; (4) cada elemento de P_q^{-1} à esquerda da diagonal principal é menor do que o elemento diagonal de sua linha. Observemos, no entanto, que $P_q^{-1}MODq$ não é mais arbitrária, e sim única; ou seja, existe apenas uma matriz $P_q^{-1}MODq$ triangular, cujas colunas formam um conjunto gerador para $D(q).M^{-1}.[\Gamma/\Lambda]_M$, tal que é possível encontrar uma matriz K triangular, de números inteiros, de modo que a matriz P_q^{-1} resultante (cujo determinante deve ser igual à $\left(\prod_{i=1}^{n} q_i\right)/|\Gamma/\Lambda|$ e cujos elementos de cada linha devem ser primos-entre-si) esteja em forma canônica triangular.

Como um exemplo ilustrativo, determinamos uma matriz geradora N para o reticulado Γ do $\mathbb{R}^n$ que possui o reticulado D_2 como subreticulado definido pela matriz geradora:

$$M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

formando uma partição Γ/D_2 com conjunto de representantes:

$$[\Gamma/D_2] = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Inicialmente, determinamos $[\Gamma/D_2]_M = [\Gamma/D_2]MOD_M D_2$:

$$[\Gamma/D_2]_M = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

(comparando-se com o exemplo ilustrativo da Seção 3.2, temos que Γ deve ser o reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$, como veremos). Em seguinda, determinaremos $M^{-1}.[\Gamma/D_2]^M$:

$$M^{-1}.[\Gamma/D_2]_M = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \right\}$$

Fazendo q_i, $1 \leq i \leq 2$, igual ao mínimo múltiplo comum dos denominadores das primeiras componentes dos elementos de $M^{-1}.[\Gamma/D_2]_M$, obtemos:

$$q = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Assim, $\left(\prod_{i=1}^{n} q_i \right) / |\Gamma/D_2| = (2x2)/2 = 2$, e:

$$[\Gamma/D_2]_M = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Confirma-se, portanto, que Γ é realmente um reticulado, pois $D(q).M^{-1}.[\Gamma/D_2]_M$ é, de fato, um subgrupo de $\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn}$, como facilmente verificável.

Fazendo:

$$P_q^{-1}MODq = \begin{bmatrix} \pi_{11} & 0 \\ \pi_{12} & \pi_{22} \end{bmatrix}$$

$$K = \begin{bmatrix} K_{11} & 0 \\ K_{12} & K_{22} \end{bmatrix}$$

$$P_q^{-1} = P_q^{-1}MODq + D(q).K = \begin{bmatrix} (\pi_{11} + K_{11}) & 0 \\ (\pi_{12} + K_{12}) & (\pi_{22} + K_{22}) \end{bmatrix}$$

Determinaremos π_{11}, π_{12}, π_{22}, K_{11}, K_{11} e K_{22} tais que:

(1) $\left\{ \begin{bmatrix} \pi_{11} \\ \pi_{12} \end{bmatrix}, \begin{bmatrix} 0 \\ \pi_{12} \end{bmatrix} \right\}$ é um conjunto gerador para $D(q).M^{-1}.[\Gamma/D_2]_M$

(2) K_{11}, K_{11}e K_{11} são números inteiros

(3) $(\pi_{11} + 2.K_{11}) \geq 0$, $(\pi_{22} + K_{22}) \geq 0$ e $0 \leq (\pi_{12} + 2.K_{12}) \leq (\pi_{22} + 2.K_{22})$

(4) $|\det(P_q^{-1})| = |\pi_{11} + 2.K_{11}|.|\pi_{22} + 2.K_{22}| = \left(\prod_{i=1}^{n} q_i \right) / |\Gamma/D_2| = 2$

(5) $(\pi_{12} + 2.K_{12})$ e $(\pi_{22} + 2.K_{22})$ são primos-entre-si.

De (1), temos $\pi_{22} = 0$; logo, de (2), (3) e (4), concluímos que $K_{22} = 1$ e que $(p_{11} + 2.K_{11}) = 1$. Sendo $(\pi_{11} + 2.K_{11}) = 1$, concluímos, de (1), que $\pi_{11} = 1$ e $K_{11} = 0$, bem como $p_{12} = 1$. Sendo $\pi_{12} = 1$, concluímos, de (3), que $K_{12} = 0$. Com estes valores, verificamos que a condição (5) também é satisfeita; logo:

$$P_q^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$$

e portanto:

$$N = M.D(q)^{-1}.P_q^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$$

é a matriz geradora obtida para Γ. Como N é unitária, confirma-se que Γ é realmente igual ao reticulado $\mathbb{Z}^2$ do $\mathbb{R}^2$.

Como exemplo adicional, determinemos uma matriz geradora para um reticulado obtido por uma construção, denominada Construção A, que será analisada em capítulo posterior. Nesta construção, $\Lambda = 2\mathbb{Z}^n$ e $[\Gamma/\Lambda]_M$ é um código binário linear C de comprimento n e dimensão K (ou seja, $|C| = 2^K$). Assim, $M = 2.I_n$ (onde I_n é a matriz identidade de ordem n) e $M^{-1}.[\Gamma/\Lambda]_M = \frac{1}{2}.C$; logo, $D(q) = 2.I_n$ e $D(q).M^{-1}.[\Gamma/\Lambda]_M = C$ (ou seja, as classes laterais de Γ/Λ são representadas pelas palavras do código C, e estas são operadas em $[\Gamma/\Lambda]_M$ módulo 2);

logo $\left(\prod_{i=1}^{n} q_i\right) / |\Gamma/\Lambda| = 2^n/2^{n-K} = 2^K$. Como todo código C possui uma matriz geradora G n x K (tal que C é o conjunto de todas as combinações lineares módulo 2 das colunas de G) em forma sistemática, ou seja:

$$G = \begin{bmatrix} I_K \\ A \end{bmatrix}$$

onde I_K é a matriz identidade de ordem K, e A é uma matriz (n-K) x K única para C, sem linhas nulas, podemos tomar:

$$P_q^{-1} \text{MOD} q = \begin{bmatrix} I_K & 0 \\ A & 0 \end{bmatrix} K = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-K} \end{bmatrix}$$

de modo que:

$$P_q^{-1} = \begin{bmatrix} I_K & 0 \\ A & 2.I_{n-K} \end{bmatrix}$$

Observe-se que: (1) os elementos de cada linha de P_q^{-1} são primos-entre-si (todas as linhas possuem pelo menos um elemento igual a 1); (2) $\det(P_q^{-1}) = 2^K$; (3) P_q^{-1} está na forma canônica triangular. Como M = D(q), temos $N = D(q).M^{-1}.P_q^{-1} = P_q^{-1}$. Assim:

$$N = \begin{bmatrix} I_K & 0 \\ A & 2.I_{n-K} \end{bmatrix}$$

é uma matriz geradora do reticulado $C + 2.\mathbb{Z}^n$, onde:

$$G = \begin{bmatrix} I_K \\ \hline A \end{bmatrix}$$

é a matriz geradora de C em forma sistemática.

3.4.2 Forma canônica diagonal

Qualquer matriz quadrada A não-singular de números inteiros pode ser conduzida, por operações elementares sobre ambas linhas e colunas (correspondendo a multiplicações à direita e à esquerda de A por matrizes unitárias), à uma matriz diagonal única $\tilde{A}$, onde todos os elementos diagonais de $\tilde{A}$ são positivos com $\tilde{A}_{ii}$ divisor de $\tilde{A}_{(i+1)(i+1)}$ (onde $\tilde{A} = [\tilde{A}_{ij}]$), conhecidos como os **fatores invariantes** de A. A obtenção desta forma canônica diagonal, conhecida como **forma normal de Smith** ([26]) é assumida conhecida na discussão abaixo, podendo ser encontrada na literatura.

Retomemos, então, o problema da determinação de um conjunto de representantes $[\Gamma/\Lambda]$ para a partição de um reticulado Γ do $\mathbb{R}^n$, para o qual se conhece uma matriz geradora N, em classes laterais de um subreticulado $\Lambda \leq \Gamma$, para o qual também se conhece uma matriz geradora M, como foi estabelecido no início da Seção 2.2.

Analisemos, inicialmente, de que modo podemos reduzir a matriz particionadora P à forma canônica diagonal. Como $P_q^{-1} = N^{-1}.M$, a redução de P à forma canônica diagonal, correspondendo a multiplicações à direita e esquerda de P, equivale a mudanças de ambas as matrizes geradoras N e M, dos reticulados Γ e Λ, respectivamente. Em vista da existência e unicidade da forma canônica diagonal, concluímos que P pode sempre ser reduzida à forma canônica diagonal, adotando-se certas (e únicas) matrizes geradoras N e M, de Γ e Λ, respectivamente.

Vejamos, então, as simplificações, na solução do problema proposto na Seção 3.2, decorrentes da conclusão acima. Vimos que, dadas N e M, podemos obter $P = N^{-1}.M$ e D(q); assim, a matriz $P_q^{-1} = D(q).P^{-1}$ pode ser determinada. Vimos, também, que $P_q^{-1}MODq$ é uma matriz cujas colunas formam um conjunto gerador para o subgrupo $D(q).M^{-1}.[\Gamma/\Lambda]_M$ do grupo $\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn}$. Assim, $[\Gamma/\Lambda]_M = \Gamma MOD_M\Lambda$ pode ser determinado, solucionando o problema. Como mencionado naquela seção, a determinação de $D(q).M^{-1}[\Gamma/\Lambda]$, quando $\prod_{i=1}^{n} q_i$ é um múltiplo distinto de $|\Gamma/\Lambda|$, redunda em calcular cada elemento de $D(q).M^{-1}.[\Gamma/\Lambda]_M$ um número de vezes igual a $\left(\prod_{i=1}^{n} q_i\right)/|\Gamma/\Lambda|$, pois na expressão:

$$D(q).M^{-1}.[\Gamma/\Lambda]_M = \left\{((P^{-1}MODq).\mu)MODq : \mu \in \mathbb{Z}_m \text{ x...x } \mathbb{Z}_m\right\}$$

a variável $\mu \in \mathbb{Z}_m$ x...x $\mathbb{Z}_m$ pode assumir $\prod_{i=1}^{n} q_i$ valores vetoriais distintos, enquanto que $D(q).M^{-1}.[\Gamma/\Lambda]_M$ possui, evidentemente, apenas $|\Gamma/\Lambda|$ elementos distintos. No entanto, se P está na forma canônica diagonal, então D(q) = P, como é facilmente verificável, e, assim, temos $P_q^{-1} = P_q^{-1}MODq = I_n$ (onde I_n é a matriz identidade de ordem n) e $[\Gamma/\Lambda]_M = N.(\mathbb{Z}_{q1}$ x...x $\mathbb{Z}_{qn})$. A solução, portanto, pode vir a ser simplificada, caso a forma canônica diagonal de P possa ser encontrada facilmente (o procedimento para a sua determinação, tem complexidade dependente da particular matriz inicial a ser reduzida à sua forma canônica diagonal). Embora exista um procedimento mais simples para a determinação da forma canônica diagonal resultante, sem a determinação das operações elementares do procedimento completo de redução, observemos que é necessário a aplicação sobre N das operações elementares sobre as linhas de P utilizadas nesta redução, de modo que, com a matriz N resultante (associada à forma canônica diagonal de P) se possa determinar:

$$[\Gamma/\Lambda]_M = N.(\mathbb{Z}_{q1} x ... x \mathbb{Z}_{qn})$$

Como um exemplo ilustrativo, determinemos $[\mathbb{Z}^2/D_2]_M$ utilizando:

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

como matrizes geradoras de $\mathbb{Z}^2$ e D_2, respectivamente, como no exemplo ilustrativo da Seção 3.2, mas reduzindo inicialmente $P = N^{-1}.M$ à forma canônica diagonal. Seguindo o procedimento de redução de P, encontrada na literatura ([26]), e aplicando sobre N apenas as operações elementares aplicadas sobre as linhas de P, obtemos:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad N = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$$

Assim:

$$D(q) = P = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad M = N.P = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}$$

de modo que:

$$[\mathbb{Z}^2/D_2] = N.(\mathbb{Z}_1 x \mathbb{Z}_2) = N.\left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}\right\}$$

Encerramos, assim, essas considerações sobre a aplicação das formas canônicas, de matrizes não-singulares de números inteiros, aos problemas abordados neste capítulo.

Capítulo 4

Aplicação de reticulados a codificação para canais AWGN limitados em banda

4.1 Introdução

Este capítulo trata das potencialidades de aplicação de reticulados em esquemas de codificação para canais AWGN limitados em banda. Inicialmente, na Seção 4.2, é estabelecido o conceito de codificador para canais AWGN limitados em banda, e são analisados os critérios para sua avaliação e comparação. Na Seção 4.3, são, então, descritos os métodos correntes de construção baseada em reticulados desses codificadores, e analisadas as suas propriedades.

4.2 Códigos para canais AWGN limitados em banda

Introduziremos inicialmente uma representação vetorial para canais AWGN limitados em banda, adequada para a descrição de transmissão digital codificada.

Um canal AWGN limitado em banda é um canal contínuo que, em resposta a um sinal de entrada x(t), produz um sinal de saída r(t) dado por:

$$r(t) = u(t) + z(t)\ ,\ -1\infty < t < +\infty$$

onde u(t) é a resposta ao sinal de entrada x(t) de um filtro passa-baixa ideal com resposta em freqüência H(f) dada por:

$$H(f) = \begin{cases} 1\ , & |f| \le W \\ 0\ , & |f| > W \end{cases}$$

onde W ($0 < W < +\infty$) é denominada a largura de banda do canal e z(t) é uma amostra (forma de onda) de um processo aleatório Z(t) gaussiano (independente do sinal de entrada x(t)) com média zero e densidade espectral de potência $S_Z(f)$ dada por:

$$S_Z(f) = \frac{N_0}{2}\ ,\ -\infty < f < +\infty.$$

A Fig. 4.2.1 ilustra o conceito de canal AWGN limitado em banda.

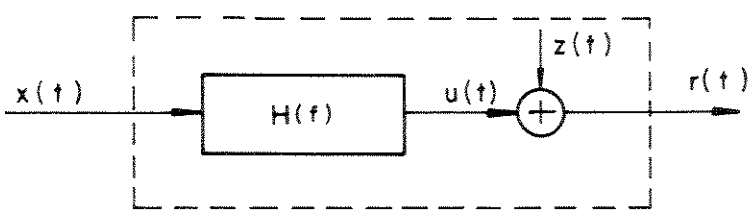

Fig. 4.2.1 - Canal AWGN limitado em banda

A utilização de um canal arbitrário para transmissão digital requer (1) o uso seqüencial do canal a uma taxa de sinalização fixa (do inglês "baud rate") de, digamos, um uso do canal a cada T segundos; (2) a seleção de um alfabeto finito de símbolos de entrada, adequadamente

selecionados dentro do espaço de entrada do canal, de onde, a cada uso do canal, é escolhido um símbolo para transmissão. No caso de um canal AWGN limitado em banda, o espaço de entrada é o espaço de todos os possíveis sinais x(t), $-\infty < t < +\infty$, e o alfabeto de símbolos de entrada será, então, um conjunto finito de sinais.

A primeira restrição que, de imediato, surge para uma seleção adequada do alfabeto de símbolos de entrada é, obviamente, que cada símbolo tenha energia finita, ou seja, se x(t), $-\infty < t < +\infty$, for um símbolo do alfabeto de entrada, devemos ter:

$$\varepsilon_x \triangleq \int_{-\infty}^{+\infty} x(t)^2 . dt < +\infty$$

Uma outra possível restrição seria a facilidade de implementação. Uma restrição bem mais difícil (a rigor, impossível) de ser satisfeita é a ausência de interferência intersimbólica na saída do canal: qualquer que seja o sinal x(t), $-\infty < t < +\infty$, na Fig. 4.2.1, o sinal u(t) é limitado em freqüência à banda (-W,W) e, portanto, necessariamente não é limitado em tempo ao intervalo (– T/2, T/2); assim, qualquer que seja o alfabeto de símbolos selecionados, a transmissão de um símbolo em um intervalo de sinalização interfere com os símbolos em outros intervalos de sinalização; portanto, rigorosamente, qualquer canal AWGN limitado em banda introduz interferência entre símbolos.

No entanto, existem sinais que são limitados em tempo ao intervalo (– T/2, T/2) e "essencialmente" limitados em freqüência à banda (-W,W). Especifiamente, é possível determinar, para cada valor de T e de W, um conjunto $\{\varphi_i(t)\}$ de $n \approx 2WT$ sinais ortonormais:

$$\int_{-\infty}^{+\infty} \varphi_i(t)\, \varphi_j(t) = \begin{cases} 1 , & i=j \\ 0 , & i \neq j \end{cases}$$

para $1 \leq i, j \leq n$, tal que o espaço de sinais, gerado por suas combinações lineares, é constituído de sinais limitados em tempo ao intervalo (– T/2, T/2) e com a maior parte de sua energia concentrada na banda (-W,W). Assim, selecionando-se o alfabeto de símbolos de entrada dentro deste espaço de sinais, cada símbolo, ao ser transmitido em um intervalo de sinalização, passa essencialmente sem distorção pela limitação de freqüência imposta pelo canal (por ser essencialmente limitado em freqüência à banda (-W,W)) e não interfere com os simbolos em outros intervalos de sinalização (por ser limitado em tempo ao intervalo (– T/2, T/2)). (Outras alternativas seriam sinais limitados em freqüência e essencialmente limitados em tempo, ou ainda sinais essencialmente limitados em tempo e essencialmente limitados em freqüência; no entanto, verifica-se que tais sinais são essencialmente idênticos a sinais do espaço gerado por $\{\varphi_i(t)\}$, para cada valor de T e de W.

Vemos, então, que a restrição da ausência de interferência intersimbólica na saída do canal pode ser essencialmente satisfeita, reduzindo, assim, o comportamento do canal AWGN limitado em banda ao de um canal AWGN sem limitação de banda, com o mesmo alfabeto de entrada e taxa de sinalização. (É interessante notar-se que, enquanto, para um canal AWGN sem

limitação de banda, os símbolos do alfabeto de entrada do canal podem ser confortavelmente dispostos em um espaço de sinais de dimensão arbitrariamente alta, temos, para um canal AWGN limitado em banda a W, que acomodar os símbolos do alfabeto de entrada do canal em um espaço de sinais de dimensão limitada, aproximadamente, em 2WT, onde T é o período de sinalização escolhido para ambos os canais; isto traduz-se, na prática, numa limitação do número de graus de liberdade na seleção de cada símbolo do alfabeto de entrada, que é uma característica própria da limitação em banda).

Em particular, podemos representar o canal contínuo original da Fig. 4.2.1 pelo canal vetorial gaussiano discreto no tempo (sem memória) da Fig. 4.2.2. Nesta representação, $x = [x_1,...,x_n]^T$ é um vetor real que representa um símbolo x(t) do alfabeto de entrada, dado por:

$$x(t) = \sum_{i=1}^{n} x_i \varphi(t) \ , \ -\infty < t < +\infty$$

pertencente ao espaço de sinais gerado por $\{\varphi_i(t)\}$; note-se, então, que x(t) é limitado em tempo ao intervalo (– T/2, T/2), e que, devido à ortonormalidade de $\{\varphi_i(t)\}$,

$$\varepsilon_x = \int_{-T/2}^{T/2} x(t)^2.dt = \sum_{i=1}^{n} x_i^2 \triangleq \|x\|^2 \qquad (4.2.1)$$

onde $\|x\|^2$ é a norma do vetor $x \in \mathbb{R}^n$; $z = [z_1,...,z_n]^T$ é um vetor cujas coordenadas são amostras de n variáveis aleatórias Z_i, $1 \le i \le n$, gaussianas independentes (e independentes do vetor x da entrada) de média $\mu_i=0$ e variância $\sigma_i^2 = N_0/2$, com densidade de probabilidade p(z) dada, portanto, por:

$$p(z) = \frac{1}{(\pi N_0)^{n/2}} . \text{Exp}\left[- \frac{\|z\|^2}{N_0} \right] , z \in \mathbb{R}^n$$

Finalmente, $r = [r_1,...,r_n]^T$ é um vetor obtido, simplesmente, por r = x + z, de modo que a densidade de probabilidade do ponto recebido r, condicionada à transmissão do ponto x, é dada por:

$$p(r/x) = \frac{1}{(\pi N_0)^{n/2}} . \text{Exp}\left[- \frac{\|r-x\|^2}{N_0} \right] \qquad (4.2.2)$$

Como $\varphi_i(t) = 0$ para $|t| > T/2$, $1 \le i \le n$, as componentes r_i do vetor recebido r são obtidas, na prática, pela filtragem de r(t) por um filtro casado a $\varphi_i(t)$, seguida por amostragem, evitando-se assim a implementação direta da correlação $r_i = \int_T r(t).\varphi_i(t).dt$ que requer a utilização de multiplicadores analógicos de precisão. Portanto, nesta representação, cada sinal no alfabeto de entrada do canal AWGN limitado em banda é representado por um ponto do $\mathbb{R}^n$ que, quando

transmitido pelo canal gaussiano discreto no tempo, sofre perturbação pela adição de um vetor de ruído, produzindo então um ponto do $\mathbb{R}^n$ na saída do canal, que representa o sinal de saída do canal AWGN limitado em banda. (A rigor, os vetores de ruído z e de saída r representam, respectivamente, as projeções de z(t) e r(t) no espaço de sinais gerado por $\{\varphi_i(t)\}$, não sendo, portanto, representações precisas; verifica-se, no entanto, que são estatisticamente suficientes para a estimação ótima de uma seqüência de pontos transmitidos, adotando-se como critério de otimização a maximização da probabilidade a posteriori da seqüência estimada).

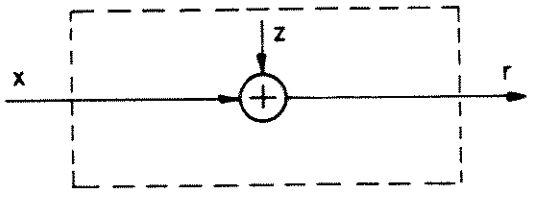

Fig. 4.2.2 - Canal gaussiano discreto no tempo

Assim, a utilização de um canal vetorial gaussiano discreto no tempo, de dimensão n a uma taxa de sinalização de um ponto a cada T segundos, corresponde à utilização de um canal AWGN limitando em banda, a uma taxa de sinalização de um sinal a cada T segundos, com largura de banda $W \approx n/2T$. Da mesma forma, um conjunto finito de pontos, no $\mathbb{R}^n$, adotado como sendo o alfabeto de símbolos de entrada para um canal gaussiano discreto no tempo, que denominaremos uma **constelação n-dimensional**, corresponde a um conjunto de sinais, no espaço de sinais gerado por $\{\varphi_i(t)\}$, adotado como sendo o alfabeto de simbolos de entrada para um canal AWGN limitado em banda a $W \approx n/2T$, onde T é o período de sinalização para ambos os canais.

Em particular, podemos conceituar um **codificador** com taxa de sinalização de um sinal a cada T segundos, para o canal AWGN limitado em banda, com largura de banda W e densidade espectral de potência uniforme $N_0/2$, em termos da conceituação de um codificador, com taxa de sinalização de um ponto a cada T segundos, para o canal vetorial gaussiano discreto no tempo, de dimensão $n \approx 2WT$ e variância de ruído por dimensão $\sigma_i^2 = N_0/2$, $1 \leq i \leq n$. Isto permite que, uma vez que seja determinado um codificador, para um canal gaussiano discreto no tempo, seja possível, em princípio, implementar um codificador, para um correspondente canal AWGN limitado em banda, com operação essencialmente livre de interferência intersimbólica e, portanto, com mesmo desempenho e complexidade. Além disso, como em breve tornar-se-á evidente, o desempenho e a complexidade de codificadores para um canal vetorial gaussiano discreto não dependem do período de sinalização T, mas de sua dimensão $n \approx 2WT$, dando assim liberdade para implementação de código de mesmo desempenho e complexidade para canais AWGN limitados em banda, com larguras de banda W de conveniência prática, ajuustando-se adequadamente o período de sinalização T e mantendo-se constante o produto $WT \approx n/2$, com a precisão permissível. (Note-se que para cada valor de T e de W, os $n \approx 2WT$ sinais do conjunto $\{\varphi_i(t)\}$ devem ser correspondentemente alterados; assim uma mesma constelação n-dimensional corresponde, em princípio, a distintos conjuntos de sinais, para seleções distintas de T e de W, com T.W constante e igual a aproximadamente n/2).

Passemos, então, à conceituação de codificador para um canal vetorial gaussiano discreto no tempo: é um dispositivo que, no decorrer de sua operação, percorre uma seqüência (s^i) de estados em um conjunto finito S de estados possíveis, tem como entrada K seqüências $(a_1^i),...,(a_K^i)$ de símbolos de um alfabeto finito A de símbolos de informação, e tem como saída uma seqüência (x^i) de pontos de uma constelação finita C de pontos do $\mathbb{R}^n$; no i-ésimo intervalo de sinalização, produz um ponto $x^i \in C$ de saída e executa uma transição para um próximo estado $s^{i+1} \in S$, que dependem do estado atual $s^i \in S$ e da K-upla $a^i = [a_1^i,...,a_K^i]^T \in A^K$ de entrada:

$$x^i = x(a^i, s^i)$$

$$s^{i+1} = s(a^i, s^i)$$

onde x(.,.) e s(.,.) são funções fixas de A^K x S em C e S, respectivamente; diz-se, então, que o codificador é **invariante no tempo** (pode-se, obviamente, estender esta definição a codificadores que variem no tempo, fazendo as funções de próximo estado e de saída dependentes, de alguma forma, da ordem i do intervalo de sinalização; nestes casos, quase sempre esta dependência é cíclica). Assumiremos, definitivamente, que as K-uplas de entrada são independentes e uniformemente distribuídas sobre A^K; diremos, então, que o codificador envia, pelo canal, $\log_2|A|^K = K.\log_2|A|$ dígitos binários (bits) de informação a cada período de sinalização de T segundos, ou ainda, a cada ponto n-dimensional transmitido. A Fig. 4.2.3 ilustra o conceito de codificador para um canal vetorial gaussiano discreto no tempo.

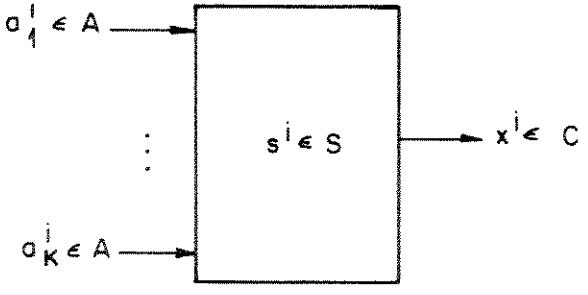

Fig. 4.2.3 - Codificador para canal gaussiano discreto no tempo

Para ilustrar os critérios de comparação e os esquemas de construção de codificadores, restringiremos nossas considerações àqueles correntemente mais utilizados. Um destes (cuja análise é relativamente complexa, mas com decodificação especialmente simples) é o denominado **codificador em treliça**: é um codificador tal que seu estado s^i, no i-ésimo intervalo de sinalização, pode ser representado pela ν-upla $(a_1^{i-1},...,a_1^{i-\nu_1};...;a_K^{i-1},...,a_K^{i-\nu_K})$, de modo que a função de próximo estado s(.,.) pode ser implementada com exatamente $\nu = \sum_{k=1}^{K} \nu_k$ elementos de memória |A|-ários dispostos na forma de K registradores de deslocamento, como na Fig. 4.2.4, onde $s^{i+1} = s(a^i, s^i)$ será o conteúdo dos registradores após a operação de deslocamento, ao final do i-ésimo intervalo de sinalização. (É interessante notar-se que cada estado é determinado pelas

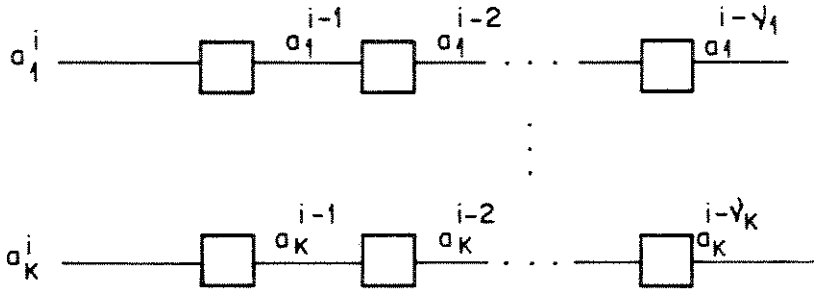

Fig.4.2.4 - Estrutura de Memória do Codificador em Treliça.

últimas $m \triangleq MAX_{1 \le k \le K} \nu_k$ K-uplas de entrada $a^{i-1},...,a^{i-m}$). Assim o número de estados possíveis é $|S| = |A|^{\nu}$, e as possíveis evoluções no tempo do estado do codificador podem ser representados, de uma maneira eficiente, por uma treliça orientada periódica (com nós e ramos representando estados e transições de estado, respectivamente) cujo período é constituído por uma única seção formada por todas as possíveis transições de estado em um período de sinalização. Diz-se que a k-ésima entrada, $1 \le k \le K$, e não-codificada se $\nu_k = 0$; se existirem K_1 entradas não-codificadas, as transições de estado, em cada seção da treliça, estarão agrupadas em feixes de $|A|^{K_1}$ transições distintas paralelas (i.e., de mesmo estado atual e mesmo próximo estado) e existirão $|A|^{K_2}$ destes feixes partindo de cada estado atual, bem como chegando em cada próximo estado, onde $K_2 = K - K_1$. Em qualquer caso, no entanto, tem-se $|A|^K$ transições de estado partindo de cada estado atual, bem como chegando em cada próximo estado, em cada seção da treliça. Note-se que a estrutura da treliça de um codificador em treliça (ou seja, a especificação de sua função de próximo estado s(.,.)) depende apenas dos valores de K e ν_k, $1 \le k \le K$, além obviamente do alfabeto de entrada A. A Fig. 4.2.5 ilustra uma seção da treliça com $A = \{0,1\}$, $K = 2$, $\nu_1 = 2$ e $\nu_2 = 0$.

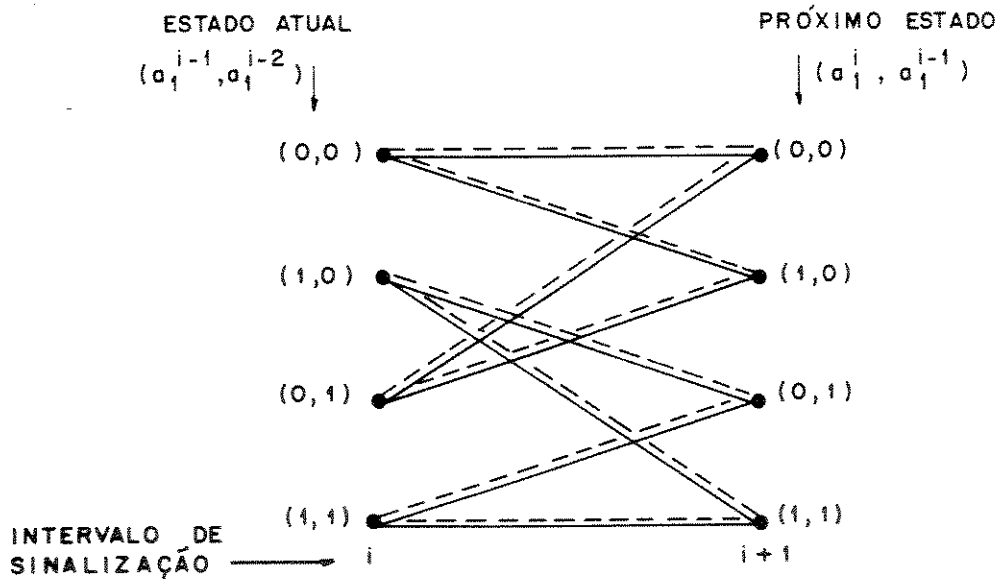

Fig. 4.2.5 - Estrutura de uma treliça com A = {0,1}, K = 2, $\nu_1 = 2$ e $\nu_2 = 0$. (——— $a_2 = 1$; - - - - - - - - - - $a_2 = 0$)

Uma vez especificada a função de próximo estado s(.,.) de um codificador em treliça, resta especificar sua função de saída x(.,.); isto equivale, obviamente, a associar a cada transição, de uma seção de sua treliça, um ponto da constelação C, mantendo esta mesma associação em todas seções; por definição, cada ponto da constelação C de um codificador deve ser utilizado pelo menos uma vez. Note-se que, embora exista um toral de $|A|^{K+\nu}$ transições de estado em cada seção da treliça, a constelação não necessariamente possui $|A|^{K+\nu}$ pontos distintos, pois um mesmo ponto de C pode estar associado a transições distintas. Em qualquer caso, no entanto, podemos expressar o ponto $x^i = x(a^i,s^i)$ transmitido no i-ésimo intervalo de sinalização, como uma função, das $K + \nu$ variáveis a_k^i, a_k^{i-1},..., $a_k^{i-\nu_k}$, $1 \le k \le K$:

$$x^i = x(a_1^i,...,a_1^{i-\nu_1};...;a_K^i,...a_K^{i-\nu_K})$$

sobrejetiva de $A^{K+\nu}$ em C, de modo que a potência média S transmitida pelo codificador (para K-uplas de entrada independentes e uniformemente distribuídas sobre A^K) é igual a

$$S = \frac{1}{T} \cdot \frac{1}{|A|^{K+\nu}} \sum \|\tilde{x}(a_1,...,a_{K+\nu})\|^2 \qquad (4.2.3)$$

onde a somatória se estende a todas as $(K+\nu)$-uplas $|A|$-árias $(a_1,...,a_{K+\nu})$, com $a_j \in A$, $1 \le j \le K + \nu$. Assim, um particular codificador em treliça fica completamente especificado.

A operação do codificador em treliça pode, então, ser interpretada como sendo a seleção, pela seqüência (a^i) de K-uplas de informação, de um caminho através de sua treliça constituído de uma seqüência contínua de transições de estado, determinando, desta forma, uma seqüência (x^i) pontos da constelação a ser transmitida pelo canal. A seqüência de pontos transmitida pelo canal é perturbada pela adição de uma seqüência (z^i) de vetores de ruído, recebendo-se uma seqüência (r^i) de pontos na saída do canal, possivelmente distinta da seqüência transmitida. Um dispositivo na recepção, denominado decodificador, deve primeiro determinar, baseando-se na seqüência recebida, uma estimativa $(\hat{x}^i)$ da seqüência transmitida, utilizando algum critério de estimação adequado, para então inferir a correspondente estimativa $(\hat{a}^i)$ da seqüência de informação.

Mais especificamente, seja $a_{[0,L)} = [a^0,a^1,...,a^{L-1}]$ uma seqüência finita, de comprimento L, de K-uplas $a^i \in A^K$, $0 \le i \le L - 1$, de informação; existem $|A|^{K.L}$ tais seqüências, todas com a mesma probabilidade de ocorrência (para K-uplas de entrada independentes e uniformemente distribuídas sobre A^K). O codificador deve produzir uma seqüência única de pontos $x_{[0,L)} = [x^0,x^1,...x^{L-1}]$ para cada uma dessas seqüências; como estados iniciais $s^0 \in S$ distintos conduzem a seqüências de pontos $x_{[0,L)}$ distintas, devemos pré-estabelecer um estado inicial definitivo para o codificador, que pode ser tomado como sendo o estado em que todos os elementos de memória $|A|$-ários, dos K registradores de deslocamento da Fig. 4.2.4, armazenam o mesmo elemento de A, que podemos, sem perda de generalidade, representar por 0; assim o estado inicial pode ser representado pela ν-upla de elementos 0. Seja $m = MAX_{1 \le k \le K} \nu_k$; de modo que cada K-upla de informação de entrada influencie um mesmo número m + 1 (denominado o **comprimento de restrição** do codificador em treliça) de pontos de saída, é conveniente que a seqüência de K-uplas de entrada seja seguida por m K-uplas pré-estabelecidas, forçando, então, o codificador a um estado final pré-estabelecido definitivo,

que podemos adotar como sendo igual ao estado inicial, $s^{L+m} = s^0$ fazendo cada uma daquelas K-uplas adicionais iguais à K-upla de elementos 0; assim, o codificador produzirá m pontos adicionais $x^L,...,x^{L+m-1}$ após a seqüência $x_{[0,L)}$, formando-se então a seqüência transmitida $x_{[0,L+m)}$. Em resumo, o codificador produz uma seqüência de pontos $x_{[0,L+m)}$ (de comprimento L + m), em resposta a uma seqüência de K-uplas de entrada $a_{[0,L)}$ (de comprimento L), partindo do estado representado pela ν -upla de elementos 0 e retornando a este estado pela inserção de uma seqüência de K-uplas de elementos 0 (de comprimento m) após a seqüência de K-uplas de entrada $a_{[0,L)}$. Cada seqüência de pontos $x_{[0,L+m)}$ transporta $K.L.\log_2|A|$ bits de informação, em (L + m) intervalos de sinalização, incorrendo-se, portanto, em uma perda relativa de $m/(L + m)$ no número de bits de informação por intervalo de sinalização, decorrente da necessidade de re-sincronização de volta ao estado inicial; esta perda relativa, no entanto, será negligenciável se L >> m, sendo isto, de fato, adotado na prática. Para um codificador cuja treliça seja a da Fig. 4.2.5, temos m = 2 e a codificação de uma seqüência finita de comprimento L >> m de K-uplas de entrada $a^i \in A^K$ (K = 2, A = {0,1}) pode ser interpretada como um caminho, desde o estado inicial $s^0 = (0,0)$ até o estado final $s^{L+2} = (0,0)$ na treliça finita da Fig. 4.2.6.

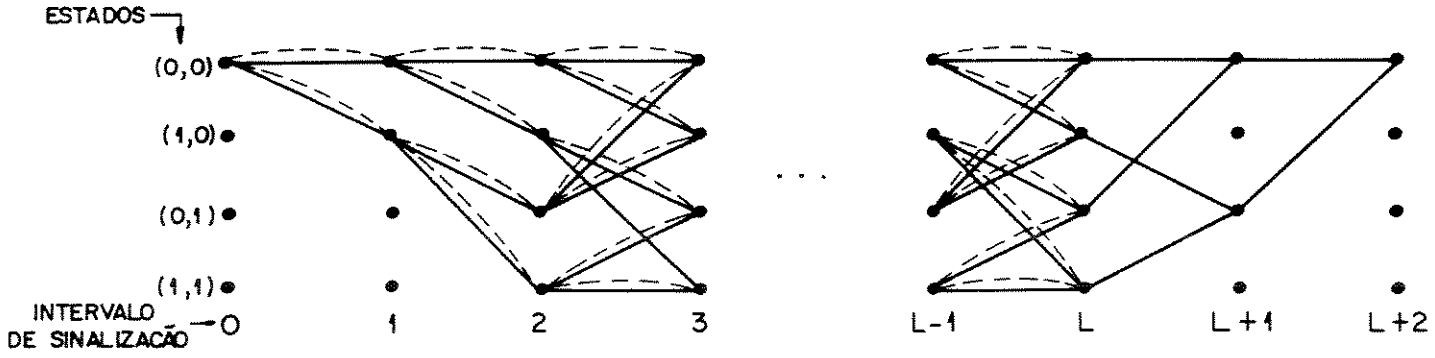

Fig. 4.2.6 - Treliça finita para codificação de seqüência de comprimento L >> m = 2

A seqüência finita de pontos $x_{[0,L+m)} = [x^0,x^1,...,x^{L+m-1}]$ é transmitida pelo canal e uma seqüência finita de pontos $r_{[0,L+m)} = [r^0,r^1,...,r^{L+m-1}]$ é recebida na saída do canal. O decodificador deve, então, selecionar uma das $|A|^{K.L}$ possíveis seqüências de pontos, de comprimento L + m que partem do estado $s^0 = (0,...,0)$ e retornam ao estado $s^{L+m} = (0,...,0)$, como uma estimativa $\hat{x}_{[0,L+m)} = [\hat{x}^0,\hat{x}^1,...,\hat{x}^{L+m-1}]$ da seqüência transmitida, baseada na seqüência recebida, utilizando algum critério da estimação adequado. O critério usualmente utilizado é o da minimização da probabilidade de erro de estimação da seqüência transmitida, que, obviamente, equivale à maximização da probabilidade a posteriori da estimativa $\hat{x}_{[0,L+m)}$, dada a seqüência recebida $r_{[0,L+m)}$; ou seja, $\hat{x}_{[0,L+m)}$ é a seqüência de pontos, em um caminho da treliça do codificador partindo de $s^0 = (0,...,0)$ e retornando a $s^{L+m} = (0,...,0)$, que maximiza:

$$p(\hat{x}_{[0,L+m)}|r_{[0,L+m)}) = \frac{p(r_{[0,L+m)}|\hat{x}_{[0,L+m)}) \cdot p(\hat{x}_{[0,L+m)})}{p(r_{[0,L+m)})}$$

ou (já que $p(r_{[0,L+m)})$ não depende de $\hat{x}_{[0,L+m)}$) que maximiza:

$$p(r_{[0,L+m)}|\hat{x}_{[0,L+m)}) \cdot p(\hat{x}_{[0,L+m)}) \tag{4.2.4}$$

Um decodificador que adote este critério é denominado um decodificador MAP (do inglês Maximum A Posteriori Probability). Com a nossa suposição usual de que as K-uplas de entrada do codificador são independentes e uniformemente distribuídas sobre A^K, temos que todas as $|A|^{K.L}$ possíveis seqüências de pontos, de comprimento L + m partindo de $s^0 = (0,...,0)$ e retornando a $s^{L+m} = (0,...,0)$, têem a mesma probabilidade de ocorrência; assim, o fator $p(\hat{x}_{[0,L+m)})$ em (4.2.4) pode ser desconsiderado na estimação, e o decodificador deve determinar a estimativa $\hat{x}_{[0,L+m)}$ que maximiza:

$$p(r_{[0,L+m)}|\hat{x}_{[0,L+m)}) \tag{4.2.5}$$

Um decodificador que adote este critério é denominado um decodificdor ML (do inglês Maximum Likelihood). A densidade de probabilidade $p(r_{[0,L+m)}|\hat{x}_{[0,L+m)})$ pode ser fatorada (em vista da ausência de memória no canal) como:

$$p(r_{[0,L+m)}|\hat{x}_{[0,L+m)}) = \prod_{i=0}^{L+m-1} p(r^i|\hat{x}^i)$$

onde, em vista de (4.2.2), tem-se:

$$p(r^i|\hat{x}^i) = \frac{1}{(\pi N_0)^{n/2}} \cdot \mathrm{Exp}\left[-\frac{\|r^i-\hat{x}^i\|^2}{N_0}\right]$$

Assim, maximizar (4.2.5) equivale a minimizar:

$$-\ln p(r_{[0,L+m)}|\hat{x}_{[0,L+m)})^{N_0} = (L+m) \cdot \ln(\pi N_0)^{n.N_0/2} + \sum_{i=0}^{L+n-1} \|r^i-\hat{x}^i\|^2 \tag{4.2.7}$$

já que $-\ln y^{N_0}$ é uma função estritamente decrescente de y (pois $N_0 > 0$). Como a primeira parcela do lado direito de (4.2.7) não depende de $\hat{x}_{[0,L+m)}$, tem-se, finalmente, que: a probabilidade de erro de estimação da seqüência transmitida é minimizada selecionando-se a estimativa $\hat{x}_{[0,L+m)}$ como sendo a seqüência de pontos, em um caminho da treliça do codificador partindo de $s^0 = (0,...,0)$ e retornando a $s^0 = (0,...,0)$ que minimiza a seguinte soma, denominada distância entre seqüências:

$$\sum_{i=0}^{L+m-1} \|r^i-\hat{x}^i\|^2 \tag{4.2.8}$$

Assim, um decodificador ML para um codificador de canal gaussiano discreto no tempo seleciona sua seqüência estimada como sendo a seqüência codificada mais próxima (em termos de distância quadrática Euclideana) da seqüência recebida, sendo, por isso, também denominado decodificador de **distância mínima**. Um decodificador ML implementado com precisão infinita, ou seja, sem quantização das componentes de r^j e sem truncamento no cálculo das distâncias quadráticas $\|r^j - x^j\|^2$, é dito operar por **decisão suave**; assumimos, definitivamente, este modo de operação para os decodificadores ML, ao compararmos o desempenho e a complexidade de codificadores específicos.

Vistos o conceito de codificador em treliça e sua decodificação ótima, passemos a considerar os aspectos de desemepenho de codificadores específicos. Obviamente, para um dado canal gaussiano discreto no tempo, um codificador será tanto melhor quanto maior for, para uma dada potência média, para um dado número de bits por intervalo da sinalização e para uma dada complexidade (que definiremos posteriormente), sua imunidade ao ruído. A primeira vista, uma possível medida de imunidade ao ruído seria a probabilidade do evento de estimar-se erroneamente a seqüência transmitida; no entanto, vimos que as seqüências transmitidas devem ser muito longas ($L >> m$) de modo que torne-se negligenciável a perda relativa $m/(L + m)$ no número de bits de informação transmitidos, por intervalo de sinalização, decorrente da re-sincronização do codificador de volta ao estado inicial; como para $L >> m$, torna-se muito próxima da unidade a probabilidade de que pelo menos um ponto da seqüência estimada ótima seja distinto do ponto correspondente transmitido, conclui-se que esta medida de imunidade ao ruído não é adequada para avaliarmos o desempenho de codificadores em treliça. De fato, para $L >> m$, a seqüência estimada ótima $\hat{x}_{[0,L+m)}$ pode diferir da seqüência transmitida $\tilde{x}_{[0,L+m)}$ em vários segmentos de tempo, cada um destes podendo ter duração de vários intervalos de sinalização, como ilustrado na Fig. 4.2.7 para o codificador cuja treliça finita seja a da Fig. 4.2.6; diz-se, então, que ocorreu um **evento de erro** no primeiro intervalo de sinalização de cada um desses segmentos de tempo.

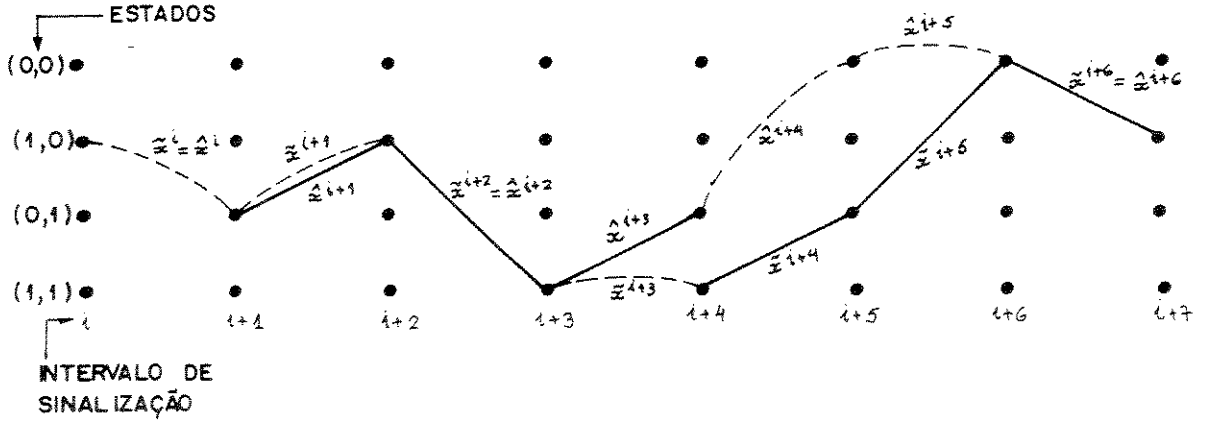

Fig. 4.2.7 - Configuração típica das seqüências transmitida ($\tilde{x}$) e estimada ($\hat{x}$)

Mais precisamente, ocorre um evento de erro, no j-ésimo intervalo de sinalização ($0 \le j \le L$), se $\hat{s}^j = \tilde{s}^j$ e $\hat{x}^j \neq \tilde{x}^j$; além disso, diz-se que o evento de erro, ocorrido no j-ésimo intervalo de

sinalização, perdura até o início do (j + l)-ésimo intervalo de sinalização se $\hat{s}^i \neq \tilde{s}^i$ para $j+1 \leq i \leq j+l-1$, e $s^{j+l} = s^{j+l}$, tendo, então, comprimento igual a l intervalos de sinalização. Assim, na Fig. 4.2.7, ocorreram eventos de erro nos intervalos i+1 e i+3, de comprimentos iguais a 1 e 3 intervalos de sinalização, respectivamente. Obviamente, tem-se sempre que $l = 1$ ou $l \geq m+1$. Isto sugere que uma medida da imunidade ao ruído, mais adequada para avaliar-se o desempenho de codificadores em treliça, pode ser definida como a probabilidade de, no j-ésimo intervalo de sinalização, ocorrer um evento de erro, condicionada à ocorrência do evento $\tilde{s}_j = \hat{s}_j$, denominada **probabilidade de evento de erro** (condicionada ao evento $\tilde{s}_j = \hat{s}_j$), no j-ésimo intervalo de sinalização, denotada por $P_e(j)$. A obtenção de uma expressão exata para $P_e(j)$ é geralmente inviável, utilizando-se, então, um limitante superior como aproximação para $P_e(j).(\hat{x}^i)$

Consideremos uma seqüência transmitida específica, digamos $(\bar{x}^i)$, que denominamos seqüência de referência. Seja $(x^i)_{[j,j+l)}$ um segmento de seqüência, de comprimento l, na treliça do codificador, que seja, para a seqüência de referência $(\bar{x}^i)$, um **possível evento de erro**, no j-ésimo intervalo de sinalização; ou seja, que satisfaça às condições: $s^j = \bar{s}^j$, $x^j \neq \bar{x}^j$, $s^{j+l} = \bar{s}^{j+l}$ e $s^i \neq \bar{s}^i$ para $j+1 \leq i \leq j+l-1$. A probabilidade de que a distância entre os segmentos de seqüência $(x^i)_{[j,j+l)}$ e $(r^i)_{[j,j+l)}$ seja menor ou igual do que a distância entre os segmentos de seqüência $(\bar{x}^i)_{[j,j+l)}$ e $(r^i)_{[j,j+l)}$, dada a seqüência de referência $(\bar{x}^i)$, é expressa por:

$$P\left[\|(x^i)_{[j,j+l)} - (r^i)_{[j,j+l)}\|^2 \leq \|(\bar{x}^i)_{[j,j+l)} - (r^i)_{[j,j+l)}\|^2 \,|\, (\bar{x}^i)\right] =$$

$$= P\left[\sum_{i=j}^{j+l-1} \|x^i - r^i\|^2 \leq \sum_{i=j}^{j+l-1} \|\bar{x}^i - r^i\|^2 \,|\, (\bar{x}^i)\right] =$$

$$= P\left[2.\sum_{i=j}^{j+l-1} \langle(\bar{x}^i - x^i), r^i\rangle \leq \sum_{i=j}^{j+l-1} (\|\bar{x}^i\|^2 - \|x^i\|^2) \,|\, (\bar{x}^i)\right]$$

onde $\|(a^i)_{[j,j+l)} - (b^i)_{[j,j+l)}\|^2 \triangleq \sum_{i=j}^{j+l-1} \|a^i - b^i\|^2$ é a distância Euclideana quadrática entre os segmentos de seqüência $(a^i)_{[j,j+l)} = [a^j,...,a^{j+l-1}]$ e $(b^i)_{[j,j+l)} = [b^j,...,b^{j+l-1}]$.

Como, dada a seqüência de referência $(\bar{x}^i)$, os pontos r^i são independentes, cada um com densidade de probabilidade como em (4.2.2), temos que:

$$y = 2 \cdot \sum_{i=j}^{j+l-1} \langle(\bar{x}^i - x^i), r^i\rangle$$

terá densidade de probabilidade, dada a seqüência de referência $(\bar{x}^i)$, expressa por:

$$p(y|(\overrightarrow{x^i})) = \frac{1}{\sqrt{2\pi\,\sigma_y^2}} \cdot \mathrm{Exp}\left[-\frac{(y-\mu_y)^2}{2.\sigma_y^2}\right]$$

onde

$$\sigma_y^2 = 2.N_0 \cdot \sum_{i=j}^{j+l-1} \|\overrightarrow{x^i} - x^i\|^2$$

$$\mu_y = 2 \cdot \sum_{i=j}^{j+l-1} \langle(\overrightarrow{x^i} - x^i), \overrightarrow{x^i}\rangle$$

Assim, sendo $\alpha = \sum_{i=j}^{j+l-1} (\|\overrightarrow{x^i}\|^2 - \|x^i\|^2)$, temos:

$$P\left[\|(x^i)_{[j,j+l)} - (r^i)_{[j,j+l)}\|^2 \le \|(\overrightarrow{x^i})_{[j,j+l)} - (x^i)_{[j,j+l)}\|^2 | (\overrightarrow{x^i})\right] =$$

$$= \int_{-\infty}^{\alpha} p(y|(\overrightarrow{x^i})).dy = Q\left(-\frac{\alpha+\mu_y}{\sigma_y}\right) = Q\left(\left[\frac{1}{2N_0} \cdot \sum_{i=j}^{j+l-1} \|\overrightarrow{x^i} - x^i\|^2\right]^{1/2}\right) =$$

$$= Q\left(\left[\frac{1}{2N_0} \cdot \|(\overrightarrow{x^i})_{[j,j+l)} - (x^i)_{[j,j+l)}\|^2\right]^{1/2}\right) \qquad (4.2.8)$$

Omitindo-se, para simplificar, o subscrito $[j,j+l)$ indicativo do intervalo de duração do possível evento de erro (no j-ésimo intervalo de sinalização e de comprimento igual a l intervalos), podemos reescrever (4.2.8) como:

$$P\left[\|(x^i) - (r^i)\|^2 \le \|(\overrightarrow{x^i}) - (r^i)\|^2 \,|\, (\overrightarrow{x^i})\right] = Q\left(\left[\frac{1}{2N_0} \cdot \|(\overrightarrow{x^i}) - (x^i)\|^2\right]^{1/2}\right) \qquad (4.2.9)$$

Assim, a probabilidade de que a distância Euclideana quadrática $\|(x^i) - (r^i)\|^2$ seja menor ou igual à distância Euclideana quadrática $\|(\overrightarrow{x^i}) - (r^i)\|^2$, dada a seqüência de referência $(\overrightarrow{x^i})$, depende exclusivamente de $\|(\overrightarrow{x^i}) - (x^i)\|^2$, a distância Euclideana quadrática do possível evento de erro (x^i) à seqüência de referência $(\overrightarrow{x^i})$. Esta probabilidade, conhecida como a **probabilidade do par** $[(\overrightarrow{x^i}) \rightarrow (x^i)]$, será útil para a obtenção do limitante superior para $P_e(j)$.

Antes, porém, analisemos os possíveis eventos de erro que podem ocorrer. Definimos, para cada possível seqüência de referência $(\overrightarrow{x^i})$: o conjunto $E_j(\overrightarrow{x^i})$ de todos os possíveis eventos de erro para $(\overrightarrow{x^i})$, no j-ésimo intervalo de sinalização $(0 \le j \le L)$; o conjunto $D_j(\overrightarrow{x^i})$ das distâncias

Euclideanas quadráticas de todos os elementos de $E_j(\overline{x}^i)$ a seqüência de referência $(\overline{x}^i)$; o subconjunto $E_j^{d^2}(\overline{x}^i)$, para cada $d^2 \in D_j(\overline{x}^i)$, de todos os elementos de $E_j(\overline{x}^i)$ cujas distâncias Euclideanas quadráticas a seqüência de referência $(\overline{x}^i)$ seja igual a d^2; o número $N_{d^2}^j(\overline{x}^i)$, para cada $d^2 \in D_j(\overline{x}^i)$, de elementos do conjunto $E_j^{d^2}(\overline{x}^i)$. Obviamente, para duas seqüências da referência que coincidam para $i \geq j$, todos esses conjuntos coincidem; assim, podemos considerar seqüências de referência como virtualmente definidas apenas para $i \geq j$. Claramente, para L finito, o valor de $N_{d^2}^j(\overline{x}^i)$ é finito, quaisquer que sejam o intervalo de sinalização j, a seqüência de referência $(\overline{x}^i)$ e a distância $d^2 \in D_j(\overline{x}^i)$; diremos que o codificador em treliça é **não-catastrófico** se, para $L \to \infty$, tem-se que o valor de $N_{d^2}^j(\overline{x}^i)$ é finito, quaisquer que sejam a seqüência de referência $(\overline{x}^i)$, o intervalo de sinalização j e a distância $d^2 \in D_j(\overline{x}^i)$. Codificadores não-catastróficos são os de interesse na prática, pois são aqueles para os quais os possíveis eventos de erro (x^i), para uma seqüência de referência qualquer $(\overline{x}^i)$, têm duração limitada, para uma dada distância Euclideana quadrática $d^2 = \|(\overline{x}^i) - (x^i)\|^2$ à seqüência de referência, mesmo quando $L \to \infty$.

Determinemos, então, o limitante superior para a probabilidade de evento de erro, no j-ésimo intervalo de sinalização, para uma dada seqüência de referência transmitida $(\overline{x}^i)$, com $L \to \infty$, de um codificador em treliça não-catastrófico; denotaremos esta probabilidade por $P_e(j|(\overline{x}^i))$. Basta notar que uma condição necessária, mas não suficiente, para que ocorra um evento de erro, no j-ésimo intervalo de sinalização, dado que $\hat{s}j = \overline{s}j$, é que a seqüênia recebida (r^i) fique mais próxima de um possível evento de erro $(x^i) \in E_j(\overline{x}^i)$ do que da seqüência de referência $(\overline{x}^i)$; assim temos:

$$P_e(j|(\overline{x}^i)) \leq P\left[\bigcup_{(x^i) \in E_j(\overline{x}^i)} \{\|(x^i) - (r^i)\|^2 \leq \|(\overline{x}^i) - (r^i)\|^2\} \mid (x^i) \right] =$$

$$= P\left[\bigcup_{d^2 \in D_j(\overline{x}^i)} \bigcup_{(x^i) \in E_j^{d^2}(\overline{x}^i)} \|(x^i) - (r^i)\|^2 \leq \|(\overline{x}^i) - (r^i)\|^2 \mid (x^i) \right] \leq$$

$$\leq \sum_{d^2 \in D_j(\overline{x}^i)} \sum_{(x^i) \in E_j^{d^2}(\overline{x}^i)} P\left[\|(x^i) - (r^i)\|^2 \leq \|(\overline{x}^i) - (r^i)\|^2 \mid (\overline{x}^i) \right] \tag{4.2.10}$$

onde a segunda desigualdade é o denominado limitante de união:

$$P\left[\bigcup_i A_i\right] \leq \sum_i P[A_i]$$

Utilizando, agora, a expressão (4.2.9) em (4.2.10), temos:

$$P_e(j|(x^i)) \leq \sum_{d^2 \in D_j(\vec{x}^i)} \sum_{(x^i) \in E_j^{d^2}(\vec{x}^i)} Q\left(\left[\frac{1}{2N_0} \cdot \|(\vec{x}^i) - (x^i)\|^2\right]^{1/2}\right) =$$

$$= \sum_{d^2 \in D_j(\vec{x}^i)} \sum_{(x^i) \in E_j^{d^2}(\vec{x}^i)} Q\left(\left[\frac{1}{2N_0} \cdot d^2\right]^{1/2}\right) =$$

$$= \sum_{d^2 \in D_j(\vec{x}^i)} N_{d^2}^j(\vec{x}^i) \cdot Q\left(\left[\frac{1}{2N_0} \cdot d^2\right]^{1/2}\right) \qquad (4.2.11)$$

A expressão (4.2.11) fornece, assim, o limitante superior para $P_e(j|(\vec{x}^i))$. Agora, para determinarmos o correspondente limitante para $P_e(j)$, calculamos a média (em relação a $(\vec{x}^i)$):

$$P_e(j) = E_{(\vec{x}^i)}[P_e(j|(\vec{x}^i))] \qquad (4.2.12)$$

Para isso, definimos:

$$D_j \triangleq \bigcup_{(\vec{x}^i)} D_j(\vec{x}^i) \qquad (4.2.3)$$

e para cada $d^2 \in D_j$, estendemos a definição de $N_{d^2}^j(\vec{x}^i)$ fazendo $N_{d^2}^j(\vec{x}^i) \triangleq 0$ quando $d^2 \notin D_j(\vec{x}^i)$; assim podemos calcular a média

$$N_{d^2}^j \triangleq E_{(\vec{x}^i)}\,[N_{d^2}^j(\vec{x}^i)] \qquad (4.2.14)$$

para cada $d^2 \in D_j$. Então, utilizando (4.2.11) e (4.2.14), podemos limitar (4.2.12) superiormente por:

$$P_e(j) \leq E_{(\vec{x}^j)}\left[\sum_{d^2 \in D_j} N^j_{d^2}(\vec{x}^j).Q\left(\left[\frac{1}{2N_0}.d^2\right]^{1/2}\right)\right] =$$

$$= \sum_{d^2 \in D_j} E_{(\vec{x}^j)}\left[N^j_{d^2}(\vec{x}^j)\right].Q\left(\left[\frac{1}{2N_0}.d^2\right]^{1/2}\right) =$$

$$= \sum_{d^2 \in D_j} N^j_{d^2}.Q\left(\left[\frac{1}{2N_0}.d^2\right]^{1/2}\right) \qquad (4.2.15)$$

A expressão (4.2.15) fornece, finalmente, o limitante superior para $P_e(j)$ de um codificador em treliça não-catastrófico, para $L \rightarrow \infty$; para L finito, no entanto, a expressão (4.2.15) também fornece um limitante superior, pois já estaremos então considerando todos os possíveis eventos de erro para este valor de L, embora igualmente muitos outros, que no entanto não invalidam, obviamente, o limitante. Veremos, mais adiante, como obter, a partir de (4.2.15), um outro limitante que, embora, de modo geral, seja mais fraco, será mais adequado para compararmos o desempenho de codificadores específicos.

Vistos estes aspectos de desempenho do codificador em treliça, consideremos os aspectos de complexidade. Usualmente, a complexidade de decodificação predomina na implementação do sistema, sendo, então, adotada como medida de complexidade decorrente da adoção de um codificador em treliça específico. Um algoritmo de decodificação eficiente é o denominado **algoritmo de Viterbi**, estensamente discutido na literatura, e cuja descrição omitiremos; apenas alguns aspectos relevantes para a avaliação de complexidade devem ser ressaltados. Trata-se de um algoritmo recursivo que, a cada intervalo de sinalização j, calcula, para cada próximo estado possível $s^{j+1} = s$, $s \in S$, uma quantidade $d^2_s(j+1)$ da seguinte forma: dentre as $|A|^{K_2}$ transições de estado ("paralelas"), que partem de um mesmo estado $s^j = s'$ (em i=j) e chegam em $s^{j+1} = s$ (em i=j+1), determina-se aquela cujo ponto associado, digamos $x^j_{s'}$, está mais próximo do ponto recebido r^j, realizando, para isto, um certo número $N_B(K_2)$ de operações binárias. Feito isso, para cada possível estado anterior s' (existem $|A|^{K_1}$ tais estados), quando então já terá realizado $N_B(K_2)$ operações binárias, o decodificador realiza as $|A|^{K_1}$ somas $d^2_{s'}(j) + \|x^j_{s'} - r^j\|^2$, e determina a menor, fazendo então $|A|^{K_1} - 1$ comparações binárias; esta menor soma é a quantidade $d^2_s(j+1)$, cuja determinação requer, então, um total de $N_B(K_2).|A|^{K_1} + 2.|A|^{K_1} - 1$. Ao determinar $d^2_s(j+1)$ para todos os $|A|^\nu$ estados $s^{j+1} = s$, o algoritmo terá, então, realizado um total, por intervalo de sinalização, de $[N_B(K_2).|A|^{K_1} + 2.|A|^{K_1} - 1]|A|^\nu$ operações binárias; como podem existir feixes distintos de transições paralelas com um mesmo conjunto de pontos associados, este total de operações binárias, por intervalo de sinalização, constitui na realidade um limitante superior (pior caso) para a complexidade de um codificador em treliça utilizando o algoritmo de decodificação ML de Viterbi; no entanto, o valor exato pode, necessariamente, ser posto na forma $N_p(K_1,K_2) + (2.|A|^{K_1} - 1).|A|^\nu$, onde $N_p(K_1,K_2)$ é o número total de operações binárias para determinar-se, para cada feixe distinto de transições paralelas de uma seção da treliça, o ponto associado mais próximo do ponto recebido. Denota-se, por N_D, a **complexidade** de um

codificador, sendo expresso em operações binárias por intervalo de sinalização, definida como o total de operações binárias necessárias, por intervalo de sinalização, para se implementar o algoritmo de decodificação de Viterbi para este codificador, como discutido acima.

Feita esta breve consideração a respeito da complexidade de um codificador de treliça, passamos à conceituação de outro possível codificador para o canal vetorial gaussiano no tempo, denominado **codificador em bloco**, também conhecido por codificador Euclideano. Este codificador pode ser conceituado como um caso particular para o codificador em treliça, para o caso $\nu = 0$; utilizaremos este enfoque, especialmente porque facilitará a comparação entre codificadores específicos. Neste enfoque, então, um codificador em bloco pode ser visto como um codificador com apenas um estado, com $|A|^K$ transições paralelas, em cada seção de sua "treliça", partindo e retornando ao mesmo (único) estado, como ilustrado na Fig. 4.2.8, para $A = \{0,1\}$, $K = 2$, $\nu_1 = \nu_2 = 0$. Em particular, a decodificação ML de uma seqüência transmitida não requer a introdução de K-uplas pré-estabelecidas de entrada (pois m=0) e, portanto, não acarreta qualquer perda relativa no número de bits de informação enviados por período de sinalização T; além disso, para qualquer comprimento L de seqüência transmitida, a seqüência estimada ótima pode, obviamente, ser obtida selecionando, em cada intervalo de sinalização, a transição paralela cujo ponto associado está mais próximo do ponto recebido naquele intervalo. Quanto aos aspectos de desempenho, nota-se que existem, apenas, para qualquer seqüência de referência "eventos de erro" de comprimento igual a um intervalo de sinalização; assim, todo codificador em bloco é necessariamente "não-catastrófico" (visto como um codificador em treliça com apenas um estado), e a soma (4.2.15), que fornece um limitante superior para $P_e(j)$, tem necessariamente um número finito de termos. Finalmente, quanto aos aspectos de complexidade, a utilização do algoritmo de Viterbi reduz-se, como já era esperado, à sua fase inicial de determinar, para o único próximo estado possível, a transição de estado, no único feixe de transições paralelas que chegam neste estado, cujo ponto associado está mais próximo do ponto recebido naquele intervalo de sinalização, como já concluído acima; assim, com a notação utilizada para os códigos de treliça, o número de operações binárias realizadas pelo decodificador, por intervalo de sinalização, seria $N_B(K)$. Em resumo, codificadores em bloco são mais simples de serem analisados; no entanto, como veremos, para um mesmo desempenho, apresentam, geralmente, maior complexidade.

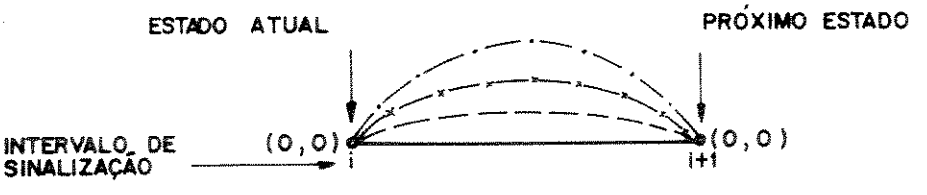

Fig. 4.2.8 - Estrutura de uma "treliça" com $A = \{0,1\}$, $K = 2$, $\nu_1 = \nu_2 = 0$. (——— $a_1 = 0$, $a_2 = 0$; - - - - - - - $a_1 = 1$, $a_2 = 0$; ____x____ $a_1 = 0$, $a_2 = 1$; ____.____ $a_1 = 1$, $a_2 = 1$)

Vistos os conceitos e as características elementares dos dois principais tipo de codificadores para o canal vetorial gaussiano discreto no tempo, passemos ao importante tópico da comparação entre estes codificadores. Como observação preliminar, lembramos que nosso objetivo nessa seção é, na realidade, a análise e comparação de codificadores para o canal AWGN limitado em banda a W, do qual um canal vetorial gaussiano discreto no tempo pode ser derivado,

cujo comportamento representa essencialmente o do canal contínuo original; no entanto, vários canais vetoriais gaussianos discretos no tempo podem ser derivados, um para cada possível valor do período de sinalização T, tendo, como vimos, dimensão limitada em $n \approx 2WT$; assim, devemos comparar codificadores não apenas para um mesmo canal vetorial gaussiano discreto no tempo, mas também codificadores para distintos canais vetoriais gaussianos discretos no tempo cuja dimensão n e taxa de sinalização T satisfaçam a condição $n/T \approx 2W$, onde W é a largura de banda do canal AWGN limitado em banda original do qual derivam; codificadores com a mesma razão n/T são ditos transmitirem o mesmo número de dimensões por segundo. Note-se que em um canal n-dimensional gaussiano discreto no tempo, derivado de um canal AWGN com densidade espectral de potência de ruído $N_0/2$ uniforme, cada coordenada do vetor de ruído tem energia média (variância), por intervalo de sinalização
$T \approx n/2W$, igual a $N_0/2$, de modo que a potência média N do vetor de ruído fica:

$$N = \frac{1}{T} \cdot E[\|z\|^2] =$$

$$= \frac{1}{T} \cdot E\left[\sum_{i=1}^{n} z_i^2\right] =$$

$$= \frac{1}{T} \cdot \sum_{i=1}^{n} E[z_i^2] =$$

$$= \frac{1}{T} \cdot \sum_{i=1}^{n} \frac{N_0}{2} =$$

$$= \frac{1}{T} .n. \frac{N_0}{2} \approx \qquad (4.2.16)$$

$$\approx N_0 . W \qquad (4.2.17)$$

utilizando-se a relação $n \approx 2WT$; da mesma forma, a **energia média** $P(\mathcal{C})$, por duas dimensões e por intervalo de sinalização T, de um codificador $\mathcal{C}$ para este canal n-dimensional, é definida por:

$$P(\mathcal{C}) \triangleq \frac{2}{n} . S(\mathcal{C}).T \qquad (4.2.18)$$

onde $S(\mathcal{C})$ é a potência média transmitida pelo codificador $\mathcal{C}$, dada por (4.2.3), de modo que:

$$S(\mathcal{C}) = \frac{1}{T} \cdot n \cdot \frac{P(\mathcal{C})}{2} \approx \tag{4.2.19}$$

$$\approx P(\mathcal{C}).W \tag{4.2.20}$$

utilizando-se, novamente, a relação $n = 2WT$; assim, a relação entre as potências médias de sinal e de ruído $S(\mathcal{C})/N$ fica, de (4.2.16) e (4.2.19) (ou, de (4.2.17) e (4.2.20)), dada por:

$$S(\mathcal{C})/N \approx P(\mathcal{C})/N_0 \tag{4.2.21}$$

ou seja, aproximadamente igual à relação entre as energias médias de sinal $P(\mathcal{C})$ e de ruído N_0, por duas dimensões e por intervalo de sinalização; assim, freqüentemente chamaremos $P(\mathcal{C})/N_0$ de relação sinal/ruído, para abreviar a terminologia.

A comparação do desempenho de codificadores específicos só é, geralmente, possível de ser analiticamente realizada em altas relações sinal/ruído; isto ocorre porque, normalmente, é neste caso que o limitante superior para P_e pode ser convenientemente simplificado, possibilitando sua manipulação analítica; além disso, neste caso, a expressão simplificada resultante para o limitante superior torna-se uma boa aproximação para P_e. Mais especificamente, ao compararmos codificadores para um dado canal vetorial gaussiano discreto no tempo, com energia média de ruído N_0 (por duas dimensões e por intervalo de sinalização), assumiremos que os pontos da constelação de um codificador $\mathcal{C}$, cujo limitante para P_e seria dado por (4.2.15), são amplificados por um fator escalar α real, de modo que para a nova probabilidade de evento de erro, denotada $P_{e(\alpha)}$, teríamos:

$$P_{e(\alpha)} \leq \sum_{d^2 \in D(\mathcal{C})} N_{d^2}.Q\left(\left[\frac{1}{2N_0} \cdot \alpha^2 \cdot d^2\right]^{1/2}\right) \tag{4.2.22}$$

com α suficientemente alto, o limitante superior em (4.2.22) é dominado pela parcela correspondente ao menor valor de $d^2 \in D(\mathcal{C})$, denotado $d^2_{MIN}(\mathcal{C})$, de modo que:

$$P_{e(\alpha)} \approx M_0(\mathcal{C}).Q\left(\left[\frac{1}{2N_0} \cdot \alpha^2 \cdot d^2_{MIN}(\mathcal{C})\right]^{1/2}\right) \tag{4.2.23}$$

onde $M_0(\mathcal{C}) = N_{d^2_{MIN}}(\mathcal{C})$ é o número de possíveis eventos de erro, que $\mathcal{C}$ possui em média, para uma seqüência de referência, iniciados em um dado intervalo da sinalização e cuja distância Euclideana quadrática à seqüência de referência seja $d^2_{MIN}(\mathcal{C})$; este número é, usualmente, denominado o **número de vizinhos mais próximos**, a uma seqüência de referência, para o codificador $\mathcal{C}$, ou ainda o **coeficiente de erro** do codificador $\mathcal{C}$. Geralmente, o limitante dado por (4.2.23) mostra-se uma aproximação aceitável para $P_{e(\alpha)}$, para α suficientemente alto; uma expressão mais adequada, no entanto, pode ser obtida, utilizando o limitante $Q(x) \leq 1/2 \cdot Exp\left[-1/2 \cdot x^2\right]$, com $x \geq 0$ (também estreito para x grande), produzindo, assim, nossa aproximação final:

$$P_{e(\alpha)} \approx \frac{M_0(\mathcal{C})}{2} \cdot \mathrm{Exp}\left[- \alpha^2 \cdot \frac{d^2_{MIN}(\mathcal{C})}{4N_0} \right] \qquad (4.2.24)$$

onde α é um fator de escala suficientemente alto; assim, $P_{e(\alpha)}$ pode ser feito arbitrariamente pequena às custas, obviamente, de uma potência média de transmissão, denotada $S_{(\alpha)}(\mathcal{C})$, maior do que a potência média original $S(\mathcal{C})$, dada por (4.2.3), pois:

$$S_{(\alpha)}(\mathcal{C}) = \alpha^2 . S(\mathcal{C}) \approx$$

$$\approx \alpha^2 . P(\mathcal{C}) . W \qquad (4.2.25)$$

onde $P(\mathcal{C})$ é a energia média original do codificador $\mathcal{C}$, por duas dimensões e por intervalo de sinalização, dada por (4.2.18); tem-se, portanto, uma alta relação sinal/ruído:

$$S_{(\alpha)}(\mathcal{C})/N = \alpha^2 . S(\mathcal{C})/N$$

$$\approx \alpha^2 . P(\mathcal{C})/N_0$$

onde a potência média N de ruído é dada por (4.2.17). Utilizaremos, então, as expressões dadas por (4.2.24) e (4.2.25), ao compararmos o desempenho de codificadores para um dado canal AWGN limitado em banda a W (ou seja, que transmitem um mesmo número $n/T \approx 2W$ de dimensões por segundo), em altas relações sinal/ruído.

A comparação entre codificadores é direta quando estes são especificados para o mesmo canal vetorial gaussiano discreto no tempo derivado do canal AWGN original (ou seja, possuem a mesma dimensão n e, portanto, o mesmo período de sinalização $T \approx n/2W$) e transmitem o mesmo número de bits de informação por intervalo de sinalização T: estipulando uma probabilidade de evento de erro P_e (suficientemente pequena) como objetivo, aquele que necessitar da menor potência média será o melhor dentre eles para o canal vetorial gaussiano discreto no tempo dado, e corresponderá, igualmente, ao melhor dentre eles para o canal AWGN original. Usualmente, as potência médias, dos codificadores a serem comparados são previamente normalizadas; uma normalização conveniente consiste em dividir cada potência média pela potência média $S_{(\alpha_R)}(\mathcal{C}_R)$ de um codificador de referência $\mathcal{C}_R$ arbitrariamente selecionado, mas que seja especificado para o mesmo canal n-dimensional gaussiano discreto no tempo, transmita o mesmo número de bits por intervalo de sinalização $T \approx n/2W$ e possua a mesma probabilidade de evento de erro P_e com aquela potência média de normalização $S_{(\alpha_R)}(\mathcal{C}_R)$, tal qual os codificadores que estão sendo comparados. Denomina-se **ganho efetivo** de um codificador, em relação a um codificador de referência, o inverso de sua potência média normalizada; é usualmente expresso em decibéis (dB), e denotado $\gamma_{eff}(\mathcal{C})$ para um dado codificador $\mathcal{C}$. Mais especificamente, seja $\mathcal{C}$ um codificador qualquer e $\mathcal{C}_R$ uma possível referência para $\mathcal{C}$ como definida acima; para uma dada P_e suficientemente pequena, teremos, usando (4.2.24) e (4.2.25), para o codificador $\mathcal{C}$:

$$P_e \approx \frac{M_0(\mathcal{C})}{2} \cdot \mathrm{Exp}\left[-\alpha^2 \cdot \frac{d^2_{MIN}(\mathcal{C})}{4.N_0}\right] \tag{4.2.27}$$

$$S_{(\alpha)}(\mathcal{C}) \approx \alpha^2 . P(\mathcal{C}) . W \tag{4.2.28}$$

enquanto que, para o codificador $\mathcal{C}_R$:

$$P_e \approx \frac{M_0(\mathcal{C}_R)}{2} \cdot \mathrm{Exp}\left[-\alpha_R^2 \cdot \frac{d^2_{MIN}(\mathcal{C}_R)}{4.N_0}\right] \tag{4.2.29}$$

$$S_{(\alpha_R)}(\mathcal{C}_R) \approx \alpha_R^2 . P(\mathcal{C}_R) . W \tag{4.2.30}$$

de modo que o ganho efetivo $\gamma_{eff}(\mathcal{C})$ do codificador $\mathcal{C}$ (isto é, em relação ao codificador de referência $\mathcal{C}_R$), é dado por:

$$\gamma_{eff}(\mathcal{C}) \triangleq \frac{S_{(\alpha_R)}(\mathcal{C}_R)}{S_{(\alpha)}(\mathcal{C})} \approx$$

$$\approx \frac{\alpha_R^2 . P(\mathcal{C}_R)}{\alpha^2 . P(\mathcal{C})} \tag{4.2.31}$$

usando (4.2.28) e (4.2.30); de (4.2.27) e (4.2.29) tiramos:

$$\alpha^2 \approx \frac{4.N_0}{d^2_{MIN}(\mathcal{C})} \cdot \ln\left(\frac{M_0(\mathcal{C})}{2.P_e}\right) \tag{4.2.32}$$

$$\alpha_R^2 \approx \frac{4.N_0}{d^2_{MIN}(\mathcal{C}_R)} \cdot \ln\left(\frac{M_0(\mathcal{C}_R)}{2.P_e}\right) \tag{4.2.33}$$

Substituindo (4.2.32) e (4.2.33) em (4.2.31), resulta:

$$\gamma_{eff}(\mathcal{C}) \approx \frac{d^2_{MIN}(\mathcal{C})}{d^2_{MIN}(\mathcal{C}_R)} \cdot \frac{P(\mathcal{C}_R)}{P(\mathcal{C})} \cdot \frac{\ln\left(\frac{M_0(\mathcal{C}_R)}{2.P_e}\right)}{\ln\left(\frac{M_0(\mathcal{C})}{2.P_e}\right)}$$

Expressando-se $\gamma_{eff}(\mathcal{C})$ em dB, teremos:

$$\gamma_{eff}(\mathcal{C})_{dB} \approx 10.\log_{10} \frac{\left(\frac{d^2_{MIN}(\mathcal{C})}{P(\mathcal{C})}\right)}{\left(\frac{d^2_{MIN}(\mathcal{C}_R)}{P(\mathcal{C}_R)}\right)} - 10.\log_{10} \frac{\ln\left(\frac{M_0(\mathcal{C})}{2.P_e}\right)}{\ln\left(\frac{M_0(\mathcal{C}_R)}{2.P_e}\right)} =$$

$$\triangleq \gamma(\mathcal{C})_{dB} - \eta_0(\mathcal{C})_{dB} \qquad (4.2.34)$$

onde $\gamma(\mathcal{C})$ é denominado o ganho nominal do codificador $\mathcal{C}$, e $\eta_0(\mathcal{C})$ é denominada a perda pelo coeficiente de erro de $\mathcal{C}$, ou ainda a perda pelo número de vizinhos mais próximos em $\mathcal{C}$, em dada P_e (suficientemente pequena). Portanto, entre codificadores de mesma dimensão n e que transmitem o mesmo número de bits por intervalo de sinalização $T \approx n/2W$, aquele que possui maior ganho efetivo, com relação a uma mesma referência compatível, é o que apresenta melhor desempenho. Note-se que o ganho efetivo, e portanto a comparação entre codificadores específicos independe do particular nível de ruído $N_0/2$ por coordenada do canal n-dimensional gaussiano discreto no tempo (bem como, portanto, da densidade espectral uniforme $N_0/2$ do canal AWGN limitado em banda, do qual aquele deriva), para altas relações sinal/ruído.

Seria conveniente adotarmos um codificador de referência definitivo $\mathcal{C}_R$, para cada dimensão n e cada número b de bits de informação enviado por intervalo de sinalização $T \approx n/2W$ (para um codificador com K-uplas de entrada independentes e uniformemente distribuídas sobre A^K, onde A é o alfabeto dos símbolos de informação de entrada, temos $b = K.\log_2|A|$, como vimos); uma referência $\mathcal{C}_R$ (conveniente, como veremos) para um tal codificador é o codificador em bloco $\mathcal{C}_{\blacksquare}(b,n)$ cuja constelação $C_{\blacksquare}$ é formada por todos os 2^b pontos do $\mathbb{R}^n$ tais que cada uma das coordenadas pode assumir qualquer um dos $2^{b/n}$ valores $\pm 1/2, \pm 3/2, \pm 5/2, \ldots, \pm \frac{2^{b/n}-1}{2}$, o alfabeto é $A_{\blacksquare} = \{0,1\}$ e o número de entradas é $K_{\blacksquare} = b$, com uma associação biunívoca arbitrária entre os 2^b pontos da constelação e as 2^b b-uplas binárias de entrada; para esta referência, tem-se:

$$P(\mathcal{C}_{\blacksquare}(b,n)) = \frac{2^{2b/n} - 1}{6} \approx \frac{2^{2b/n}}{6} \qquad (4.2.35)$$

$$d_{MIN}(\mathcal{C}_{\blacksquare}(b,n)) = 1 \qquad (4.2.36)$$

$$M_0(\mathcal{C}_{\blacksquare}(b,n)) = 2.n.\left[1 - \frac{1}{2^{b/n}}\right] \approx 2.n \qquad (4.2.37)$$

de modo que, de (4.2.34), o ganho nominal de um codificador $\mathcal{C}$, compatível com este codificador de referência (ou seja, com dimensão n e enviando b bits de informação por intervalo $T = n/2W$), fica:

$$\gamma(\mathcal{C})_{dB} = 10.\log_{10}\left(\frac{2^{2b/n}}{6} \cdot \frac{d^2_{MIN}(\mathcal{C})}{P(\mathcal{C})}\right) \qquad (4.2.38)$$

enquanto que a perda pelo coeficiente de erro $M_0(\mathcal{C})$ fica:

$$\eta_0(\mathcal{C})_{dB} = 10.\log_{10}\frac{\ln\left(\frac{M_0(\mathcal{C})}{2.P_e}\right)}{\ln\left(\frac{n}{P_e}\right)} \qquad (4.2.39)$$

Assumindo que P_e é escolhida suficientemente pequena de modo que $n << 1/P_e$ e $M_0(\mathcal{C}) << 1/P_e$, é possível simplificar a expressão em (4.2.39) para $\eta_0(\mathcal{C})_{dB}$, utilizando a aproximação $\ln(x+1) \approx x$ (para $|x| << 1$), como segue:

$$\eta_0(\mathcal{C})_{dB} = \frac{10}{\ln 10}.\ln\frac{\ln\left(\frac{M_0(\mathcal{C})}{2}\right) + \ln\left(\frac{1}{P_e}\right)}{\ln(n) + \ln\left(\frac{1}{P_e}\right)} =$$

$$= \frac{10}{\ln 10}.\ln\frac{\frac{\ln\left(\frac{M_0(\mathcal{C})}{2}\right)}{\ln\left(\frac{1}{P_e}\right)} + 1}{\frac{\ln(n)}{\ln\left(\frac{1}{P_e}\right)} + 1} =$$

$$= \frac{10}{\ln 10}.\left[\ln\left(\frac{\ln\left(\frac{M_0(\mathcal{C})}{2}\right)}{\ln\left(\frac{1}{P_e}\right)} + 1\right) - \ln\left(\frac{\ln(n)}{\ln\left(\frac{1}{P_e}\right)} + 1\right)\right] =$$

$$\approx \frac{10}{\ln 10} \cdot \left[\frac{\ln\left(\frac{M_0(\mathcal{C})}{2}\right)}{\ln\left(\frac{1}{P_e}\right)} - \frac{\ln(n)}{\ln\left(\frac{1}{P_e}\right)} \right] =$$

$$= \frac{10.\log_{10}2}{\ln\left(\frac{1}{P_e}\right)} \cdot \log_2\left(\frac{M_0(\mathcal{C})}{2.n}\right) =$$

$$= \frac{3.01}{\ln\left(\frac{1}{P_e}\right)} \cdot \log_2 \frac{\tilde{M}_0(\mathcal{C})}{4} \qquad (4.2.40)$$

onde $\tilde{M}_0(\mathcal{C}) \triangleq 2/n \, . \, M_0(\mathcal{C})$, na última expressão, é o número de vizinhos mais próximos, por duas dimensões, em $\mathcal{C}$, denominado o **coeficiente de erro normalizado**, por duas dimensões do codificador $\mathcal{C}$. Por exemplo, para $P_e \approx 10^{-6}$ temos $\eta_0(\mathcal{C})_{dB} \approx 0.22.\log_2\tilde{M}_0(\mathcal{C})/4$, ou seja, uma perda de 0.22 dB cada vez que $\tilde{M}_0(\mathcal{C})$ dobra, concordando com o que é sugerido na literatura. Note-se que, quando $P_e \to 0$, tem-se que $\eta_0(\mathcal{C})_{dB} \to 0$ (por (4.2.34)), de modo que $\gamma_{eff}(\mathcal{C}) \to \gamma(\mathcal{C})$; por esta razão, $\gamma(\mathcal{C})$ é usualmente também denominado de **ganho assintótico**. Em qualquer caso, no entanto, com $n << 1/P_e$ e $\tilde{M}_0(\mathcal{C}) << 1/P_e$, temos:

$$\gamma_{eff}(\mathcal{C})_{dB} \approx 10.\log_{10}\left(\frac{2^{2b/n}}{6} \cdot \frac{d^2_{MIN}(\mathcal{C})}{P(\mathcal{C})}\right) - \frac{3.01}{\ln\left(\frac{1}{P_e}\right)} \cdot \log_2 \frac{\tilde{M}_0(\mathcal{C})}{4} \qquad (4.2.41)$$

A conveniência, mencionada acima, de adotarmos o particular codificador $\mathcal{C}_\blacksquare(b,n)$ como referência, para codificadores n-dimensionais que enviam b bits de informação por intervalo de sinalização $T \approx n/2W$, reside no fato de que podemos utilizar o ganho efetivo, dado por (4.2.41), para comparar o desempenho de codificadores, mesmo entre codificadores com diferentes valores de n e b. Para fundamentar esta afirmação, basta calcularmos o maior ganho efetivo possível para um codificador, em relação a sua respectiva referência $\mathcal{C}_\blacksquare(b,n)$. A capacidade do canal n-dimensional gaussiano discreto no tempo, da Fig. 4.2.2, com variância de ruído por dimensão igual a $N_0/2$, cuja entrada está restrita a uma energia média máxima P, por duas dimensões e por intervalo de sinalização $T \approx n/2W$, é dada por:

$$C = \frac{n}{2} \cdot \log_2\left(\frac{P}{N_0} + 1\right)$$

onde C está expresso em unidades de bits de informação por intervalo de sinalização T. Assim, para P_e suficientemente pequena, todo codificador $\mathcal{C}$ para este canal, que envia b bits de informação por intervalo de sinalização T, com probabilidade de evento de erro P_e, deve possuir

energia média $\alpha^2.P(\mathcal{C})$ (já incluído o fator de amplificação α), por duas dimensões e por intervalo de sinalização, maior ou igual a $(\alpha^2.P(\mathcal{C}))_{MIN}$, dado por:

$$b = \frac{n}{2} \cdot \log_2 \left(\frac{(\alpha^2.P(\mathcal{C}))_{MIN}}{N_0} + 1 \right)$$

ou seja, para valores ao menos moderados de 2b/n, teremos:

$$(\alpha^2.P(\mathcal{C}))_{MIN} \approx N_0.2^{2b/n} \tag{4.2.42}$$

Para o codificador $\mathcal{C}_{\blacksquare}(b,n)$ de referência, temos, de (4.2.33) e (4.2.35-37), que:

$$\alpha_{\blacksquare}^2.P(\mathcal{C}_{\blacksquare}(b,n)) \approx \frac{2}{3} \cdot \ln\left(\frac{n}{P_e}\right).N_0.2^{2b/n} \tag{4.2.43}$$

Substituindo-se (4.2.42) e (4.2.43) em (4.2.31), tiramos:

$$\gamma_{eff}(\mathcal{C})_{max} \approx \frac{\frac{2}{3} \cdot \ln\left(\frac{n}{P_e}\right).N_0.2^{2b/n}}{N_0.2^{2b/n}} =$$

$$= \frac{2}{3} \cdot \ln\left(\frac{n}{P_e}\right)$$

Assim, para $n << 1/P_e$, o ganho efetivo máximo $\gamma_{eff}(\mathcal{C})_{max}$, expresso em dB, fica:

$$\gamma_{eff}(\mathcal{C})_{maxdB} \approx 10.\log_{10}\left(\frac{2}{3} \cdot \ln\left(\frac{1}{P_e}\right)\right) \tag{4.2.44}$$

Por exemplo, em $P_e = 10^{-6}$, teríamos $\gamma_{eff}(\mathcal{C})_{max} \approx 9.64$ dB. Note-se que a expressão (4.2.44) para o ganho efetivo máximo de um codificador n-dimensional, que envie b bits de informação por intervalo de sinalização $T \approx n/2W$, em relação ao respectivo codificador $\mathcal{C}_{\blacksquare}(b,n)$ de referência, num dado nível de probabilidade de evento de erro P_e, essencialmente, não depende de n nem de b, mas apenas de P_e. Como um codificador terá tão melhor desempenho, para um dado valor de P_e, quanto menor for a diferença entre o máximo ganho que poderia alcançar, dado por (4.2.44), e o ganho que de fato alcança, dado por (4.2.41), e como aquele máximo em (4.2.44) não depende essencialmente de n nem de b, concluímos que codificadores podem ser comparados em desempenho baseando-se exclusivamente na comparação de ganhos efetivos, para uma mesma P_e dada, em relação aos respectivos codificadores $\mathcal{C}_{\blacksquare}(b,n)$ de referência: simplesmente o melhor dentre os codificadores comparados, em termos do desempenho, será aquele que apresentar o maior ganho efetivo, como afirmado inicialmente.

Como último comentário a respeito da expressão (4.2.41), note-se que dentre codificadores com os mesmos valores dos parâmetros $d_{MIN}(\mathcal{C})$, $P(\mathcal{C})$ e $\tilde{M}_0(\mathcal{C})$, aquele para o

qual a relação 2b/n for a maior terá, obviamente, o maior ganho efetivo $\gamma_{eff}(\mathcal{C})$, em qualquer P_e suficientemente pequena; esta relação 2b/n é denominada a **taxa normalizada** (para duas dimensões) do codificador $\mathcal{C}$, denotada $\tilde{b}(\mathcal{C})$ (ou simplesmente $\tilde{b}$), sendo expressa em bits de informação enviados por duas dimensões e por intervalo de sinalização $T \approx n/2W$; é também conhecida como **eficiência espectral**, por ser igual a $(b/T)/W$, podendo então ser também expressa em bit's de informação por unidade de tempo e por unidade de freqüência.

Finalmente, passamos, agora, a analisar a comparação da complexidade de codificadores específicos. Note-se que, dentre codificadores com o mesmo desempenho, aquele para o qual o número de operações binárias por unidade de tempo, realizado pelo respectivo decodificador, for menor será, obviamente, o mais adequado. Seria conveniente, portanto, definir-se a **complexidade normalizada** $\tilde{N}_D$, de um codificador como sendo a relação N_D/T, onde N_D é a complexidade de decodificação do codificador; no entanto, como estamos comparando codificadores para utilização em um mesmo canal AWGN limitado em banda a W, podemos definir a complexidade normalizada como sendo a relação $2N_D/n \approx (N_D/T)/W$, semelhantemente à definição de taxa normalizada; esta é a definição que adotamos, como na literatura. Obviamente, este conceito pode ser usado para compararmos a complexidade de codificadores com desempenhos diferentes, sendo que, neste caso, a maior adequação de um codificador em relação aos demais dependerá dos aspectos econômicos da aplicação específica para o sistema de codificação em estudo.

Concluímos esta seção reiterando que a comparação do desempenho e da complexidade de codificadores específicos não depende dos particulares valores de T ou W, mas apenas de seu produto $TW \approx n/2$. De fato, podemos, notar que os parâmetros relevantes para a avaliação de um codificador $\mathcal{C}$ (isto é, $d^2_{MAX}(\mathcal{C})$, $P(\mathcal{C})$, $\tilde{M}_0(\mathcal{C})$, $\tilde{b}(\mathcal{C})$ e $\tilde{N}_D(\mathcal{C})$) dependem da estrutura de sua treliça, da configuração das distâncias entre os pontos associados às transições de estado na treliça e da dimensão n desses pontos, sendo irrelevantes o período de sinalização T e, por conseguinte, a largura de banda $W = n/2T$ correspondente. Isto, entre outras coisas, permite o uso de um mesmo codificador em diferentes canais AWGN, com mesmo desempenho e mesma complexidade, visto que a densidade espectral de potência uniforme é também irrelevante para a avaliação de um codificador, em altas relações sinal/ruído.

4.3 Construção de codificadores baseados em reticulados

Nesta seção descrevemos brevemente as técnicas correntes, propostas na literatura, de construção de codificadores, para o canal AWGN limitado em banda, que utilizam reticulados e suas partições; o propósito desta descrição é evidenciar as características relevantes dos reticulados para esta aplicação, de modo a orientar a seleção de reticulados adequados para utilização nestes esquemas de construção de codificadores baseados em reticulados.

Para simplificar a exposição, sem comprometer essencialmente sua generalidade, nos restringimos aos codificadores cujo alfabeto de entrada seja binário, digamos $A=\{0,1\}$; assim, um codificador n-dimensional com taxa normalizada $\bar{b} = 2b/n$ bits, por duas dimensões e por intervalo de sinalização, terá b entradas binárias, assumindo-se que as b-uplas binárias de entrada são independentes e uniformemente distribuídas sobre A^b, como usualmente fazemos.

O esquema de construção mais simples é, possivelmente, aquele que meramente mapeia cada uma das 2^b b-uplas binárias distintas possíveis de entrada, por intervalo de sinalização, em um ponto distinto de uma constelação constituída por 2^b pontos de um reticulado n-dimensional Λ (ou de uma classe lateral $c + \Lambda$ deste, onde $c \in \mathbb{R}^n$ é selecionado de modo a tornar a constelação simétrica em relação à origem, geralmente um buraco profundo de Λ como definido na seção 2.4) pertencentes a uma região limitada $\mathfrak{R}$ do $\mathbb{R}^n$ (igualmente simétrica em relação à origem); o codificador em bloco n-dimensional $\mathcal{C}$ obtido é geralmente denominado um **codificador reticulado** (Conway e Sloane [8], Forney e Wey [15]), e utilizaremos a notação $\mathcal{C}(\Lambda,\mathfrak{R})$ quando desejarmos evidenciar o reticulado Λ e a região $\mathfrak{R}$ utilizados na construção; a figura 4.3.1 ilustra a forma geral da construção. Com a suposição usual de b-uplas binárias de entrada independentes e uniformemente distribuídas sobre A^b, onde $A=\{0,1\}$, o desempenho do codificador reticulado resultante obviamente não depende do particular mapeamento biunívoco estabelecido entre estas b-uplas binárias de entrada e os pontos da constelação; a seleção deste mapeamento biunívoco deve, então, ser orientada pela simplicidade da sua implementação na transmissão e da implementação de seu mapeamento inverso na recepção, após a fase de decodificação; em geral, a complexidade dessas implementações depende tanto de Λ como de $\mathfrak{R}$, sendo algumas soluções propostas na literatura (Conway e Sloane [10], Forney e Wei [15], Forney [16]). Como na seção anterior, não consideramos a complexidade dessas implementações na avaliação da complexidade desses codificadores, que fica então resumida à complexidade de decodificação (estimação da sequência de pontos n-dimensionais transmitidos a partir da sequência de pontos n-dimensionais recebida).

Fig 4.3.1 - Codificador $\mathcal{C}(\Lambda,\mathfrak{R})$

É particularmente simples relacionar, de forma aproximada, os parâmetros do codificador reticulado $\mathcal{C}(\Lambda,\mathfrak{R})$ a parâmetros do reticulado Λ e da região $\mathfrak{R}$; por exemplo, $d^2_{MIN}(\mathcal{C}) = d^2_{MIN}(\Lambda)$. Forney e Wei [15] demonstraram, utilizando uma aproximação conhecida por **aproximação contínua** (também conhecida como **aproximação integral** - ver Forney e outros [18]), que, quando o volume $V(\mathfrak{R})$ da região $\mathfrak{R}$ é consideravelmente maior que o volume fundamental $V(\Lambda)$ do reticulado Λ (e, portanto, quando o número $2^b \approx V(\mathfrak{R})/V(\Lambda)$ de pontos da constelação é consideravelmente grande), o ganho nominal $\gamma(\mathcal{C})_{dB}$ do codificador reticulado $\mathcal{C}(\Lambda,\mathfrak{R})$ pode ser expresso por

$$\gamma(\mathcal{C})_{dB} = \gamma(\Lambda)_{dB} + \gamma_s(\mathfrak{R})_{dB} \tag{4.3.1}$$

onde $\gamma(\Lambda)$ é o parâmetro de Hermite de Λ, definido na seção 2.4, que passamos a denominar por **ganho nominal de codificação** de Λ (também conhecido como **ganho de codificação fundamental** de Λ, após Forney [13]), e $\gamma_s(\mathfrak{R})$ é o denominado **ganho de forma** de $\mathfrak{R}$ (que é maior quanto mais "esférica" em torno da origem for a região $\mathfrak{R}$), definido por:

$$\gamma_s(\mathfrak{R}) \triangleq \frac{V(\mathfrak{R})^{2/n}}{[6.P(\mathfrak{R})]} = \frac{1}{[12.G(\mathfrak{R})]}$$

onde $P(\mathfrak{R})$ é a **norma média** (energia média), por duas dimensões, de uma distribuição contínua, uniforme sobre os pontos do $\mathbb{R}^n$ dentro de $\mathfrak{R}$ e nula fora de $\mathfrak{R}$, e G(R) é o **segundo momento normalizado** de $\mathfrak{R}$, definido por $G(\mathfrak{R}) \triangleq P(\mathfrak{R})/[2.V(\mathfrak{R})^{2/n}]$. Assim, para $V(\mathfrak{R}) >> V(\Lambda)$, o ganho nominal $\gamma(\mathcal{C})$ de $\mathcal{C}(\Lambda,\mathfrak{R})$ decompõe-se em dois fatores, um essencialmente dependente apenas de Λ e outro essencialmente dependente apenas de $\mathfrak{R}$, dentro da precisão fornecida pela aproximação contínua.

Definimos, por analogia, a **perda pelo número de vizinhos mais próximos** em Λ, ou ainda a **perda pelo coeficiente de erro** de Λ, em uma dada P_e (suficientemente pequena tal que $M_0(\Lambda) << 1/P_e$ e $n << 1/P_e$), por:

$$\eta_0(\Lambda)_{dB} \triangleq \frac{3.01}{\ln(1/P_e)} \cdot \log_2\left(\frac{M_0(\Lambda)}{2n}\right) \tag{4.3.2}$$

onde $M_0(\Lambda)$ é o número de vizinhos mais próximos em Λ definido na seção 2.4, que passamos a denominar igualmente por **coeficiente de erro** de Λ, e $\tilde{M}_0(\Lambda) \triangleq (2/n).M_0(\Lambda)$ é o número de vizinhos mais próximos, por duas dimensões, em Λ, que denominamos por **coeficiente normalizado de erro** de Λ. Da mesma forma, definimos o **ganho efetivo de codificação** de Λ, analogamente, por:

$$\gamma_{eff}(\Lambda)_{dB} \triangleq \gamma(\Lambda)_{dB} - \eta_0(\Lambda)_{dB} \tag{4.3.3}$$

Nestes termos, observando que, para grandes constelações (ou seja, grandes valores de 2^b), temos $M_0(\mathcal{C}) \approx M_0(\Lambda)$, teríamos de (4.3.2), neste caso:

$$\eta_0(\mathcal{C}) \approx \eta_0(\Lambda) \tag{4.3.4}$$

numa dada P_e. Finalmente, ainda neste caso, teríamos de (4.3.1), (4.3.3) e (4.3.4):

$$\gamma_{eff}(\mathcal{C})_{dB} \approx \gamma_{eff}(\Lambda)_{dB} + \gamma_s(\mathfrak{R})_{dB} \qquad (4.3.5)$$

para altas relações sinal/ruído, em virtude de (4.2.34). Assim, para $V(\mathfrak{R}) >> V(\Lambda)$, o ganho efetivo $\gamma_{eff}(\mathcal{C})$ de $\mathcal{C}(\Lambda,\mathfrak{R})$ decompõe-se igualmente em dois fatores, um essencialmente dependente apenas de Λ e outro essencialmente dependente apenas de $\mathfrak{R}$, em altas relações sinal/ruído.

Fica, então, evidente que para obter codificadores reticulados $\mathcal{C}(\Lambda,\mathfrak{R})$ de alto desempenho, em altas relações sinal/ruído, devemos utilizar reticulados Λ com máximo ganho efetivo de codificação $\gamma_{eff}(\Lambda)$ e regiões com máximo ganho de forma $\gamma_s(\mathfrak{R})$ (a região n-dimensional de maior ganho de forma é a hiperesfera centrada na origem, com valor assintótico de 1.52 dB, quando n tende a infinito).

Voltemo-nos, agora, aos aspectos de complexidade de um codificador reticulado $\mathcal{C}(\Lambda,\mathfrak{R})$. Um algoritmo de decodificação por máxima verossimilhança (ML) para $\mathcal{C}(\Lambda,\mathfrak{R})$ define em torno de cada ponto λ de sua constelação uma região de decisão $R_{\mathcal{C}}(\lambda)$ formada por todos os pontos do $\mathbb{R}^n$ mais próximos de λ do que de qualquer outro ponto desta constelação. Vemos, então, que para a maioria dos pontos λ da constelação (exceto para um número relativamente pequeno de pontos próximos à fronteira da região $\mathfrak{R}$) a região de decisão $R_{\mathcal{C}}(\lambda)$ coincide com a região de Voronoi $V_\Lambda(\lambda)$ de Λ em torno de $\lambda \in \Lambda$. Portanto, podemos estimar a complexidade de decodificação de $\mathcal{C}(\Lambda,\mathfrak{R})$ como sendo a complexidade de um algoritmo que, dado um ponto $r \in \mathbb{R}^n$, determina o ponto $\lambda \in \Lambda$ tal que $r \in V_\Lambda(\lambda)$, denominado um algoritmo de **decodificação por ML do reticulado** Λ; denotaremos sua complexidade por $N_D(\Lambda)$, e sua complexidade normalizada $(2/n).N_D(\Lambda)$ por $\tilde{N}_D(\Lambda)$. Assim, para obtermos codificadores reticulados $\mathcal{C}(\Lambda,\mathfrak{R})$ de baixa complexidade normalizada $\tilde{N}_D(\mathcal{C})$ devemos utilizar reticulados Λ de baixa complexidade normalizada $\tilde{N}_D(\Lambda)$.

Observa-se, infelizmente, que em geral alto ganho efetivo de codificação e baixa complexidade normalizada são objetivos conflitantes na seleção do reticulado utilizado, o que já era esperado pelo trabalho de Shannon ([29]). A seleção de reticulados para utilização em codificadores para o canal AWGN limitado em banda deve ser orientada para alcançar o melhor compromisso possível entre esses objetivos conflitantes. Uma possível abordagem é a utilização de algoritmos de decodificação sub-ótimos (assintoticamente ML, em altas relações sinal/ruído) que, embora reduzam o ganho efetivo de codificação fornecido pelo reticulado, podem apresentar redução substancial de complexidade em relação aos algoritmos de decodificação ótimos (ML); trataremos dessa abordagem no capítulo 6.

Para finalizar esses breves comentários sobre as características relevantes dos reticulados para a construção de codificadores reticulados $\mathcal{C}(\Lambda,\mathfrak{R})$ mencionamos que igual conflito existe na seleção da região $\mathfrak{R}$ utilizada, ou seja, quanto maior o ganho de forma $\gamma_s(\mathfrak{R})$ da região $\mathfrak{R}$, maior será a complexidade de implementação do mapeamento biunívoco entre as 2^b b-uplas binárias de entrada e os 2^b pontos da constelação, tanto na transmissão (mapeamento direto) quanto na recepção (mapeamento inverso) após a decodificação. Além disso, na prática, outras restrições podem limitar ainda mais a seleção de regiões $\mathfrak{R}$ que maximizam $\gamma_s(\mathfrak{R})$, para uma

dada dimensão n (por exemplo: relação de energias pico/média e taxa de expansão das constelações bidimensionais constituintes - ver Wei [35], Forney e Wei [15]).

Passemos, agora, a um esquema de construção de codificadores baseados em reticulados que, de certa forma, generaliza o esquema de construção acima. Deve-se a Calderbank e Sloane ([5]) a observação de que muitos dos bons esquemas de construção de codificadores (em bloco e em treliça) propostos na literatura podem ser descritos em termos de partições de reticulados; posteriormente Forney ([13]) formalizou essas idéias, criando o conceito de **código de classes laterais** (do tipo reticulado), denotado $\mathbf{C}(\Gamma/\Lambda,C)$, que utiliza uma partição de reticulados n-dimensionais Γ/Λ, de ordem $|\Gamma/\Lambda| = 2^{K+r}$, e um codificador binário C, de taxa K/(K+r), como ilustrado na figura 4.3.2. Para uma sequência infinita de K-uplas binárias $[a_1^i,\ldots,a_K^i]^T$ de entrada, o codificador binário C gera uma sequência infinita de (K+r)-uplas binárias $[c_1^i,\ldots,c_{K+r}^i]^T$ de saída; o seletor de classe lateral implementa um mapeamento biunívoco fixo entre o conjunto de todas as 2^{K+r} possíveis (K+r)-uplas binárias e a partição Γ/Λ constituída de 2^{K+r} classes laterais; assim, para cada sequência infinita de K-uplas binárias $[a_1^i,\ldots,a_K^i]^T$ na entrada do codificador binário C, é gerada uma sequência infinita de classes laterais $c^i + \Lambda$ na saída do seletor de classe lateral em Γ/Λ; considerando cada sequência infinita de classes laterais $(c^0 + \Lambda, c^1 + \Lambda,\ldots)$ como o conjunto de todas as sequências de pontos $(c^0 + \lambda^0, c^1 + \lambda^1,\ldots)$ com $\lambda^i \in \Lambda$ arbitrários, podemos definir o código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ como a união de todas as possíveis sequências infinitas de classes laterais na saída do seletor de classe lateral, variando-se a sequência infinita de K-uplas binárias na entrada do codificador binário

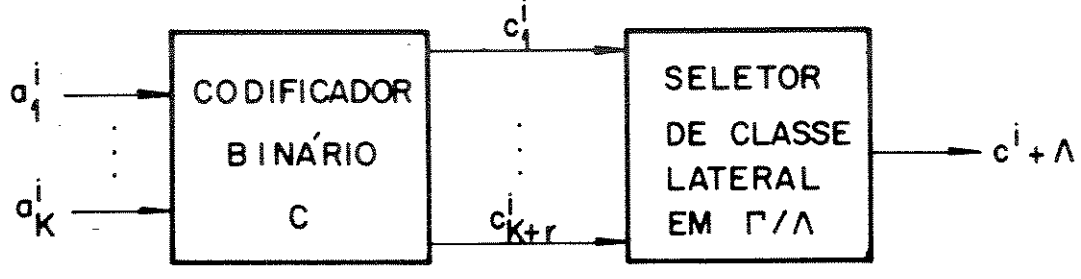

Fig. 4.3.2 - Código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$

Se C for um codificador binário em treliça (por exemplo, um codificador convolucional), tendo um registrador de deslocamento com ν_k elementos de memória binária associado à sua k-ésima entrada, $1 \le k \le K$, podemos então representar a operação de C, em resposta a uma sequência infinita de K-uplas binárias de entrada, por um caminho em uma treliça como na figura 4.2.5, ou seja, com 2^ν estados distintos, onde $\nu = \sum_{k=1}^{K} \nu_k$, e tendo em cada uma de suas seções (idênticas) 2^K transições de estado partindo de um mesmo estado presente (e chegando em um mesmo estado futuro), agrupadas em 2^{K_2} feixes com 2^{K_1} transições paralelas (com mesmo estado

presente e futuro), onde K_1 é o número de entradas não-codificadas de C (com $v_k = 0$) e $K_2 = K - K_1$; cada transição estaria rotulada pela correspondente (K+r)upla binária, dependente do estado presente e da K-upla binária de entrada. Um código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ que utilize um codificador binário em treliça C pode, então, ser representado pela mesma treliça de C, onde a (K+r)-upla binária que rotula cada transição é substituída pela classe lateral determinada pelo mapeamento biunívoco fixo implementado pelo seletor de classes laterais em Γ/Λ. Códigos de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ que utilizam codificadores binários em treliça são denominados **códigos de classes laterais em treliça**.

Se C for um codificador binário em bloco, podemos então representar a operação de C, em resposta a uma sequência infinita de K-uplas binárias de entrada, por um caminho em uma "treliça" como na figura 4.2.8, ou seja, com apenas um único estado e tendo em cada uma de suas seções (idênticas) 2^K transições de estado partindo do único estado presente (e chegando no único estado futuro), agrupadas em um único feixe de transições paralelas (por existir apenas um único estado presente e futuro); cada transição estaria rotulada pela correspondente (K+r)upla binária, dependente apenas da K-upla binária de entrada. Um código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ que utilize um codificador binário em bloco C pode, então, ser também representado pela mesma "treliça" de C, onde a (K+r)-upla binária que rotula cada transição é substituída pela classe lateral determinada pelo mapeamento biunívoco fixo implementado pelo seletor de classes laterais em Γ/Λ. Códigos de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ que utilizam codificadores binários em bloco são denominados **códigos de classes laterais em bloco**.

Note-se que a união das 2^K classes laterais de Λ que rotulam as transições de estado que partem de um dado estado s^i=s em um código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ (em treliça ou em bloco), formam um empacotamento periódico Π_s (que independe de i), cujo volume é igual à $2^{-K}.V(\Lambda)$ (independentemente de s); este volume é definido como sendo o **volume fundamental** $V(\mathbf{C})$ de $\mathbf{C}(\Gamma/\Lambda,C)$. Como $V(\Lambda) = 2^{K+r}.V(\Gamma)$, tem-se que $V(\mathbf{C}) = 2^r.V(\Gamma)$. [É interessante notar que, como um codificador binário em bloco possui apenas um estado $s^i = 0$ quando visto como um codificador binário em treliça (com $v = 0$), os códigos de classe lateral $\mathbf{C}(\Gamma/\Lambda,C)$ que utilizam codificadores binários em bloco C podem ser vistos como simples seletor de classe lateral de Λ contidas em Π_0, onde Π_0 é o empacotamento periódico (único) associado ao estado (único) de $\mathbf{C}(\Gamma/\Lambda,C)$, como descrito acima.]

Define-se a **distância quadrática mínima** $d^2_{MIN}(\mathbf{C})$ de $\mathbf{C}(\Gamma/\Lambda,C)$ como sendo a menor distância quadrática entre duas sequências de pontos em $\mathbf{C}$. Igualmente, define-se o **coeficiente de erro** ou o **número de vizinhos mais próximos** $M_0(\mathbf{C})$ como sendo o número médio de sequências de pontos em $\mathbf{C}$ que diferem, de uma dada sequência de pontos em $\mathbf{C}$, por $d^2_{MIN}(\mathbf{C})$ e que diferem pela primeira vez desta em um dado intervalo de sinalização; da mesma forma, $\tilde{M}_0(\mathbf{C}) \triangleq (2/n).M_0(\mathbf{C})$ é denominado o **coeficiente normalizado de erro**. Assim, definem-se o **ganho nominal de codificação** (ou **ganho de codificação fundamental**) $\gamma(\mathbf{C}) \triangleq d^2_{MIN}(\mathbf{C})/V(\mathbf{C})^{2/n}$, a **perda pelo coeficiente de erro** (ou a **perda pelo número de vizinhos mais próximos**) $\eta_0(\mathbf{C})_{dB} \triangleq (3.01/\ln(1/P_e)).\log_2(\tilde{M}_0(\mathbf{C})/4)$ (para uma dada P_e suficientemente pequena tal que $M_0(\mathbf{C}) << 1/P_e$ e $n << 1/P_e$), e o **ganho efetivo de codificação** $\gamma_{eff}(\mathbf{C})_{dB} \triangleq \gamma(\mathbf{C})_{dB} - \eta_0(\mathbf{C})_{dB}$ de $\mathbf{C}(\Gamma/\Lambda,C)$, analogamente aos respectivos parâmetros de um reticulado n-dimensional.

Visto o conceito de código de classes laterais, vejamos como utilizá-lo na construção de codificadores (em bloco e em treliça); a figura 4.3.3 ilustra a forma geral da construção. Cada b-upla binária de entrada é dividida em duas partes: uma K-upla binária de entrada para o código de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$, e uma (b-K)-upla binária de entrada para o seletor de ponto, onde $c^i + \Lambda$ é a classe lateral gerada por $\mathbf{C}(\Gamma/\Lambda,C)$ e $\Re$ é uma região limitada do $\mathbb{R}^n$; a saída $c^i + \lambda^i$ é determinada pelo mapeamento biunívoco, entre o conjunto de todas as 2^{b-K} possíveis (b-K)-uplas binárias e o conjunto de todos os 2^{b-K} pontos distintos de $c^i + \Lambda$ pertencentes a $\Re$, implementado pelo seletor de ponto. Denotaremos o codificador obtido por $\mathcal{C}(\mathbf{C},\Re)$

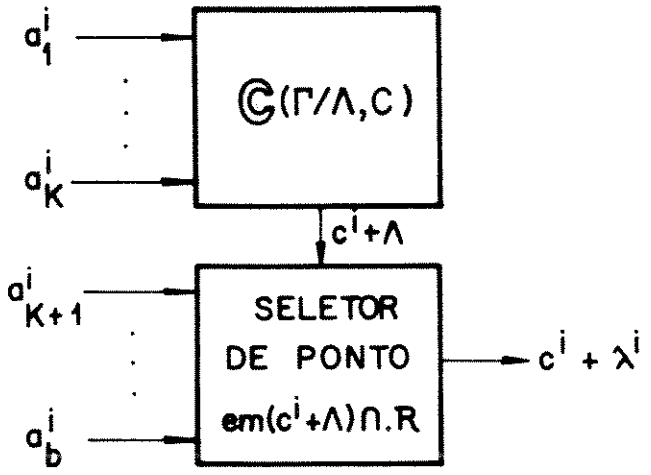

Fig. 4.3.3 - Codificador $\mathcal{C}(\mathbb{C},\Re)$

Note-se que se $\mathbf{C}(\Gamma/\Lambda,C)$ utilizar um codificador em bloco C com K=0 (sem entradas codificadas), a saída de $\mathbf{C}(\Gamma/\Lambda,C)$ será uma sequência de classes laterais de Λ em Γ/Λ todas iguais a uma classe lateral fixa (digamos, o próprio Λ); assim, neste caso, $\mathcal{C}(\mathbf{C},\Re)$ coincide com $\mathcal{C}(\Lambda,\Re)$. Mais genericamente, para um codificador binário em bloco C qualquer, vimos que $\mathbf{C}(\Gamma/\Lambda,C)$ funciona como um simples seletor de classes laterais de Λ contidas num empacotamento periódico Π_0 (descrito acima); assim $\mathcal{C}(\mathbf{C},\Re)$ constitui um codificador "reticulado" $\mathcal{C}(\Pi_0,\Re)$, ou seja, um codificador que mapeia cada uma das 2^b b-uplas binárias distintas possíveis de entrada em um ponto distinto da constelação constituída por 2^b pontos de Π_0 pertencentes a $\Re$. Isto evidencia que, para codificadores binários em treliça C, $\mathbf{C}(\Gamma/\Lambda,C)$ seja uma generalização dos conceitos de reticulado e empacotamento periódico, sugerindo que existam relacionamentos entre parâmetros de $\mathcal{C}(\mathbf{C},\Re)$, $\mathbf{C}(\Gamma/\Lambda,C)$ e $\Re$, de forma similar aos existentes entre $\mathcal{C}(\Lambda,\Re)$, Λ e $\Re$; de fato, existem relacionamentos idênticos àqueles vistos acima, ou seja, para $\mathcal{C}(\mathbf{C},\Re)$ temos $d^2_{MIN}(\mathcal{C}) = d^2_{MIN}(\mathbf{C})$, $\tilde{M}_0(\mathcal{C}) = \tilde{M}_0(\mathbf{C})$, $\eta_0(\mathcal{C}) = \eta_0(\mathbf{C})$, $\gamma(\mathcal{C})_{dB} \approx \gamma(\mathbf{C})_{dB} + \gamma_s(\Re)_{dB}$ e $\gamma_{eff}(\mathcal{C})_{dB} \approx \gamma_{eff}(\mathbf{C})_{dB} + \gamma_s(\Re)_{dB}$ como estabelecido por Forney e Wei ([15]), para $V(\mathbf{C}) << V(\Re)$ e altas relações sinal/ruído.

Quanto à complexidade normalizada de $\mathcal{C}(\mathbf{C},\Re)$, podemos aproximar $\tilde{N}_D(\mathcal{C})$ por $\tilde{N}_D(\mathbf{C})$, como fizemos para os codificadores reticulados, sendo $\tilde{N}_D(\mathbf{C}) = (2/n).(2^{K+1} - 1).2^v + \tilde{N}_D(\Gamma/\Lambda)$, onde a segunda parcela correspondente à complexidade normalizada de se determinar, em cada intervalo de sinalização, o ponto de cada classe lateral em Γ/Λ mais próximo do ponto recebido (denomina-se esta fase por decodificação

da partição, como veremos na secão 6.2), e a primeira parcela corresponde à complexidade normalizada de se determinar, em cada intervalo de sinalização, a melhor transição de estado para cada estado futuro; este é, como vimos, o algoritmo de decodificação de Viterbi. Esta expressão para $\tilde{N}_D(\mathbf{C})$ é apenas um limitante superior no caso de códigos de classes laterais em bloco (ν=0) pois, para certas partições Γ/Λ e certos mapeamentos biunívocos implementados pelo seletor de classes laterais, outros procedimentos para decodificação de $\mathbf{C}(\Gamma/\Lambda,C)$ podem apresentar menor complexidade normalizada.

Infelizmente, no entanto, não existe relacionamentos simples entre os parâmetros de $\mathbf{C}(\Gamma/\Lambda,C)$ e os parâmetros de Γ, Λ e C (exceto para o volume fundamental $V(\mathbf{C})$ e a complexidade normalizada $\tilde{N}_D(\mathbf{C})$, como já vimos) que evidenciam as características que Γ, Λ e C devem possuir para que se gerem bons códigos de classes laterais $\mathbf{C}(\Gamma/\Lambda,C)$ para utilização em codificadores $\mathcal{C}(\mathbf{C},\Re)$. Isto se deve, evidentemente ao fato de que os relacionamentos dependem, em geral, do mapeamento biunívoco implementado pelo seletor de classes laterais. No entanto, para certos códigos binários C e certas partições Γ/Λ, os relacionamentos podem ser explícitos; alguns exemplos são dados abaixo, para ilustrar isto.

Tomemos, como primeiro exemplo, a família de códigos de classes laterais denominada Classe II, proposta por Forney ([13]), reproduzida na figura 4.3.4. A partição Λ/Λ' utilizada tem ordem 2^{2K}, e $d^2_{MIN}(\Lambda') = 2.d^2_{MIN}(\Lambda)$. Independente do mapeamento implementado pelo seletor de classes laterais, verifica-se facilmente que $d^2_{MIN}(\mathbf{C}) = d^2_{MIN}(\Lambda')$, $\tilde{M}_0(\mathbf{C}) = \tilde{M}_0(\Lambda')$ e $V(\mathbf{C}) = 2^{-K}.V(\Lambda')$; assim, $\gamma(\mathbf{C}) = 2^{2K/n}.\gamma(\Lambda')$, $\eta_0(\mathbf{C}) = \eta_0(\Lambda')$, e $\gamma_{eff}(\mathbf{C})_{dB} = (K/n)\,.\,6.02 + \gamma_{eff}(\Lambda')_{dB}$. Se ainda impusermos que $\Lambda' = R.\Lambda$, onde R é uma similaridade dobradora de norma com $|\det(R)| = 2^{n/2}$, teremos da condição $|\Lambda/\Lambda'| = 2^{2K}$ que K=n/4, e portanto $\gamma_{eff}(\mathbf{C})_{dB} = 1{,}51 + \gamma_{eff}(\Lambda')_{dB}$. Assim, para este código de classes laterais, com $\Lambda' = R.\Lambda$, devemos utilizar um reticulado Λ' de máximo ganho efetivo; obviamente, o preço pago, pela utilização de reticulados Λ' de ganho efetivo crescente, será uma complexidade correspondentemente crescente, pois neste caso tem-se $\tilde{N}_D(\mathbf{C}) = (2/n).(2^{n/4+1} - 1).2^{n/2} + \tilde{N}_D(\Lambda/\Lambda')$, sendo portanto crescente com a complexidade normalizada $\tilde{N}_D(\Lambda/\Lambda')$ de decodificação da partição Λ/Λ'.

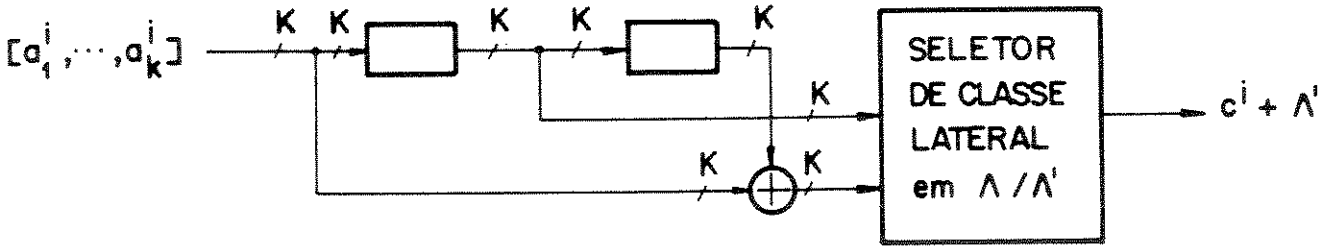

Fig. 4.3.4 - Códigos de classes laterais da família de Classe II

Como outro exemplo, tomemos o caso em que $\Gamma = \Gamma_0^{K+r}$ e $\Lambda = \Gamma_1^{K+r}$, sendo Γ_0/Γ_1 uma partição de ordem 2, com $\Gamma_0 = \Gamma_1 \bigcup (\alpha_0 + \Gamma_1)$, e façamos o mapeamento biunívoco, das 2^{K+r} (K+r)-uplas binárias nas 2^{K+r} classes laterais de Γ/Λ, da seguinte forma: a (K+r)-upla binária $(c_1,\ldots,c_{K+r})$ é mapeada na classe lateral $(c_1.\alpha_0,\ldots,c_{K+r}.\alpha_0) + \Gamma_1^{K+r}$, onde $c_i.\alpha_0 = 0$ se $c_i = 0$ e $c_i.\alpha_0 = \alpha_0$ se $c_i = 1$. A figura 4.3.5 ilustra a construção. Utilizando um codificador binário C tal que $d_H(C).d^2_{MIN}(\Gamma_0) > d^2_{MIN}(\Gamma_1)$, obtemos um código de classes laterais $\mathbb{C}(\Gamma/\Lambda,C)$ com parâmetros $d^2_{MIN}(\mathbb{C}) = d^2_{MIN}(\Gamma_1)$, $\tilde{M}_0(\mathbb{C}) = \tilde{M}_0(\Gamma_1)$ e $V(\mathbb{C}) = 2^{-K}.V(\Gamma_1)^{K+r}$; assim, $\gamma(\mathbb{C}) = 2^{2K/n}.\gamma(\Gamma_1)$, $\eta_0(\mathbb{C}) = \eta_0(\Gamma_1)$ e $\gamma_{eff}(\mathbb{C})_{dB} = \frac{K}{n}.6.02 + \gamma_{eff}(\Gamma_1)_{dB}$ onde n é a dimensão do reticulado Λ. Assim, para este código de classes laterais, com $d_H(C).d^2_{MIN}(\Gamma_0) > d^2_{MIN}(\Gamma_1)$, devemos utilizar um reticulado Γ_1 de máximo ganho efetivo; novamente, o preço pago, pela utilização de reticulados Γ_1 de ganho efetivo crescentes, será uma complexidade correspondentemente crescente, pois neste caso tem-se $\tilde{N}_D(\mathbb{C}) \approx (2/n).N_D(C) + \tilde{N}_D(\Gamma_0/\Gamma_1)$, como veremos na seção 6.3, onde $N_D(C)$ é a complexidade de decodificação com decisão suave por ML do codificador binário C (Observamos que Wolf [37] desenvolveu uma representação em treliça de códigos binários lineares em bloco de taxa K/(K+r), com 2^r estados e K+r seções, e utilizou o algoritmo de Viterbi para decodificação com decisão suave por ML desses códigos; embora esta representação seja adequada para a construção deste exemplo, algoritmos mais eficientes para decodificação de códigos binários lineares em bloco podem ser utilizados no algoritmo de decodificação de reticulados proposto na seção 6.3). Vemos que $\tilde{N}_D(\mathbb{C})$ cresce com $\tilde{N}_D(\Gamma_0/\Gamma_1)$, limitando a seleção de Γ_1.

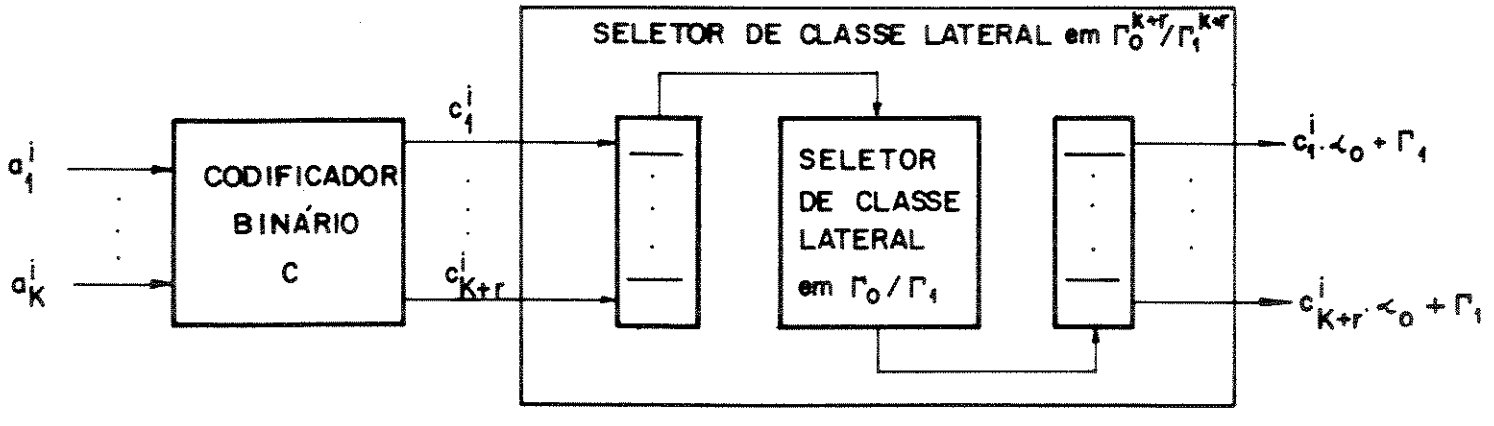

Fig. 4.3.5 - Códigos de classes laterais utilizando partições $\Gamma_0^{K+r}/\Gamma_1^{K+r}$

Estes exemplos evidenciam que os reticulados utilizados na construção de códigos de classes laterais $\mathbb{C}(\Gamma/\Lambda,C)$ devem, como na construção de códigos reticulados $\mathbb{C}(\Lambda,\Re)$, possuir geralmente ganho efetivo substancial, desde que não comprometam a complexidade de decodificação de $\mathbb{C}(\Gamma/\Lambda,C)$.

Para finalizar esta breve descrição das técnicas de construção de codificadores para o canal AWGN limitado em banda, utilizando reticulados mencionamos as denominadas **construções multi-nível** de codificadores baseados em reticulados. São generalizações das construções procedentes que utilizam uma **cadeia de partições** de reticulados $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$, onde qualquer ponto $\gamma_0 \in \Gamma_0$ pode ser expresso como $\gamma_0 = c_0 + \ldots + c_{b-1} + \gamma_b$ com $\gamma_b \in \Gamma_b$ e $c_i \in [\Gamma_i/\Gamma_{i+1}]$, $0 \le i \le b-1$, sendo $[\Gamma_i/\Gamma_{i+1}]$ um conjunto qualquer de representantes de classes laterais para a partição Γ_i/Γ_{i+1}. A forma geral dessas construções está ilustrada na figura 4.3.6. Muitas construções multi-nível de codificadores baseados em reticulados tem sido propostas; dentre estas podemos citar as de Cusack ([11]), Sayegh ([28]), Tanner ([32]) e Calderbank ([4]).

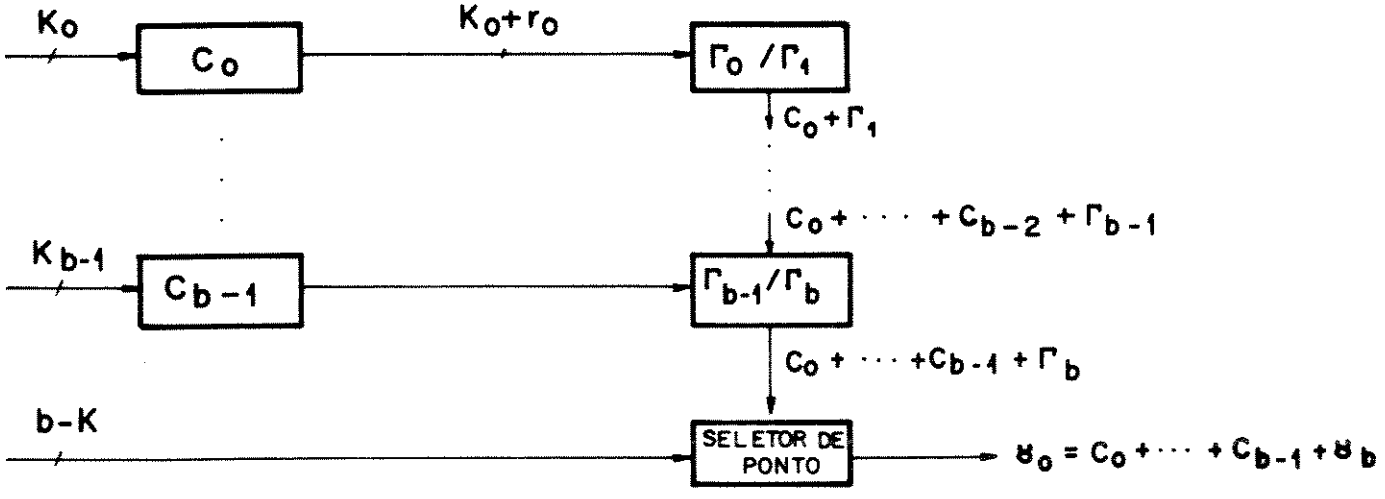

Fig. 4.3.6 - Construções multi-nível de codificadores baseados em reticulados.

Finalizando este capítulo, reproduzimos na Fig. 4.3.7 (Forney[13]) os códigos de classes laterais conhecidos que apresentam os melhores compromissos entre desempenho e complexidade de decodificação. Note-se que esta fronteira de eficiência máxima apresenta uma tendência linear, com inclinação de 0.4 dB por oitava, em grande extensão de seu trecho central.

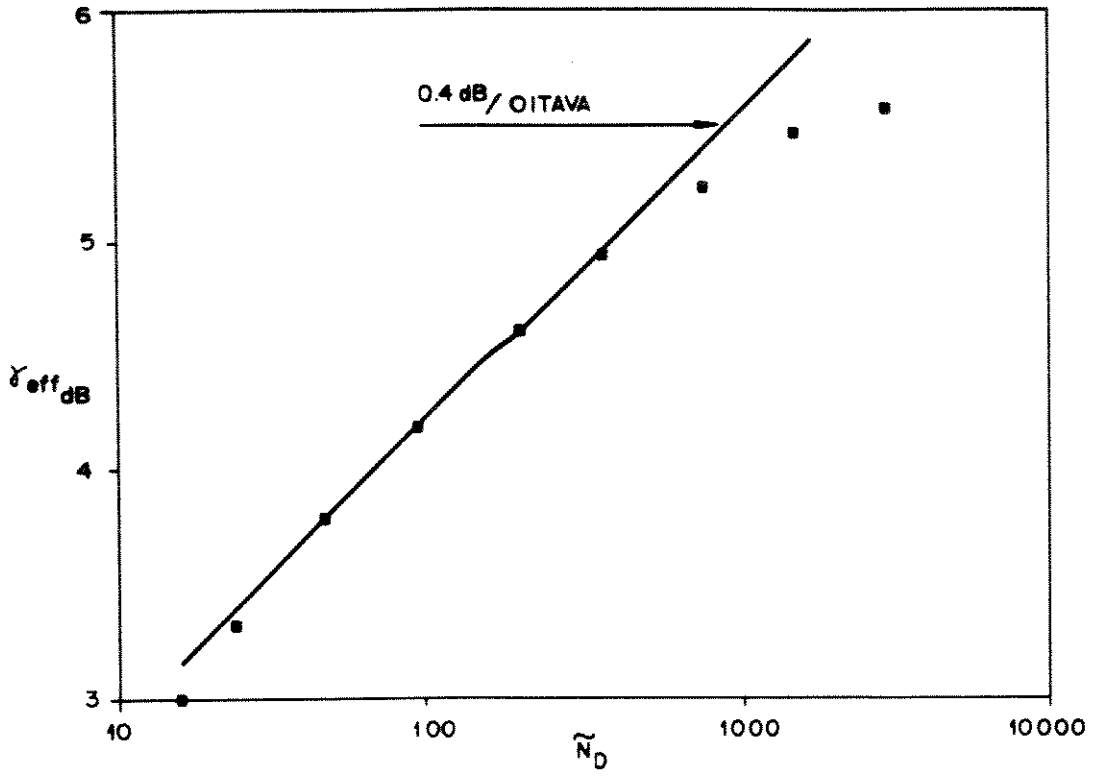

Fig. 4.3.7 - Melhores códigos de classes laterais.

Capítulo 5

Construção de reticulados

5.1 Introdução

Este capítulo trata da construção de reticulados com características adequadas para utilização em esquemas de construção, baseados em reticulados, de codificadores para canais AWGN limitados em banda. Inicialmente, na seção 5.2, são feitos alguns comentários preliminares sobre a construção de reticulados. Na seção 5.3, uma nova construção de reticulados é, então, proposta e suas características relevantes são avaliadas, sendo apresentados alguns exemplos de reticulados que podem ser obtidos com esta construção.

5.2 Comentários preliminares

As mais importantes construções de reticulados formam novos reticulados pela união de um número finito de elementos de uma partição de reticulados Γ/Λ (ou seja, pela união de um número finito de classes laterais de Λ contidas em Γ).

Mais especificamente, sejam Γ/Λ uma partição de reticulados, $[\Gamma/\Lambda]$ um conjunto de representantes e $A \subseteq \Gamma$ um subconjunto de Γ; do exposto na seção 2.2, temos que $A + \Lambda = A \text{ MOD}_{[\Gamma/\Lambda]} \Lambda + \Lambda$,onde $A \text{ MOD}_{[\Gamma/\Lambda]}\Lambda = \{a \text{ MOD}_{[\Gamma/\Lambda]}\Lambda : a \in A\} \subseteq [\Gamma/\Lambda]$; assim, ao considerarmos somas da forma $L = A + \Lambda$, poderemos, sem perda de generalidade, restringir-nos a subconjuntos $A \subseteq \Gamma$ do reticulado Γ, constituídos exclusivamente de representantes de classes laterais em um conjunto $[\Gamma/\Lambda]$ previamente estabelecido. Nestes termos, dado um subconjunto $A \subseteq [\Gamma/\Lambda]$ de representantes para uma partição Γ/Λ, a soma $L = A + \Lambda$ é um reticulado se e somente se A é um subgrupo do grupo $[\Gamma/\Lambda]$ sob adição módulo Λ, como pode ser inferido igualmente do exposto na seção 2.2. Mais genericamente, sendo $[\Gamma^m/\Lambda^m]$ um conjunto de representantes para a partição Γ^m/Λ^m, temos que, para um subconjunto de representantes $A \subseteq [\Gamma^m/\Lambda^m]$, a soma $L = A + \Lambda^m$ é um reticulado se e somente se A é um subgrupo do grupo $[\Gamma^m/\Lambda^m]$, sob adição módulo Λ em cada uma das n componentes. É importante notar que qualquer reticulado L para a cadeia de partições $\Gamma/L/\Lambda$ (resp., $\Gamma^m/L/\Lambda^m$) pode ser obtido independentemente do conjunto de representantes $[\Gamma/\Lambda]$ (resp., $[\Gamma^m/\Lambda^m]$) utilizado; em particular, podemos utilizar o conjunto básico de representantes $[\Gamma/\Lambda]_M$ (resp., $[\Gamma/\Lambda]_M^m$) associado a uma matriz geradora M (resp., M^m) de Λ (resp., Λ^m), utilizando a solução obtida na seção 3.2.

Como um exemplo ilustrativo, seja a partição de reticulados $\mathbb{Z}^2/4\mathbb{Z}^2$, tomando-se para $[\mathbb{Z}^2/4\mathbb{Z}^2]$ o conjunto básico de representantes $[\mathbb{Z}^2/4\mathbb{Z}^2]_M$, definido na seção 3.2, associado à matriz geradora $M = 4I_2$ de $4\mathbb{Z}^2$, onde I_2 é a matriz identidade de ordem 2; ou seja, $[\mathbb{Z}^2/4\mathbb{Z}^2]_M = \mathbb{Z}_4 \times \mathbb{Z}_4$ (módulo $4\mathbb{Z}^2$). Assim qualquer subgrupo A de $\mathbb{Z}_4 \times \mathbb{Z}_4$ é um subconjunto de representantes de classes laterais, de $4\mathbb{Z}^2$ contidas em $\mathbb{Z}^2$, cuja união constitui um reticulado, que evidentemente tem $4\mathbb{Z}^2$ como subreticulado e que por sua vez é

subreticulado de $\mathbb{Z}^2$. Um exemplo está ilustrado na figura 5.2.1, onde o subgrupo de $\mathbb{Z}_4 \times \mathbb{Z}_4$ selecionado é dado por $A = \{(0,0)^T,(2.0)^T,(1,2)^T,(3,2)^T\}$; note-se que o reticulado $L = A + 4\mathbb{Z}^2$ obtido tem o mesmo ganho nominal (0. dB) que $\mathbb{Z}^2$, mas, por possuir apenas dois vizinhos mais próximos a cada um de seus pontos, tem um maior ganho efetivo (0.22 dB em $P_e = 10^{-6}$) do que $\mathbb{Z}^2$ (0. dB em $P_e = 10^{-6}$), que possui quatro vizinhos mais próximos a cada um de seus pontos.

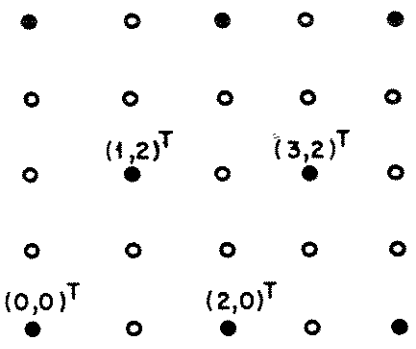

Fig. 5.2.1 - Reticulado obtido pela soma $L = A + 4\mathbb{Z}^2$, onde $A = \{(0,0)^T,(2,0)^T,(1,2)^T,(3,2)^T\}$ é um subgrupo (módulo $4\mathbb{Z}^2$) de $\mathbb{Z}_4 \times \mathbb{Z}_4$.

A busca de reticulados $L = A + \Lambda$, onde A é um subgrupo de $[\Gamma/\Lambda]$, que sejam melhores do que Γ e Λ para utilização em esquemas de codificação para canais AWGN limitados em banda, como no exemplo acima, torna-se evidentemente bastante complexa, especialmente quando a ordem $|\Gamma/\Lambda|$ da partição Γ/Λ for muito grande. Em particular, somas da forma $L = A + \Lambda^m$, onde A é um subgrupo de $[\Gamma^m/\Lambda^m]$, que sejam melhores do que Γ e Λ, no sentido acima, são difíceis de serem encontradas para valores grandes de m, mesmo quando Γ/Λ possui poucas classes laterais. Essa abordagem, então, presta-se somente à busca computacional exaustiva, não sendo, portanto, passível de tratamento analítico.

Uma extratégia alternativa para a busca sistemática de reticulados da forma $L = A + \Lambda^m$, onde A é um subgrupo de $[\Gamma^m/\Lambda^m]$ sob adição módulo Λ^m, consiste em, dado um refinamento $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ de Γ/Λ (isto é, uma cadeia com $\Gamma_0 = \Gamma$ e $\Gamma_b = \Lambda$), selecionar conjuntos de representantes $[\Gamma_i^m/\Gamma_{i+1}^m]$ e construir reticulados da forma:

$$L_i = A_i + \ldots + A_{b-1} + \Gamma_b^m \qquad (5.2.1)$$

onde $A_i \subseteq [\Gamma_i^m/\Gamma_{i+1}^m]$ são subgrupos de $[\Gamma_i^m/L_{i+1}]$ sob adição módulo L_{i+1}, desde i = b - 1 até i = 0, sendo $L_b = \Gamma_b^m$. Constrói-se, assim, o reticulado $L = L_0$ da forma desejada, ou seja, tal que $\Gamma^m/L/\Lambda^m$ seja uma cadeia de partições de reticulados. É importante salientar que os reticulados L que podem ser obtidos, com um dado refinamento $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$, dependem da particular seleção dos conjuntos de representantes $[\Gamma_i^m/\Gamma_{i+1}^m]$; no entanto, variando-se esta

seleção, qualquer reticulado L para a cadeia $\Gamma^m/L/\Lambda^m$ pode ser obtido, adotando-se um único refinamento $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ fixo. Infelizmente porém, não existem, presentemente, métodos gerais para a seleção ótima (no sentido de maximizar $\gamma(L)$ e/ou minimizar $M_0(L)$) de $[\Gamma_i^m/\Gamma_{i+1}^m]$ e A_i, $0 \le i \le b-1$.

As construções propostas na literatura adotam, então, certas seleções particulares para $[\Gamma_i^m/\Gamma_{i+1}^m]$ e A_i, $0 \le i \le b-1$, baseadas em um refinamento $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ construído de partições Γ_i/Γ_{i+1} com uma estrutura específica, denominada **partição p-elementar**: uma partição de reticulados é dita ser p-elementar se as ordens de todas as suas classes laterais forem iguais a um mesmo número primo p. Uma condição equivalente para que Γ_i/Γ_{i+1} seja p-elementar é que $p.\Gamma_i \le \Gamma_{i+1}$ (p primo); assim, do exposto na seção 3.2, tem-se que Γ_i/Γ_{i+1} é isomórfica a $\mathbb{Z}_p \times \ldots \times \mathbb{Z}_p$ (com n_i fatores, $n_i \le n$, onde n é a dimensão dos reticulados); em particular, a ordem de Γ_i/Γ_{i+1} é igual a p^{n_i}. Note-se que uma condição suficiente, mas não necessária, para que cada Γ_i/Γ_{i+1}, $0 \le i \le b-1$, seja p-elementar é que Γ_0/Γ_b seja p-elementar.

Uma vez estabelecido um refinamento $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ de Γ/Λ constituído de partições p-elementares Γ_i/Γ_{i+1}, selecionam-se conjuntos de representantes $[\Gamma_i/\Gamma_{i+1}]$ particulares (usualmente constituídos por líderes de classes laterais - ver seção 2.4), e adotam-se $[\Gamma_i/\Gamma_{i+1}]^m$ como os conjuntos de representantes $[\Gamma_i^m/\Gamma_{i+1}^m]$ para as partições $\Gamma_i^m/\Gamma_{i+1}^m$, de onde serão selecionados os conjuntos $A_i \le \left[\Gamma_i^m/L_{i+1}\right]$, $0 \le i \le b-1$. Como cada partição Γ_i/Γ_{i+1}, $0 \le i \le b-1$, é p-elementar e, portanto, isomórfica ao grupo aditivo $\mathbb{Z}_p \times \ldots \times \mathbb{Z}_p$ (n_i fatores) de $GF(p^{n_i})$, tem-se que A_i, $0 \le i \le b-1$, corresponde a um código C_i (usualmente aditivo ou linear - ver seção 2.2) sobre $GF(p^{n_i})$, de comprimento m, visto que $A_i \subseteq [\Gamma_i^m/\Gamma_{i+1}^m] = [\Gamma_i/\Gamma_{i+1}]^m$; assim, a cada $\overline{c}_i$ e C_i fica associado um elemento $\gamma_i \in A_i \subseteq [\Gamma_i/\Gamma_{i+1}]^m$, $0 \le i \le b-1$. Cada soma $\gamma_0 + \ldots + \gamma_{b-1}$ então obtida constitui o representante de uma classe lateral de $\Gamma_b^m = \Lambda^m$ na partição $\Gamma_0^m/\Gamma_b^m = \Gamma^m/\Lambda^m$; a união de todas as classes laterais de $\Gamma_b^m = \Lambda^m$ assim obtidas constitui o reticulado $L = A_0 + \ldots + A_{b-1} + \Gamma_b^m$ resultante da construção. Em geral, a condição $A_i \le [\Gamma_i^m/L_{i+1}]$ (sob adição módulo L_{i+1}) impõe restrições sobre os códigos C_i, $0 \le i \le b-1$, que dependem da cadeia de partições $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ e dos conjuntos de representantes $[\Gamma_i/\Gamma_{i+1}]$, $0 \le i \le b-1$, selecionados. A figura 5.2.2 representa a forma geral dessas construções. Note-se a correspondência entre esta forma de construção de reticulados e a construção multi-nível de codificadores para o canal gaussiano discreto no tempo na seção 4.3; de fato, estas formas de construção parecem ter surgido independentemente em contextos distintos, e esta forma geral de construção de reticulados é, por vezes, denominada **construção multi-nível** de reticulados.

Várias construções propostas na literatura podem ser expressas nesta forma geral; dentre elas podemos citar a Construção A (Leech e Sloane [23]), a Construção B (Leech e Sloane [23]), a Construção D (Barnes e Sloane [1]), a Construção E (Bos, Conway e Sloane [3]) e a Construção

Quadrática Iterada (Forney [14]). Na seção seguinte, propomos uma nova construção que segue esta forma geral, e analisamos as suas propriedades.

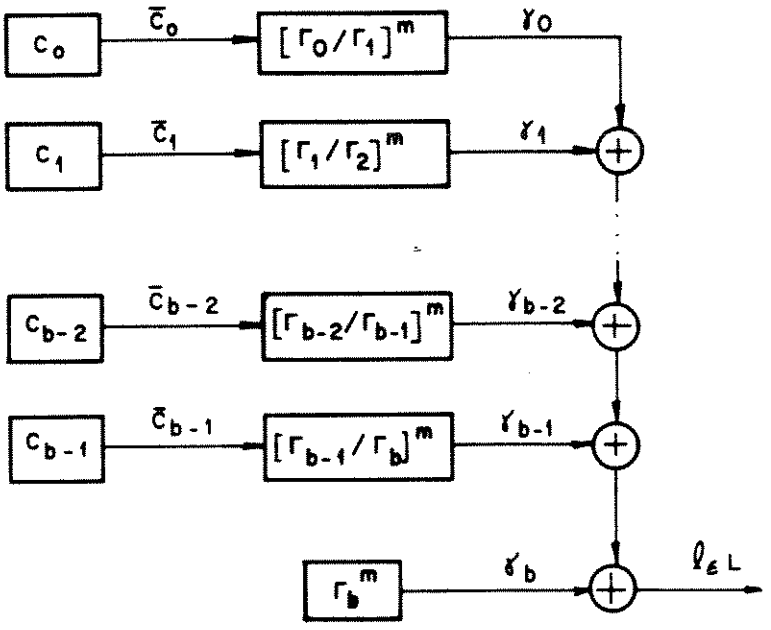

Fig. 5.2.2 - Forma geral de uma construção de reticulados utilizando códigos C_i sobre $GF(p^n)$, $0 \le i \le b-1$, de comprimento m e uma cadeia de partições elementares Γ_i/Γ_{i+1} de ordem p^n, $0 \le i \le b-1$, com $\Gamma_0 = \Gamma$ e $\Gamma_b = \Lambda$.

5.3 Construção Binária Multinível

Seja Γ/Λ uma partição binária elementar de ordem 2^b, e seja $\{\alpha_0,\ldots,\alpha_{b-1}\}$ um conjunto de vetores que constitua uma base binária para a partição Γ/Λ; isto é, tal que:

$$[\Gamma/\Lambda] = \{c_0 \cdot \alpha_0 + \ldots + c_{b-1} \cdot \alpha_{b-1} :\ c_i \in \{0;1\} \subseteq \mathbb{Z},\ i = 0,\ldots,b-1\}$$

seja um conjunto de representantes das classes laterais de Λ em Γ/Λ.

Seja $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ uma cadeia de b sub-partições de Γ/Λ, elementares de ordem 2, fazendo-se $\Gamma_0 = \Gamma, \Gamma_b = \Lambda$ e $[\Gamma_i/\Gamma_{i+1}] = \{0;\alpha_i\} = \{c_i \cdot \alpha_i :\ c_i \in \{0;1\} \subseteq \mathbb{Z}\}$, i=0, ... ,b-1. Note-se que, como $\alpha_i \notin \Gamma_{i+1}$, temos que todo ponto de $\alpha_i + \Gamma_{i+1}$ é não-nulo (i.e., distinto da origem) e pertence a $\Gamma_i = \Gamma_{i+1} \bigcup (\alpha_i + \Gamma_{i+1})$, tendo norma maior ou igual a $N(\alpha_i + \Gamma_{i+1})$.

Definição: Construção Binária Multinível

Sendo $\Gamma_0/\Gamma_1/\ldots/\Gamma_{b-1}/\Gamma_b$ uma cadeia de partições de reticulados como acima, e sendo $C_0, C_1,\ldots, C_{b-1}$ códigos binários lineares de comprimento m, definamos o seguinte arranjo periódico L:

$$L \triangleq \{ \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} : \bar{c}_i \in C_i, i=1,\ldots, b-1\} + \Gamma_b^m$$

onde, para $\bar{c}_i = (c_{i,1},\ldots, c_{i,m})^T \in C_i$, $\bar{c}_i \otimes \alpha_i \triangleq (c_{i,1} \cdot \alpha_i,\ldots, c_{i,m}.\alpha_i)^T$, sendo $\bar{c}_i$ interpretado como um vetor real de coordenadas $c_{i,j} \in \{0;1\} \subseteq \mathbb{Z}$, i=0,...,b-1.

Indicaremos a Construção Binária Multinível pela expressão:

$$L = C_0 \otimes [\Gamma_0/\Gamma_1] + \ldots + C_{b-1} \otimes [\Gamma_{b-1}/\Gamma_b] + \Gamma_b^m$$

Teorema 5.3.1: L é um reticulado

Prova: Seja $l \in L$. Logo $l = \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$, com $\bar{c}_i \in C_i$, i = 0, ... ,b-1, e $\gamma_b \in \Gamma_b^m$. Como Γ/Λ é uma partição binária elementar, $2.\alpha_i \in \Lambda = \Gamma_b$, e logo $2.\bar{c}_i \otimes \alpha_i \in \Gamma_b^m$. Assim, fazendo $\gamma'_b = 2.\bar{c}_1 \otimes \alpha_1 + \ldots + 2.\bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$ e $l' = \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} - \gamma'_b$, temos $\gamma'_b \in \Gamma_b^m$, $l' \in L$, e $l + l' = 0$. Logo, para todo $l \in L$, existe um $l' \in L$ tal que $l + l' = 0$.

Sejam l, $l' \in L$. Logo $l = \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$ e $l' = \bar{c}_0' \otimes \alpha_0 + \ldots + \bar{c}'_{b-1} \otimes \alpha_{b-1} + \gamma'_b$, com $\bar{c}_i, \bar{c}'_i \in C_i$, i=0, ... ,b-1 e $\gamma_b, \gamma'_b \in \Gamma_b^m$. Assim $l + l' = (\bar{c}_0 + \bar{c}'_0) \otimes \alpha_0 + \ldots + (\bar{c}_{b-1} + \bar{c}'_{b-1}) \otimes \alpha_{b-1} + (\gamma_b + \gamma'_b)$. Observando que $\bar{c}_i + \bar{c}'_i = (\bar{c}_i \oplus \bar{c}'_i) + 2.(\bar{c}_i * \bar{c}'_i)$, onde $\bar{c}_i \oplus \bar{c}'_i$ é a soma binária (módulo 2) de $\bar{c}_i$ com $\bar{c}'_i$ e $(\bar{c}_i * \bar{c}'_i)$ é o produto componente-por-componente de $\bar{c}_i$ com $\bar{c}'_i$, temos que $l + l' = \bar{c}''_0 \otimes \alpha_0 + \ldots + \bar{c}''_{b-1} \otimes \alpha_{b-1} + \gamma''_b$, onde $\bar{c}''_i = (\bar{c}_i \oplus \bar{c}'_i) \in C_i$, já que C_i é binário linear, e $\gamma''_b = 2.(\bar{c}_0 * \bar{c}'_0) \otimes \alpha_0 + \ldots + 2.(\bar{c}_{b-1} * \bar{c}'_{b-1}) \otimes \alpha_{b-1} + (\gamma_b + \gamma'_b) \in \Gamma_b^m$, já que Γ/Λ é binária elementar. Logo, para quaisquer l, l' e L, tem-se que $l + l' \in L$.

Logo, L é um reticulado. [CQD]

O teorema seguinte determina a distância mínima no reticulado obtido pela Construção Binária Multinível, para o caso $C_0 \subseteq \ldots \subseteq C_{b-1}$.

Teorema 5.3.2

$$d^2_{MIN}(L) = MIN[d_H(C_0).N(\alpha_0 + \Gamma_1), \ldots, d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b), d^2_{MIN}(\Gamma_b)]$$

Prova: Sendo L um reticulado, basta determinar a norma não-nula mínima de um vetor em L.

Seja, então, $l \in L$. Logo l pode ser expresso univocamente como:

$$l = \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$$

com $\bar{c}_i \in C_i$, $0 \le i \le b-1$, e $\gamma_b \in \Gamma_b^m$.

Seja i o menor inteiro, se existir, tal que $\bar{c}_i \neq 0$ na expressão de l acima, i=0, ... ,b-1.

Caso exista tal inteiro i, então pelo menos $d_H(C_i)$ componentes de l pertencem a $\alpha_i + \Gamma_{i+1}$, sendo assim não-nulas e pertencentes a Γ_i. Logo $\|l\|^2 \ge d_H(C_i).N(\alpha_i + \Gamma_{i+1})$, i=0, ... ,b-1, podendo a igualdade ser satisfeita trivialmente, visto que $C_0 \subseteq \ldots \subseteq C_{b-1}$, tomando-se l como sendo $\bar{c}_i \otimes \gamma_i$, onde $\bar{c}_i \in C_i$ é selecionado como sendo uma das palavras de peso mínimo em C_i, e $\gamma_i \in (\alpha_i + \Gamma_{i+1})$ como sendo um dos vetores de norma mínima $N(\alpha_i + \Gamma_{i+1})$ em $\alpha_i + \Gamma_{i+1}$.

Caso não exista $\bar{c}_i \neq 0$ na expressão de l, então $l \in \Gamma_b^m$. Logo, claramente, $\|l\|^2 \ge d^2_{MIN}(\Gamma_b)$, podendo a igualdade ser satisfeita trivialmente tomando-se l como tendo apenas uma componente não-nula e igual a um vetor de norma $d^2_{MIN}(\Gamma_b)$ em Γ_b.

Como todos os casos acima são mutuamente excludentes, tem-se:

$$d^2_{MIN}(L) = \min[d_H(C_0).N(\alpha_0 + \Gamma_1), \ldots, d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b), d^2_{MIN}(\Gamma_b)]. \text{ [CQD]}$$

O teorema seguinte determina, de forma explícita, o coeficiente de erro (número de vizinhos mais próximos) de um reticulado obtido pela Construção Binária Multinível, para o caso $C_0 \subseteq \ldots \subseteq C_{b-1}$.

Teorema 5.3.3

$$M_0(L) = \sum_{\substack{i=0 \\ d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)}}^{b-1} \sum_{\substack{\bar{c}_i \in C_i \\ w_H(\bar{c}_i) = d_H(C_i)}} M^{(i)}(\bar{c}_i) + \overline{M}_b$$

onde

$$M^{(i)}(\bar{c}_i) = \begin{cases} M_0(\alpha_{b-1} + \Gamma_b)^{d_H(C_{b-1})}, & i = b-1 \\ \displaystyle\sum_{\substack{\bar{c}_j \in C_j \\ \bar{c}_j * \bar{c}_i = \bar{c}_j \\ i < j \le b-1}} \prod_{\substack{k=1 \\ c_{i,j} \neq 0}}^{m} M_0^{(j)}(\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b), & 0 \le i \le b-1 \end{cases}$$

sendo $M_0^{(j)}(S)$ o número de pontos de S com norma igual a $N(\alpha_i + \Gamma_{i+1})$, e:

$$\overline{M}_b = \begin{cases} m.M_0(\Gamma_b), & \text{se } d^2_{MIN}(\Gamma_b) = d^2_{MIN}(L) \\ 0, & \text{se } d^2_{MIN}(\Gamma_b) < d^2_{MIN}(L) \end{cases}$$

Prova: Novamente, como L é um reticulado, basta determinar o número de vizinhos mais próximos da origem, ou seja, o número de vetores de L com norma não-nula mínima $d^2_{MIN}(L)$. Como na prova do teorema anterior, seja $l \in L$ de norma não-nula, e seja i o menor inteiro, se existir, tal que $\bar{c}_i \neq 0$ na expressão de l, i=0,...,b-1.

Caso exista tal inteiro i, a norma de l pode ser igual a $d^2_{MIN}(L)$ somente se $w_H(\bar{c}_i) = d_H(C_i)$ e $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)$, pois $C_0 \subseteq \ldots \subseteq C_{b-1}$. Assim, neste caso, um vetor l, com norma $d^2_{MIN}(L)$, poderia assumir

$$\sum_{\substack{i=0 \\ d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L_i)}} \; \sum_{\substack{\bar{c}_i \in C_i \\ w_H(\bar{c}_i) = d_H(C_i)}} M^{(i)}(\bar{c}_i)$$

valores possíveis, onde $M^{(i)}(\bar{c}_i)$ é o número de vetores l possíveis, com norma $d^2_{MIN}(L)$, para i e $\bar{c}_i$ dados tal que $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)$ e $w_H(\bar{c}_i) = d_H(C_i)$. Agora, já que as $d_H(C_i)$ componentes l_k de l, para as quais $c_{i,k} \neq 0$, contribuem com pelo menos $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)$ para a norma de l, podemos afirmar que um vetor l com norma $d^2_{MIN}(L)$ deve ter suas demais componentes l_k, para as quais $c_{i,k} = 0$, iguais a zero. Assim, as demais palavras $\bar{c}_j \in C_j$, $i < j \leq b-1$, na expressão de l podem ser não-nulas, desde que $c_{j,k} = 0$ sempre que $c_{i,k} = 0$, para $1 \leq k \leq m$, ou seja, $\bar{c}_j * \bar{c}_i = \bar{c}_j$. Finalmente, temos que a componente l_k, para a qual $c_{i,k} \neq 0$, pode ser qualquer vetor em $\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b$ com norma igual a $N(\alpha_i + \Gamma_{i+1})$, para termos l com norma $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)$. Logo, para $0 \leq i < b-1$, temos

$$M^{(i)}(\bar{c}_i) = \sum_{\substack{\bar{c}_j \in C_j \\ \bar{c}_j * \bar{c}_i = \bar{c}_j \\ i < j \leq b-1}} \; \prod_{\substack{k=1 \\ c_{j,k} \neq 0}}^{m} M_0^{(j)}(\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b)$$

possibilidades para l, dados i e $\bar{c}_i \in C_i$. No caso i = b - 1, a expressão acima reduz-se a

$$M^{(b-1)}(\bar{c}_{b-1}) = \prod_{\substack{k=1 \\ c_{b-1,k} \neq 0}}^{m} M_0^{(b-1)}(\alpha_{b-1} + \Gamma_b) = M_0(\alpha_{b-1} + \Gamma_b)^{d_H(C_{b-1})}$$

Nas expressões acima para $M^{(i)}(\bar{c}_i)$, indicamos por $M_0^{(i)}(S)$ o número de pontos de S com norma igual a $N(\alpha_i + \Gamma_{i+1})$.

Caso não exista $\bar{c}_i \neq 0$ na expressão de l, então $l \in \Gamma_b^m$. Logo, claramente, se $d^2_{MIN}(\Gamma_b) = d^2_{MIN}(L)$ então l deve possuir apenas uma componente não-nula. Assim, neste caso, temos

$$\overline{M}_b = m.M_0(\Gamma_b)$$

possibilidades para l, de modo que l tenha norma igual a $d^2_{MIN}(L)$. Se $d^2_{MIN}(\Gamma_b) > d^2_{MIN}(L)$, obviamente l não pode ter norma igual a $d^2_{MIN}(L)$, ou seja, $\overline{M}_b = 0$.

Como todos os casos acima são mutuamente excludentes, segue-se o resultado enunciado no teorema. [CQD]

Observe-se, pelo teorema 5.3.3, que a determinação de $M_0(L)$ exige um conhecimento bastante detalhado dos códigos C_i e dos reticulados Γ_i, diferentemente da determinação de $d^2_{MIN}(L)$, dada pelo teorema 5.3.2, que exige apenas as distâncias mínimas $d_H(C_i)$ e $N(\alpha_i + \Gamma_{i+1})$. No entanto, uma vez que estes detalhes sobre C_i e Γ_i sejam determinados, o processo de cálculo é direto. Em particular, observe-se que $M^{(i)}(\bar{c}_i)$ deve ser calculado apenas para aqueles valores de i tais que $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)$ e, dados estes valores de i, apenas para aquelas palavras $\bar{c}_i \in C_i$, tais que $w_H(\bar{c}_i) = d_H(C_i)$. Além disso, observe-se que $M_0^{(i)}(\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b)$, k=1, ... ,m, onde $c_{i,k} \neq 0$, são calculados apenas para aquelas palavras $\bar{c}_j \in C_j$, j=i+1, ... ,b-1, tais que $\bar{c}_j * \bar{c}_i = \bar{c}_j$, ou seja, que estejam contidas em $\bar{c}_i$. Observe-se, finalmente, que $M_0^{(i)}(\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b)$ é igual a zero se a norma mínima $N(\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b)$ em $\alpha_i + c_{i+1,k} \cdot \alpha_{i+1} + \ldots + c_{b-1,k} \cdot \alpha_{b-1} + \Gamma_b$ (necessariamente maior ou igual a $N(\alpha_i + \Gamma_{i+1})$) for maior que $N(\alpha_i + \Gamma_{i+1})$, reduzindo ainda mais, portanto, o número de palavras $\bar{c}_j \in C_j$ utilizadas no cálculo de $M^{(i)}(\bar{c}_i)$.

O volume fundamental de L, V(L), é igual ao volume fundamental de Γ_b^m dividido por $|L/\Gamma_b^m|$. Como $|L/\Gamma_b^m| = \prod_{i=0}^{b-1} |C_i|$, $|C_i| = 2^{k_i}$, onde k_i é a dimensão do código binário linear C_i, $0 \leq i \leq b-1$, e $V(\Gamma_b^m) = V(\Gamma_b)^m$, resulta:

$$V(L) = V(\Gamma_b)^m . 2^{-\sum_{i=0}^{b-1} k_i}$$

Observando-se que a dimensão de L é m vezes a dimensão de Γ_b, ou seja:

$$D(L) = m.D(\Gamma_b)$$

podemos, com os valores de $d^2_{MIN}(L)$ e $M_0(L)$ dados pelos teoremas acima, calcular os ganhos nominal e efetivo de codificação do reticulado L obtido pela construção binária multinível.

Alguns exemplos de reticulados L que podem ser obtidos com esta construção estão descritos nas Tabelas I-VII, a seguir. As cadeias de reticulados utilizadas referem-se ao sistema de coordenadas utilizadas por Forney ([14]); em particular utilizamos o reticulado $J_8 = D_8^1 \bigcup (D_8^1 + (10100000)^T)$. Para cada cadeia utilizada, selecionou-se os códigos binários conhecidos que maximizam o ganho nominal de L, sob a condição $C_0 \subseteq \ldots \subseteq C_{b-1}$, utilizando a referência [34].

Tabela I

$L = C_0 \otimes [\mathbb{Z}^2/D_2] + C_1 \otimes [D_2/2\mathbb{Z}^2] + (2\mathbb{Z}^2)^m$

C_0	C_1	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(8,4,4)	(8,7,2)	H_{16}	2^5	4	4,14	284
(16,11,4)	(16,15,2)	X_{32}	2^6	4	4,89	1244

Tabela II

$L = C_0 \otimes [\mathbb{Z}^4/D_4] + C_1 \otimes [D_4/R\mathbb{Z}^4] + C_2 \otimes [R\mathbb{Z}^4/RD_4] + C_3 \otimes [RD_4/2\mathbb{Z}^4] + (2\mathbb{Z}^4)^m$

C_0	C_1	C_2	C_3	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(4,1,4)	(4,3,2)	(4,3,2)	(4,4,1)	H_{16}	2^5	4	4,14	284
(8,4,4)	(8,7,2)	(8,7,2)	(8,8,1)	X_{32}	2^6	4	4,89	1244

Tabela III

$L = C_0 \otimes [D_4/R\mathbb{Z}^4] + C_1 \otimes [R\mathbb{Z}^4/RD_4] + C_2 \otimes [RD_4/2\mathbb{Z}^4] + C_3 \otimes [2\mathbb{Z}^4/2D_4] + (2D_4)^m$

C_0	C_1	C_2	C_3	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	Λ_{16}	2^{12}	8	4,52	540
(8,4,4)	(8,4,4)	(8,7,2)	(8,7,2)	Y_{32}	2^{18}	8	5,64	3036

Tabela IV

$L = C_0 \otimes [\mathbb{Z}^8/D_8] + C_1 \otimes [D_8/J_8] + C_2 \otimes [J_8/D_8^{\perp}] + C_3 \otimes [D_8^{\perp}/E_8] + C_4 \otimes [E_8/RD_8] + C_5 \otimes [RD_8/RJ_8] + C_6 \otimes [RJ_8/RD_8^{\perp}] + C_7 \otimes [RD_8^{\perp}/2\mathbb{Z}^8] + (2\mathbb{Z}^8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,4,1)	(4,4,1)	(4,4,1)	(4,4,1)	X_{32}	2^6	4	4,89	1244

Tabela V

$L = C_0 \otimes [E_8/RD_8] + C_1 \otimes [RD_8/RJ_8] + C_2 \otimes [RJ_8/RD_8^{\perp}] + C_3 \otimes [RD_8^{\perp}/RE_8] + C_4 \otimes [RE_8/2D_8] + C_5 \otimes [2D_8/2J_8] + C_6 \otimes [2J_8/2D_8^{\perp}] + C_7 \otimes [2D_8^{\perp}/2E_8] + (2E_8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(2,0,∞)	(2,1,2)	(2,1,2)	(2,1,2)	(2,2,1)	(2,2,1)	(2,2,1)	(2,2,1)	$R.H_{16}$	2^{13}	8	4,14	284
(2,1,2)	(2,1,2)	(2,1,2)	(2,1,2)	(2,2,1)	(2,2,1)	(2,2,1)	(2,2,1)	Λ_{16}	2^{12}	8	4,52	540
(4,0,∞)	(4,0,∞)	(4,0,∞)	(4,0,∞)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	$2\,W_{32}$	2^{38}	16	4,89	732
(4,0,∞)	(4,1,4)	(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	$R.H_{32}$	2^{33}	16	5,83	5084
(4,1,4)	(4,1,4)	(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	Λ_{32}	2^{32}	16	6,02	9180

Tabela VI

$L = C_0 \otimes [J_8/D_8^{\perp}] + C_1 \otimes [D_8^{\perp}/E_8] + C_2 \otimes [E_8/RD_8] + C_3 \otimes [RD_8/RJ_8] + C_4 \otimes [RJ_8/RD_8^{\perp}] + C_5 \otimes [RD_8^{\perp}/RE_8] + C_6 \otimes [RE_8/2D_8] + C_7 \otimes [2D_8/2J_8] + (2J_8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	V(L)	$d^2_{MIN}(L)$	$\gamma(L)_{dB}$	$\tilde{M}_0(L)$
(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	(4,4,1)	(4,4,1)	Y_{32}	2^{18}	8	5,64	3036

Os resultados apresentados nessas tabelas são obtidos diretamente pelas fórmulas desenvolvidas nesta seção. Por exemplo, para o reticulado H_{16} da Tabela I, temos:

$$V(H_{16}) = \frac{V(2\mathbb{Z}^2)^8}{2^{4+7}} = \frac{4^8}{2^{11}} = 2^5$$

$$d^2_{MIN}(H_{16}) = MIN[4.N([1,0]^t + D_2),\ 2.N([1,1,]^t + 2\mathbb{Z}^2),\ d^2_{MIN}(2\mathbb{Z}^2)] =$$

$$= MIN[4.1, 2.2, 4] = 4$$

$$\gamma(H_{16})_{dB} = 10.\log_{10}\frac{4}{(2^5)^{2/16}} = 4{,}14$$

$$M_0(H_{16}) = M_0(8,4,4).[M_0^{(0)}([2,1]^t + 2\mathbb{Z}^2)^4 +$$

$$+ \binom{4}{2}.M_0^{(0)}([2,1]^t + 2\mathbb{Z}^2)^2.M_0^{(0)}([1,0]^t + 2\mathbb{Z}^2)^2 +$$

$$+ M_0^{(0)}([1,0]^t + 2\mathbb{Z}^2)^4] +$$

$$+ M_0(8,7,2).[M_0^{(1)}([1,1]^t + 2\mathbb{Z}^2)^2] +$$

$$+ [8.M_0(2\mathbb{Z}^2)] =$$

$$= 14.[2^4+6.2^2.2^2+2^2] + 28.[4^2] + [8.4] = 2.272$$

$$\tilde{M}_0(H_{16}) = (2/16)\ 2.272 = 284$$

Dentre os reticulados obtidos pela Construção Binária Multinível, dois reticulados novos de 32 dimensões, W_{32} e Y_{32}, apresentam um bom compromisso entre desempenho e complexidade, quando comparados com os melhores reticulados conhecidos nesta dimensão, como mostra Fig.5.3.1. No próximo capítulo, onde abordamos o problema da decodificação de reticulados, veremos como utilizar essa construção para reduzir a complexidade de decodificação dos reticulados obtidos nas várias dimensões, em relação aos algorítmos propostos por Forney ([14]).

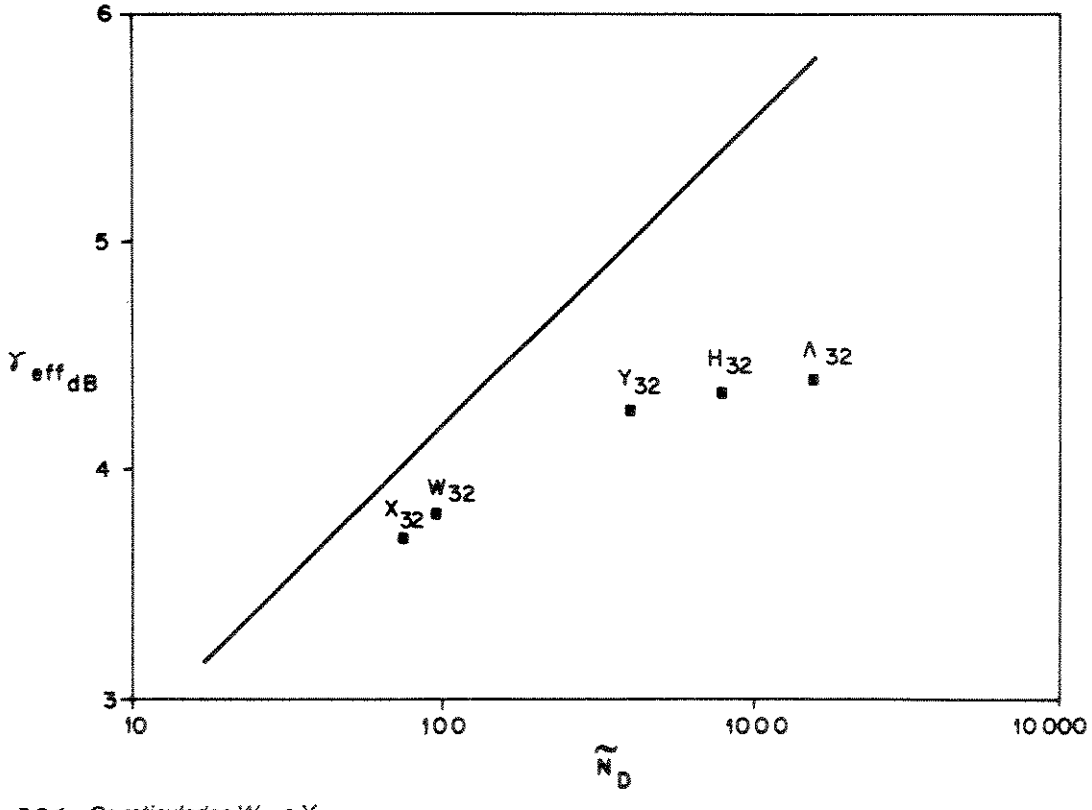

Fig.5.3.1 - Os reticulados W_{32} e Y_{32}.

Capítulo 6

Decodificação de reticulados

6.1 Introdução

Este capítulo trata do problema da decodificação de reticulados utilizando algoritmos de complexidade reduzida. Inicialmente, na seção 6.2, são feitos alguns comentários preliminares sobre algoritmos de decodificação de reticulados, introduzindo-se o conceito de decodificação com distância limitada. Na seção 6.3, um algoritmo de decodificação com distância limitada é proposto para os reticulados obtidos pela Construção Binária Multinível, analizando-se suas propriedades e avaliando-se sua complexidade.

6.2 Comentários preliminares

A decodificação por máxima verosimilhança (ML) de um reticulado Γ do $\mathbb{R}^n$ consiste, como definida na seção 4.3, na determinação do ponto $\gamma \in \Gamma$ mais próximo de um ponto $r \in \mathbb{R}^n$ dado; diremos, no mesmo sentido, que $r \in \mathbb{R}^n$ é decodificado, em relação a Γ, no ponto $\gamma \in \Gamma$. Igualmente, define-se a decodificação por ML de uma classe lateral $c + \Gamma$ de Γ como sendo a determinação do ponto $(c + \gamma) \in (c + \Gamma)$ mais próximo de um ponto $r \in \mathbb{R}^n$ dado. Evidentemente, se r-c é decodificado por ML em relação a Γ no ponto $\gamma \in \Gamma$, então r é decodificado por ML em relação a $c + \Gamma$ no ponto $(c + \gamma) \in (c + \Gamma)$; assim, qualquer algoritmo para decodificação por ML de um reticulado pode ser utilizado como um algoritmo para decodificação por ML de uma classe lateral desse reticulado, desde que devidamente adaptado. Estamos interessados, geralmente, na determinação de algoritmos de decodificação por ML eficientes, ou seja, que requeiram o menor número possível de operações.

Algoritmos de decodificação por ML são, necessariamente, **invariantes por translação**, ou seja, se r é decodificado por ML em relação a Γ (resp.,$c + \Gamma$) no ponto $\gamma \in \Gamma$ (resp.,$(c + \gamma) \in (c + \Gamma)$) então, para qualquer $\gamma' \in \Gamma$, $r + \gamma'$ é decodificado por ML em relação a Γ (resp.,$c + \Gamma$) no ponto $(\gamma + \gamma') \in \Gamma$ (resp.,$(c + \gamma + \gamma') \in (c + \Gamma)$); é importante notar que esta invariância por translação só é válida, em geral, para translações γ' pertencentes a Γ.

Algoritmos eficientes de decodificação por ML são conhecidos para alguns reticulados simples, tais como $\mathbb{Z}^n$, D_n e A_n (Conway e Sloane [8],[9]). Para reticulados Γ mais complexos, a estratégia comumente empregada é a denominada **decodificação por ML de classes laterais** (Conway e Sloane [8],[9]), que envolve a determinação de um sub-reticulado Λ do reticulado Γ para o qual se conhece um algoritmo de decodificação por ML. Com um conjunto selecionado $[\Gamma/\Lambda]$ de representantes de classes laterais para a partição Γ/Λ, decodifica-se o ponto dado $r \in \mathbb{R}^n$ no ponto $c + \lambda$ mais próximo em cada classe lateral $c + \Lambda$, $c \in [\Gamma/\Lambda]$, utilizando o algoritmo de decodificação por ML para Λ; este procedimento inicial é denominado **decodificação por ML da partição** Γ/Λ; o ponto $c + \lambda$, $c \in [\Gamma/\Lambda]$, mais próximo de r é, então, o ponto de Γ mais próximo de r. Algoritmos de decodificação por ML gerados desta forma somente serão eficientes se o algoritmo de decodificação por ML, para o sub-reticulado utilizado, for eficiente e o número de classes laterais na partição correspondente for muito pequeno; isto ocorre porque o número de operações requeridas pelo algoritmo gerado é pelo menos o número de classes laterais vezes o número de operações requeridas pelo algoritmo para a decodificação do sub-reticulado, ou seja, o número de operações requeridas para decodificação por ML da partição utilizada.

Uma técnica de redução do número de operações requeridas para a decodificação por ML da partição utilizada é selecionar adequadamente esta partição e os representantes de suas classes laterais de modo que estes apresentem alguma estrutura que possibilite sua decodificação de forma paralela. A primeira proposta neste sentido foi feita por Conway e Sloane ([9]) para a decodificação por ML de reticulados obtidos pela Construção A (uma generalização deste algoritmo será dada na próxima seção). Em seguida Forney ([14]) representando sua Construção Quadrática iterada por treliças regulares, propôs um algoritmo recursivo para decodificação dos reticulados obtidos. O reticulado de Leech ([22]) tem sido objeto de intensa aplicação dessas idéias ([9],[14],[2],[21]).

Uma forma efetiva para reduzir a complexidade de decodificação de um reticulado Γ consiste em nos contentarmos com um algoritmo que garanta decodificação por ML ao menos para pontos suficientemente próximos de Γ; esta forma de decodificação é denominada **decodificação com distância limitada**, e a distância, a partir de um ponto do reticulado Γ, dentro do qual o algoritmo A garante decodificação por ML é denominado **raio de correção efetivo** do algoritmo A para decodificação de Γ, denotado $r_{eff}^{A}(\Gamma)$; obviamente, $r_{eff}^{A}(\Gamma) \leq r(\Gamma) \triangleq \frac{1}{2}.d_{MIN}(\Gamma)$. Um algoritmo A, para decodificação de Γ, define em torno de cada ponto $\gamma \in \Gamma$ uma **região de decisão** $R_{\Gamma}^{A}(\gamma)$, formada por todos os pontos que são decodificados por A no ponto γ, que contém a hiperesfera de raio $r_{eff}^{A}(\Gamma)$ centrada em γ (note-se que para um algoritmo de decodificação por ML, a região de decisão em torno de $\gamma \in \Gamma$ seria a região de Voronoi $V_{\Gamma}(\gamma)$); se o algoritmo for invariante por translação, então todas as regiões de decisão serão isométricas, e $R_{\Gamma}^{A}(\gamma) = \gamma + R_{\Gamma}^{A}(0), \gamma \in \Gamma$, onde $R_{\Gamma}^{A}(0)$ é a região de decisão na origem, sendo portanto uma região fundamental para Γ; neste caso, o número de pontos na fronteira de $R_{\Gamma}^{A}(0)$ com norma igual a $r_{eff}^{A}(\Gamma)^2$ (ou seja, o número de pontos de contato da hiperesfera de raio $r_{eff}^{A}(\Gamma)$ centrada na origem com a fronteira de $R_{eff}^{A}(\Gamma)$) é denominado o **coeficiente efetivo de erro** de Γ para decodificação por A, denotado $M_{0_{eff}}^{A}(\Gamma)$; em geral, $M_{0_{eff}}^{A}(\Gamma) \geq M_0(\Gamma)$. Se um codificador reticulado $\mathcal{C}$, formado pelos pontos de Γ dentro de uma região R limitada, for decodificado utilizando um algoritmo A com distância limitada, a probabilidade de erro P_e terá uma aproximação idêntica, em forma, a (4.24) com $M_{0_{eff}}^{A}(\Gamma)$ e $(2.r_{eff}^{A}(\Gamma))^2$ em lugar de $M_0(\mathcal{C}) \approx M_0(\Gamma)$ e $d_{MIN}^2(\mathcal{C}) = d_{MIN}^2(\Gamma)$, respectivamente, reduzindo assim seu ganho efetivo; em particular, se $r_{eff}^{A}(\Gamma) = \frac{1}{2}.d_{MIN}(\Gamma)$ então a redução no ganho efetivo do codificador reticulado $\mathcal{C}$ é devido apenas ao aumento efetivo no coeficiente de erro, sendo dada aproximadamente por:

$$\Delta\gamma_{eff}^{A}(\mathcal{C}) = \gamma_{eff}(\mathcal{C})_{dB} - \gamma_{eff}^{A}(\mathcal{C})_{dB} \approx 0.22 \times \log_2 \frac{M_{0_{eff}}^{A}(\Gamma)}{M_0(\Gamma)}$$

em $P_e = 10^{-6}$; como o ganho de forma $\gamma_s(\mathcal{C})$ depende apenas da região R, temos, igualmente, em $P_e = 10^{-6}$:

$$\Delta\gamma_{eff}^{A}(\Gamma) = \gamma_{eff}(\Gamma)_{dB} - \gamma_{eff}^{A}(\Gamma)_{dB} \approx 0.22 \times \log_2 \frac{M_{0_{eff}}^{A}(\Gamma)}{M_0(\Gamma)}$$

Assim, se o aumento no coeficiente de erro não for muito acentuado, um algoritmo simples de decodificação para Γ, com raio de correção efetivo igual a $\frac{1}{2}.d_{MIN}(\Gamma)$, pode ser mais vantajoso do que um algoritmo complexo de decodificação por ML para Γ.

A Fig.6.2.1 ilustra a redução de complexidade obtida com novos algorítmos de decodificação, por ML e com distância limitada neste último caso, um algorítmo de decodificação com distância limitada é compensador se a redução do ganho efetivo for inferior à 0.4 dB por oitava de redução de complexidade, levando-o para mais próximo da fronteira linear de eficiência máxima correntemente conhecida.

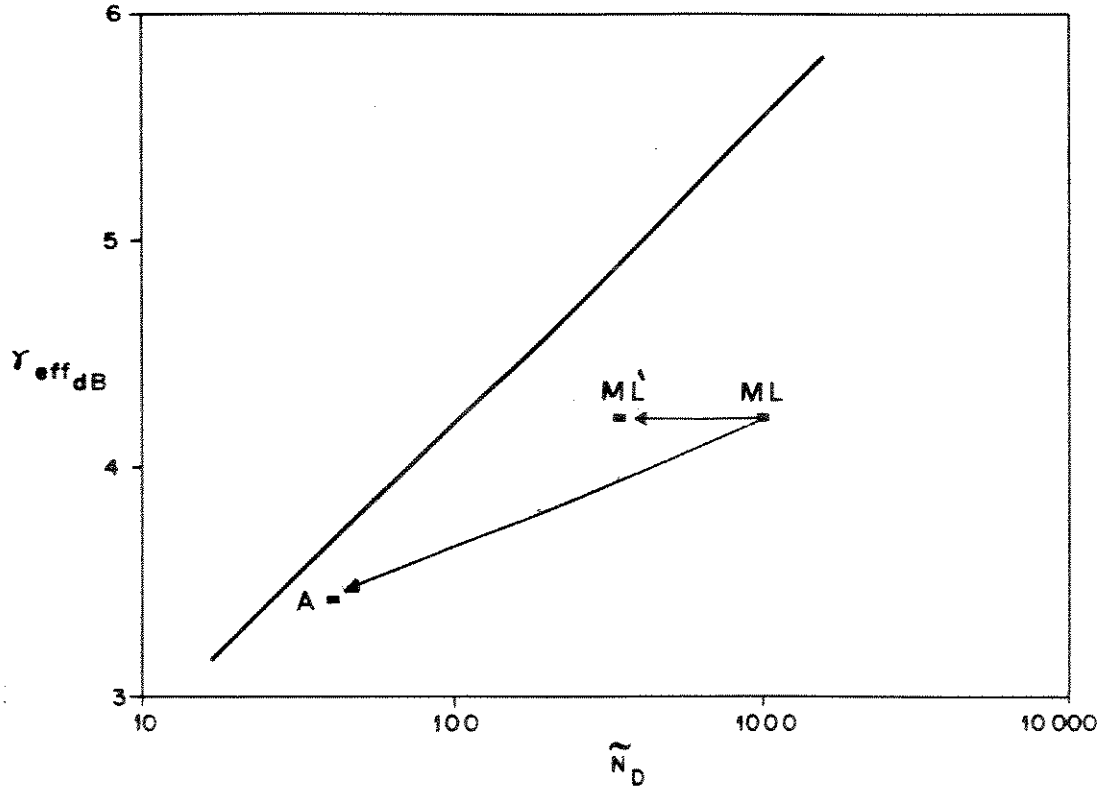

Fig.6.2.1. Redução de complexidade de decodificação.

Como exemplo elementar, seja $\Gamma = C_0 + 2.(m,m-1,2) + (4\mathbb{Z})^m$ um reticulado obtido pela Construção B (Leech e Sloane [23]), onde C_0 é um código binário linear de comprimento m e (m,m-1,2) é o código binário de comprimento m formado por todas as possíveis palavras de peso de Hamming par. A figura 6.2.2 ilustra Γ, para $C_0 = (2,1,2)$, exibindo a região de Voronoi de Γ na origem; note-se que $M_0(\Gamma) = 2$.

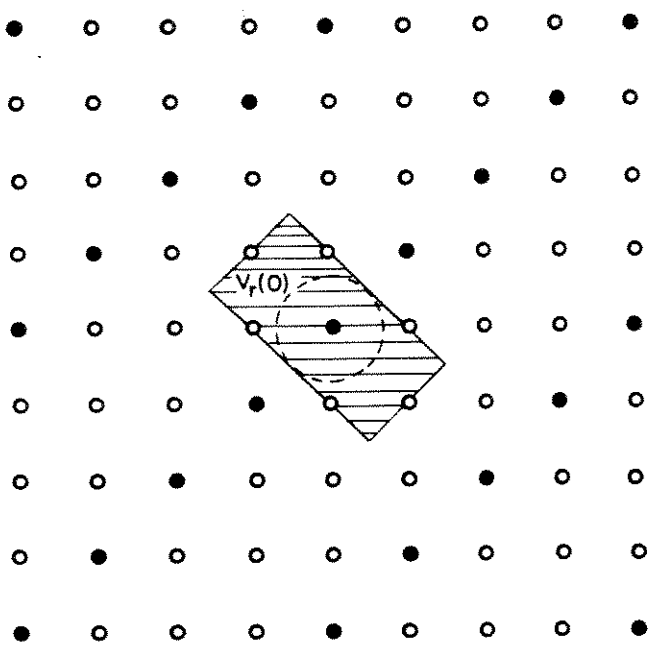

Fig.6.2.2. O reticulado $(2,1,2) + 2.(2,1,2) + (4\mathbb{Z})^2$ *(círculos cheios)*

Forney ([17]) propôs o seguinte algoritmo com distância limitada para a Construção B: dado $r \in \mathbb{R}^m$, decodifique r por ML no ponto $\gamma_0 = \bar{c}_0 + 2z_0$ mais próximo no reticulado $C_0 + (2.\mathbb{Z})^m$; verifique a paridade de z_0: se z_0 possuir um número par de coordenadas ímpares, então $\gamma_0 \in C_0 + 2.(m,m-1,2) + (4\mathbb{Z})^m$, e aceite γ_0 como o resultado da decodificação; caso contrário, identifique a componente z_{0j} de z_0, para o qual $|r_j - (c_{0j} + 2.z_{0j})|$ é máximo, e faça $z'_{0j} = z_{0j} + 1$ se $r_j > c_{0j} + 2.z_{0j}$, $z'_{0j} = z_{0j} - 1$ se $r_j < c_{0j} + 2.z_{0j}$, e $z'_{0j} = z_{0j}$ para $i \neq j$, $0 \leq i \leq m$, de modo que $\gamma'_0 = c_0 + 2.z'_0 \in C_0 + 2.(m,m-1,2) + (4\mathbb{Z})^m$, e aceite γ'_0 com o resultado da decodificação (esta técnica corretiva é conhecida como a **regra de Wagner** [20]). A figura 6.2.3 exibe a região de decisão na origem $R_\Gamma^A(0)$ desse algoritmo para o reticulado Γ da figura 6.2.2; note-se que $r_{eff}^A(\Gamma) = \frac{1}{2}.d_{MIN}(\Gamma)$ e $M_{0_{eff}}^A(\Gamma) = 4$, sendo $R_\Gamma^A(0)$ uma região fundamental para Γ.

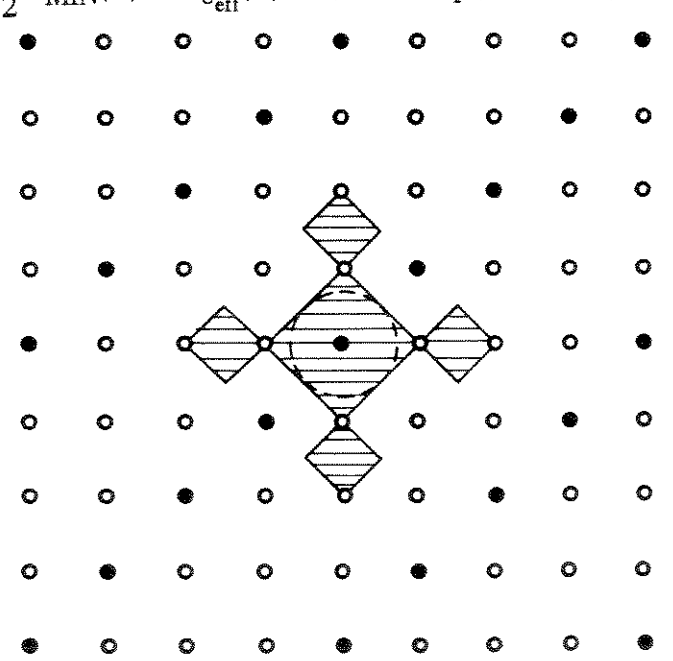

Fig.6.2.3. A região de decisão $R_\Gamma^A(0)$

Forney ([17]) extendeu o algoritmo acima a reticulados da forma $\Gamma = C_0 + 2.C_1 + \ldots + 2^{b-1}.C_{b-1} + (2^b\mathbb{Z})^m$, mantendo a condição $r_{eff}^{A}(\Gamma) = \frac{1}{2}.d_{MIN}(\Gamma)$. Algoritmos com esta estrutura são exemplos particulares da classe de algoritmos denominados de **decodificação por estágios**, que se aplicam a construção multiníveis de reticulados e, mais genericamente, de codificadores. Na próxima seção, propomos um algoritmo de decodificação por estágios para a Construção Binária Multinível, determinando seu desempenho e complexidade.

6.3 Decodificação por estágios da Construção Binária Multinível

Seja $L = C_0 \otimes [\Gamma_0/\Gamma_1] + \ldots + C_{b-1} \otimes [\Gamma_{b-1}/\Gamma_b] + \Gamma_b^m$ como definido na seção 5.3.

Algoritmo: Dado r,

0-A) Faça $r_0 = r$.

0-B) Decodifique r_0 no ponto $\hat{c}_0 \otimes \alpha_0 + \hat{\gamma}_1$ mais próximo no reticulado $\Lambda_0 = C_0 \otimes [\Gamma_0/\Gamma_1] + \Gamma_1^m$.

1-A) Faça $r_1 = r_0 - \hat{c}_0 \otimes \alpha_0$.

1-B) Decodifique r_1 no ponto $\hat{c}_1 \otimes \alpha_1 + \hat{\gamma}_2$ mais próximo no reticulado $\Lambda_1 = C_1 \otimes [\Gamma_1/\Gamma_2] + \Gamma_2^m$.

.
.
.

(b-1)-A) Faça $r_{b-1} = r_{b-2} - \hat{c}_{b-2} \otimes \alpha_{b-2}$.

(b-1)-B) Decodifique r_{b-1} no ponto $\hat{c}_{b-1} \otimes \alpha_{b-1} + \hat{\gamma}_b$ mais próximo no reticulado $\Lambda_{b-1} = C_{b-1} \otimes [\Gamma_{b-1}/\Gamma_b] + \Gamma_b^m$.

Assim, r fica decodificado no ponto $\hat{l} = \hat{c}_0 \otimes \alpha_0 + \ldots + \hat{c}_{b-1} \otimes \alpha_{b-1} + \hat{\gamma}_b$ do reticulado $L = C_0 \otimes [\Gamma_0/\Gamma_1] + \ldots + C_{b-1} \otimes [\Gamma_{b-1}/\Gamma^m] + \Gamma_b^m$.

Mostraremos que este algoritmo tem raio de correção efetivo igual a $\frac{1}{2}.d_{MIN}(L)$; antes, no entanto, estabelecemos o lema a seguir.

Lema: O algoritmo é invariante por translação.

Prova: Seja $\hat{l} \in L$ o resultado da decodificação de r pelo algoritmo.

Seja $l \in L$. Provemos que o algoritmo decodifica $r' = r + l$ em $\hat{l}' = \hat{l} + l$.

Sendo $\hat{l} = \hat{\bar{c}}_0 \otimes \alpha_0 + \ldots + \hat{\bar{c}}_{b-1} \otimes \alpha_{b-1} + \hat{\gamma}_b$ e $l = \bar{c}_0 \otimes \alpha_0 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$, com $\hat{\bar{c}}_i$, $\bar{c}_i \in C_i$, $0 \le i \le b-1$, e $\hat{\gamma}_b$, $\gamma_b \in \Gamma_b^m$, definamos, para $0 \le i \le b-1$, δ_{i+1} e l_i dados por:

$$\delta_{i+1} = (\hat{\bar{c}}_0 * \bar{c}_0) \otimes (2\alpha_0) + \ldots + (\hat{\bar{c}}_{i-1} * \bar{c}_{i-1}) \otimes (2\alpha_{i-1}) + \bar{c}_{i+1} \otimes \alpha_{i+1} + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$$

$$l_i = \hat{\bar{c}}_i \otimes \alpha_i + \delta_{i+1}$$

Assim, $\delta_{i+1} \in \Gamma_{i+1}^m$ (pois $2\alpha_j \in \Gamma_b \le \Gamma_{i+1}$, para $0 \le j \le i-1$) e, portanto, $l_i \in \Lambda_i = C_i \otimes [\Gamma_i/\Gamma_{i+1}] + \Gamma_{i+1}^m$. Provaremos inicialmente que na aplicação do algoritmo a decodificação de $r' = r + l$, r'_i é dada por $r'_i = r_i + l_i$, para $0 \le i \le b-1$. A prova é por indução sobre i, sendo trivial para i=0, pois $r'_0 = r'$, $r_0 = r$ e $l_0 = l$ (visto que $\delta_1 = \bar{c}_1 \otimes \alpha_1 + \ldots + \bar{c}_{b-1} \otimes \alpha_{b-1} + \gamma_b$). Para o passo de indução, suponhamos que $r'_i = r_i + l_i$ e provemos que $r'_{i+1} = r_{i+1} + l_{i+1}$: como decodificação ML é invariante por translação, se r_i é decodificado (ML) no ponto mais próximo $\hat{\bar{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1}$ do reticulado $\Lambda_i = C_i \otimes [\Gamma_i/\Gamma_{i+1}] + \Gamma_{i+1}^m$, então $r'_i = r_i + l_i$ será decodificado (ML) no ponto mais próximo $\hat{\bar{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1} + l_i$ do reticulado $\Lambda_i = C_i \otimes [\Gamma_i/\Gamma_{i+1}] + \Gamma_{i+1}^m$, pois $l_i \in \Lambda_i = C_i \otimes [\Gamma_i/\Gamma_{i+1}] + \Gamma_{i+1}^m$; este ponto pode, no entanto, ser expresso por:

$$\hat{\bar{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1} + l_i = \bar{c}_i \otimes \alpha_i + \hat{\gamma}_{i+1} + (\hat{\bar{c}}_i \otimes \alpha_i + \delta_{i+1}) =$$

$$= (\hat{\bar{c}}_i + \bar{c}_i) \otimes \alpha_i + (\hat{\gamma}_{i+1} + \delta_{i+1}) =$$

$$= (\hat{\bar{c}}_i \oplus \bar{c}_i) \otimes \alpha_i + (\hat{\gamma}_{i+1} + \delta_{i+1} + (\hat{\bar{c}}_i * \bar{c}_i) \otimes (2\alpha_i)) =$$

$$= \hat{\bar{c}}'_i \otimes \alpha_i + \hat{\gamma}'_{i+1}$$

onde $\hat{\bar{c}}'_i \in C_i$ e $\hat{\gamma}'_i \in \Gamma_{i+1}^m$. Assim $r'_{i+1} = r'_i - \hat{\bar{c}}'_i \otimes \alpha_i$; como $r'_i = r_i + l_i$, r'_{i+1} fica:

$$r'_{i+1} = r_i + l_i - \hat{\bar{c}}'_i \otimes \alpha_i =$$

$$= r_i + l_i - (\hat{\bar{c}}_i + \bar{c}_i - 2.(\hat{\bar{c}}_i * \bar{c}_i)) \otimes \alpha_i =$$

$$= (r_i - \hat{\bar{c}}_i \otimes \alpha_i) + (l_i - \bar{c}_i \otimes \alpha_i + (\hat{\bar{c}}_i * \bar{c}_i) \otimes (2\alpha_i)) =$$

$$= r_{i+1} + (\delta_{i+1} + (\hat{\bar{c}}_i * \bar{c}_i) \otimes (2\alpha_i)) =$$

$$= r_{i+1} + (\delta_{i+2} + \bar{c}_{i+1} \otimes \alpha_{i+1}) =$$

$$= r_{i+1} + l_{i+1}$$

tendo sido utilizado, na penúltina igualdade, a identidade $\delta_{i+1} + (\hat{c}_i * \bar{c}_i) \otimes (2\alpha_i) = \delta_{i+2} + \bar{c}_{i+1} \otimes \alpha_{i+1}$; isto conclui o passo de indução, valendo então que $r'_i = r_i + l_i$, para $0 \leq i \leq b-1$. Além disso, vemos, na prova acima, que $\hat{c}'_i = \hat{c}_i \oplus \bar{c}_i$, para $0 \leq i \leq b-1$, e que, em i=b-1, $\hat{\gamma}'_b = \hat{\gamma}_b + \delta_b + (\hat{c}_{b-1} * \bar{c}_{b-1}) \otimes (2\alpha_{b-1})$, de modo que:

$$\hat{l}' = (\hat{c}_0 \oplus \bar{c}_0) \otimes \alpha_0 + \ldots + (\hat{c}_{b-1} \oplus \bar{c}_{b-1}) \otimes \alpha_{b-1} + \hat{\gamma}_b + \delta_b + (\hat{c}_{b-1} * \bar{c}_{b-1}) \otimes (2\alpha_{b-1})$$

Como $\delta_b = (\hat{c}_0 * \bar{c}_0) \otimes (2\alpha_0) + \ldots + (\hat{c}_{b-2} * \bar{c}_{b-2}) \otimes (2\alpha_{b-2}) + \gamma_b$, $\hat{l}'$ fica:

$$\hat{l}' = (\hat{c}_0 \oplus \bar{c}_0 + 2.(\hat{c}_0 * \bar{c}_0)) \otimes \alpha_0 + \ldots + (\hat{c}_{b-1} \otimes \bar{c}_{b-1} + 2.(\hat{c}_{b-1} * \bar{c}_{b-1})) \otimes \alpha_{b-1} + (\hat{\gamma}_b + \gamma_b) =$$

$$= (\hat{c}_0 \oplus \bar{c}_0) \otimes \alpha_0 + \ldots + (\hat{c}_{b-1} + \bar{c}_{b-1}) \otimes \alpha_{b-1} + (\hat{\gamma}_b + \gamma_b) =$$

$$= \hat{l} + l \qquad \text{[CQD]}$$

Teorema 6.3.1: Sendo $l \in L$, todo r tal que $\| r - l \| < \frac{1}{2}.d_{MIN}(L)$ é decodificado corretamente pelo algoritmo, ou seja $\hat{l} = l$, no caso $C_0 \subseteq \ldots \subseteq C_{b-1}$.

Prova: Em virtude da invariância por translação do algoritmo, basta provar para $l = 0$:

Como $r_0 = r$, temos:

$$\| r_0 - (\hat{c}_0 \otimes \alpha_0 + \hat{\gamma}_1) \|^2 \leq \| r_0 \|^2 < \frac{1}{4}.d^2_{MIN}(L) \leq \frac{1}{4}.d_H(C_0).N(\alpha_0 + \Gamma_1)$$

Logo:

$$\| \hat{c}_0 \otimes \alpha_0 + \hat{\gamma}_1 \|^2 \leq (\| r_0 - (\hat{c}_0 \otimes \alpha_0 + \hat{\gamma}_1) \| + \| r_0 \|)^2 <$$

$$< \left(\left(\frac{1}{4}.d_H(C_0).N(\alpha_0 + \Gamma_1) \right)^{1/2} + \left(\frac{1}{2}.d_H(C_0).N(\alpha_0 + \Gamma_1) \right)^{1/2} \right)^2 =$$

$$= d_H(C_0).N(\alpha_0 + \Gamma_1)$$

Assim, temos necessariamente $\hat{c}_0 \otimes \alpha_0 = 0$; e portanto $r_1 = r_0$.

Igualmente, para $1 \leq i < b-1$, temos:

$$\| r_i - (\hat{\tilde{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1}) \|^2 \leq \| r_i \|^2 < \frac{1}{4}.d^2_{MIN}(L) \leq \frac{1}{4}.d_H(C_i).N(\alpha_i + \Gamma_{i+1})$$

Logo:

$$\| \hat{\tilde{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1} \|^2 \leq (\| r_i - (\hat{\tilde{c}}_i \otimes \alpha_i + \hat{\gamma}_{i+1})\| + \| r_i \|)^2 <$$

$$< \left(\left(\frac{1}{4}.d_H(C_i).N(\alpha_i + \Gamma_{i+1})\right)^{1/2} + \left(\frac{1}{4}.d_H(C_i).N(\alpha_i + \Gamma_{i+1})\right)^{1/2}\right)^2 =$$

$$= d_H(C_i).N(\alpha_i + \Gamma_{i+1})$$

Assim, temos necessariamente $\hat{\tilde{c}}_i \otimes \alpha_i = 0$; e portanto $r_{i+1} = r_i$, para $1 \leq i < b-1$.

Finalmente, em $i = b - 1$, temos também:

$$\| r_{b-1} - \hat{\gamma}_b \|^2 \leq \|r_{b-1}\|^2 < \frac{1}{4}.d^2_{MIN}(L) \leq \frac{1}{4}.d^2_{MIN}(\Gamma_b)$$

Logo:

$$\|\hat{\gamma}_b\|^2 \leq (\|r_{b-1} - \hat{\gamma}_b\| + \|r_{b-1}\|)^2 <$$

$$< \left(\left(\frac{1}{4}.d^2_{MIN}(\Gamma_b)\right)^{1/2} + \left(\frac{1}{4}.d^2_{MIN}(\Gamma_b)\right)^{1/2}\right)^2 = d^2_{MIN}(\Gamma_b)$$

Assim, temos necessariamente $\hat{\gamma}_b = 0$.

Concluimos, então, que $\hat{l} = 0$, verificando o teorema. [CQD]

Assim, a degradação de desempenho ocasionado pelo algorítmo é devido apenas ao aumento efetivo do número de vizinhos mais próximos, dado pelo teorema a seguir.

Teorema 6.3.2: Para o caso $C_0 \subseteq \ldots \subseteq C_{b-1}$, temos:

$$M^{A}_{0_{eff}}(L) = \sum_{\substack{i=0 \\ d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d^2_{MIN}(L)}}^{b-1} M_0(C_i).M_0(\alpha_i + \Gamma_{i+1})^{d_H(C_i)} + \overline{M}_b \qquad (6.3.1)$$

$$\text{onde: } \overline{M}_b = \begin{cases} m.M_0(\Gamma_b), & d^2_{MIN}(\Gamma_b) = d^2_{MIN}(L) \\ 0, & d^2_{MIN}(\Gamma_b) = d^2_{MIN}(L) \end{cases}$$

Prova: Vê-se claramente da definição do algoritmo que a região de decisão na origem $R_L^A(0)$ é dada por:

$$R_L^A(0) = (V_{\Lambda_0}(0) + \Gamma_1^m) \bigcap \ldots \bigcap (V_{\Lambda_{b-2}}(0) + \Gamma_{b-1}^m) \bigcap V_{\Lambda_{b-1}}(0)$$

de modo que o número de pontos de contato da hiperesfera de raio $\frac{1}{2}.d_{MIN}(L)$ centrada na origem (contida em $R_L^A(0)$) com a fronteira de $R_L^A(0)$ é a soma do número de pontos de contato dessa hiperesfera com a fronteira de cada "região periódica" $V_{\Lambda_i}(0) + \Gamma_{i+1}^m$, $0 \le i \le b-2$ e de $V_{\Lambda_{b-1}}(0)$. Note-se que para o caso $C_0 \subseteq \ldots \subseteq C_{b-1}$, temos que:

$$d_{MIN}^2(L) = \min[d_H(C_0).N(\alpha_0 + \Gamma_1), \ldots, d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b), d_{MIN}^2(\Gamma_b)]$$

O número de pontos de contato na fronteira de $V_{\Lambda_i}(0) + \Gamma_{i+1}^m$ é igual a $M_0(C_i).M_0(\alpha_i + \Gamma_{i+1})^{d_H(C_i)}$ se $d_H(C_i).N(\alpha_i + \Gamma_{i+1}) = d_{MIN}^2(L)$, e igual a zero, caso contrário, para $0 \le i \le b-2$.

O número de pontos de contato na fronteira de $V_{\Lambda_{b-1}}(0)$ é igual a:

(1) $M_0(C_{b-1}).M_0(\alpha_{b-1} + \Gamma_b)^{d_H(C_{b-1})}$, se $d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b) = d_{MIN}^2(L) \neq d_{MIN}^2(\Gamma_b)$;

(2) $m.M_0(\Gamma_b)$, se $d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b) \neq d_{MIN}^2(L) = d_{MIN}^2(\Gamma_b)$;

(3) $M_0(C_{b-1}).M_0(\alpha_{b-1}+\Gamma_b)^{d_H(C_{b-1})}+m.M_0(\Gamma_b)$, se $d_H(C_{b-1}).N(\alpha_{b-1}+\Gamma_b) = d_{MIN}^2(L)=d_{MIN}^2(\Gamma_b)$;

(4) zero, se $d_H(C_{b-1}).N(\alpha_{b-1} + \Gamma_b) \neq d_{MIN}^2(L) \neq d_{MIN}^2(\Gamma_b)$. [CQD]

A complexidade total do algoritmo dependerá da complexidade dos algoritmos que forem utilizados para a decodificação por ML dos reticulados Λ_i, $0 \le i \le b-1$. Se definirmos para cada palavra $\bar{c}_i \in C_i$, $0 \le i \le b-1$, uma métrica $D^{(i)}(\bar{c}_i)$, dado r_i, por:

$$D^{(i)}(\bar{c}_i) = \min_{\gamma_{i+1} \in \Gamma_{i+1}^m} \|r_i - (\bar{c}_i \otimes \alpha_i + \gamma_{i+1})\|^2$$

e pudermos expressá-la como uma métrica aditiva em relação às suas m componentes $\bar{c}_{i,j}$, $1 \le j \le m$, poderemos decodificar Λ_i utilizando um algoritmo de decodificação com decisão suave por ML para C_i, $0 \le i \le b-1$ (por exemplo, aqueles propostos por Forney [14]).

Temos:

$$D^{(i)}(\bar{c}_i) = \underset{\gamma_{i+1} \in \Gamma_{i+1}^m}{\text{MIN}} \sum_{j=1}^{m} \|r_{i,j} - (\bar{c}_{i,j}.\alpha_i + \gamma_{i+1,j})\|^2 =$$

$$= \sum_{j=1}^{m} \left[\frac{1}{2}.\left(\underset{\gamma_{i+1,j}^0 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\gamma_{i+1,j}^0)\|^2 + \underset{\gamma_{i+1,j}^1 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\alpha_i + \gamma_{i+1,j}^1)\|^2 + \right.\right.$$

$$\left.\left. + \frac{1}{2}.(-1)^{\bar{c}_{i,j}}.\left(\underset{\gamma_{i+1,j}^0 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\gamma_{i+1,j}^0)\|^2 - \underset{\gamma_{i+1,j}^1 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\alpha_i + \gamma_{i+1,j}^1)\|^2 \right) \right] \right] =$$

$$= \frac{1}{2}.\sum_{j=1}^{m} s_j^{(i)} + \frac{1}{2}.\sum_{j=1}^{m} (-1)^{\bar{c}_{i,j}}.m_j^{(i)}$$

onde:

$$s_j^{(i)} = d_{j_0}^{(i)} + d_{j_1}^{(i)}$$

$$m_j^{(i)} = d_{j_0}^{(i)} - d_{j_1}^{(i)}$$

e:

$$d_{j_0}^{(i)} = \underset{\gamma_{i+1,j}^0 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\alpha_i + \gamma_{i+1,j}^0)\|^2$$

$$d_{j_1}^{(i)} = \underset{\gamma_{i+1,j}^1 \in \Gamma_{i+1}}{\text{MIN}} \|r_{i,j} - (\alpha_i + \gamma_{i+1,j}^1)\|^2$$

Assim a minimização de $D^{(i)}(\bar{c}_i)$ com $\bar{c}_i \in C_i$ equivale a minimizar:

$$\sum_{j=1}^{m} (-1)^{\bar{c}_{i,j}}.m_j^{(i)}$$

(note-se que o algoritmo de decodificação por estágios para L requer apenas a palavra-código $\bar{c}_i \in C_i$, que minimiza $D^{(i)}(\bar{c}_i)$). Assim podemos utilizar um algoritmo de decodificação com decisão suave por ML para C_i na decodificação de Λ_i, como desejávamos. Esta forma de decodificação pode ser vista como uma generalização daquela proposta por Conway e Sloane para a decodificação da Construção A, citada na seção anterior.

A obtenção das métricas $d_{j_0}^{(i)}$ e $d_{j_1}^{(i)}$ requer a decodificação por ML de cada componente $r_{i,j} \in \mathbb{R}^n$, $1 \le j \le m$, do vetor r_i (dado pelo passo anterior do algoritmo de decodificação por estágios de L) em relação a Γ_{i+1} e $\alpha_i + \Gamma_{i+1}$; isto equivale a decodificação por ML da partição Γ_i/Γ_{i+1} para cada $r_{i,j}$, $1 \le j \le m$. Concluímos então que o número de operações necessárias para a decodificação de Λ_i, $0 \le i \le b-1$, é dado aproximadamente por:

$$N_D(\Lambda_i) = m.[N_D(\Gamma_i/\Gamma_{i+1}) + 1] + N_D(C_i) \tag{6.3.2}$$

onde $N_D(\Gamma_i/\Gamma_{i+1})$ é a complexidade do algoritmo utilizado para decodificar por ML a partição Γ_i/Γ_{i+1}, e $N_D(C_i)$ é a complexidade do algoritmo utilizado para decodificação com decisão suave por ML de C_i. Assim, globalmente, a complexidade do algoritmo de decodificação por estágios para L é dado aproximadamente por:

$$N_D^A(L) = \sum_{i=0}^{b-1} N_D(\Lambda_i) + b.m.n \tag{6.3.3}$$

onde m é a dimensão dos Γ_i's, sendo a última parcela referente aos passos do algoritmo que calcula cada vetor r_i a ser decodificado em relação a Λ_i, no passo subsequente.

Assim, utilizando (6.3.1), (6.3.2) e (6.3.3) podemos determinar o desempenho e a complexidade do algoritmo proposto para os reticulados obtidos na seção 5.3. As Tabelas I-VI apresentam os resultados encontrados; foram utilizados para decodificação das partições Γ_i/Γ_{i+1}, $0 \le i \le b-1$, e dos códigos C_i, $0 \le i \le b-1$, os algoritmos propostos por Forney ([14]).

Tabela I

$L = C_0 \otimes [\mathbf{Z}^2/D_2] + C_1 \otimes [D_2/2\mathbf{Z}^2] + (2\mathbf{Z}^2)^m$

C_0	C_1	L	$\tilde{N}_D(L)$	$\tilde{M}_{eff}^A(L)$	$\tilde{N}_D^A(L)$	$\beta^A(L)$
(8,4,4)	(8,7,2)	H_{16}	31,88	508	14,75	0,16
(16,11,4)	(16,15,2)	X_{32}	75,94	2364	18,88	0,10

TABELA II

$L = C_0 \otimes [\mathbf{Z}^4/D_4] + C_1 \otimes [D_4/R\mathbf{Z}^4] + C_2 \otimes [R\mathbf{Z}^4/RD_4] + C_3 \otimes [RD_4/2\mathbf{Z}^4] + (2\mathbf{Z}^4)^m$

C_0	C_1	C_2	C_3	L	$\tilde{N}_D(L)$	$\tilde{M}^A_{eff}(L)$	$\tilde{N}^A_D(L)$	$\beta^A(L)$
(4,1,4)	(4,3,2)	(4,3,2)	(4,4,1)	H_{16}	31,88	764	21,63	0,56
(8,4,4)	(8,7,2)	(8,7,2)	(8,8,1)	X_{32}	75,94	4156	23,38	0,22

TABELA III

$L = C_0 \otimes [D_4/R\mathbf{Z}^4] + C_1 \otimes [R\mathbf{Z}^4/RD_4] + C_2 \otimes [RD_4/2\mathbf{Z}^4] + C_3 \otimes [2\mathbf{Z}^4/2D_4] + (2D_4)^m$

C_0	C_1	C_2	C_3	L	$\tilde{N}_D(L)$	$\tilde{M}^A_{eff}(L)$	$\tilde{N}^A_D(L)$	$\beta^A(L)$
(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	Λ_{16}	63,88	8956	27,50	0,73
(8,4,4)	(8,4,4)	(8,7,2)	(8,7,2)	Y_{32}	399,94	61500	30,75	0,26

TABELA IV

$L = C_0 \otimes [\mathbf{Z}^8/D_8] + C_1 \otimes [D_8/J_8] + C_2 \otimes [J_8/D_8^\dagger] + C_3 \otimes [D_8^\dagger/E_8] + C_4 \otimes [E_8/RD_8] + C_5 \otimes [RD_8/RJ_8] + C_6 \otimes [RJ_8/RD_8^\dagger] + C_7 \otimes [RD_8^\dagger/2\mathbf{Z}^8] + (2\mathbf{Z}^8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	$\tilde{N}_D(L)$	$\tilde{M}^A_{eff}(L)$	$N^A_d(L)$	$\beta^A(L)$
(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,4,1)	(4,4,1)	(4,4,1)	(4,4,1)	X_{32}	75,94	6172	59,25	1,42

TABELA V

$L = C_0 \otimes [E_8/RD_8] + C_1 \otimes [RD_8/RJ_8] + C_2 \otimes [RJ_8/RD_8^\dagger] + C_3 \otimes [RD_8^\dagger/RE_8] + C_4 \otimes [RE_8/2D_8] + C_5 \otimes [2D_8/2J_8] + C_6 \otimes [2J_8/2D_8^\dagger] + C_7 \otimes [2D_8^\dagger/2E_8] + (2E_8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	$\tilde{N}_D(L)$	$\tilde{M}^A_{eff}(L)$	$\tilde{N}^A_D(L)$	$\beta^A(L)$
(2,0,∞)	(2,1,2)	(2,1,2)	(2,1,2)	(2,2,1)	(2,2,1)	(2,2,1)	(2,2,1)	R.H_{16}	31,88	732	50,13	-0,46
(2,1,2)	(2,1,2)	(2,1,2)	(2,1,2)	(2,2,1)	(2,2,1)	(2,2,1)	(2,2,1)	Λ_{16}	63,88	2788	64,00	-192,42
(4,0,∞)	(4,0,∞)	(4,0,∞)	(4,0,∞)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	2 W_{32}	91,94	2076	64,25	0,64
(4,0,∞)	(4,1,4)	(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	R.H_{32}	791,94	$2{,}2\times10^6$	114,53	0,69
(4,1,4)	(4,1,4)	(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	Λ_{32}	1583,94	$1{,}9\times10^7$	120,75	0,65

TABELA VI

$L = C_0 \otimes [J_8/D_8^\dagger] + C_1 \otimes [D_8^\dagger/E_8] + C_2 \otimes [E_8/RD_8] + C_3 \otimes [RD_8/RJ_8] + C_4 \otimes [RJ_8/RD_8^\dagger] + C_5 \otimes [RD_8^\dagger/RE_8] + C_6 \otimes [RE_8/2D_8] + C_7 \otimes [2D_8/2J_8] + (2J_8)^m$

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	L	$\tilde{N}_D(L)$	$\tilde{M}^A_{eff}(L)$	$\tilde{N}^A_D(L)$	$\beta^A(L)$
(4,1,4)	(4,1,4)	(4,3,2)	(4,3,2)	(4,3,2)	(4,3,2)	(4,4,1)	(4,4,1)	Y_{32}	399,94	77852	106,38	0,54

Os resultados apresentados nessas tabelas são obtidos diretamente pelas fórmulas desenvolvidas nesta seção. Por exemplo, para o reticulado H_{16} da Tabela I, temos:

$$N_D(H_{16}) = 2.N_D(D_8/E_8) + 2.|D_8/E_8| - 1 = 2.120 + 2.8 - 1 = 255$$

$$\tilde{N}_D(H_{16}) = (2/16).255 = 31{,}88$$

$$M^{A}_{0_{eff}}(H_{16}) = M_0(8,4,4).M_0([1,0]^t + 2\mathbb{Z}^2)^4 + M_0(8,7,2).M_0([1,1]^t + 2\mathbb{Z}^2)^2 + 8.M_0(2\mathbb{Z}^2) =$$

$$= 14.4^4 + 28.4^2 + 8.4 = 4.064$$

$$\tilde{M}^{A}_{0_{eff}}(H_{16}) = (2/16).4.064 = 508$$

$$N^{A}_{D}(H_{16}) = [8.(N_D(\mathbb{Z}^2/D_2) + 1) + N_D(8,4,4)] + [8.(\mathrm{NsunD}(D_2/2\mathbb{Z}^2) + 1) + N_D(8,7,2)]$$

$$+ [2.8.2] =$$

$$= [8.(2+1) + 31] + [8.(2+1) + 7) + [2.8.2] = 118$$

$$\tilde{N}^{A}_{D}(H_{16}) = (2/16).118 = 14{,}75$$

Para os dois reticulados novos, W_{32} e Y_{32}, os valores de $\tilde{N}_D(L)$ são aqueles associados aos algorítmos propostos por Forney ([14]), utilizando as construções $Y_{32} = |D_4^2/E_8/RE_8|^4$ e $2W_{32} = |RE_8/2D_8/2E_8|^4$.

Vemos que a grande maioria dos reticulados apresenta uma relação de degradação de desempenho (em dB) por redução de complexidade (em oitavas), $\beta^A(L) \triangleq \Delta\gamma^A_{eff}(L)_{dB}/\log_2(N^A_D(L)/N_D(L))$, inferior ou ao menos próximo a 0.4, que é o valor desta relação observada ao longo dos melhores códigos de classes laterais conhecidos com alto ganho efetivo (Forney e Wei [15]). Observa-se que os melhores resultados ocorrem quando o comprimento m dos códigos binários utilizados não é muito reduzido (em m=2, por exemplo, o algoritmo não apresentou efetividade). Espera-se comportamento semelhante do algorítmo para outros reticulados obtidos pela mesma construção, ou extensões desta que vierem a ser propostas.

Capítulo 7

Conclusões e sugestões para futuras investigações

7.1 Conclusões

Neste trabalho, vários aspectos teóricos e aplicados de reticulados e suas partições foram tratados, visando sua utilização em esquemas de codificação baseados em reticulados para o canal AWGN limitado em banda. Os principais resultados e conclusões podem ser resumidos como se segue.

No Capítulo 2 as caracterizações algébrica e geométrica dos reticulados e suas partições foram introduzidas utilizando conceitos definidos para grupos abelianos arbitrários, indicando que grande parte das técnicas propostas nos capítulos seguintes podem ser extendidas a outros grupos abelianos úteis para a construção de codificadores para o canal AWGN limitado em banda.

No Capítulo 3, a análise algébrica teórica da partição de reticulados, conceituada no Capítulo 2, foi completamente desenvolvida, com a utilização, inclusive, das formas canônicas de Hermite e Smith para matrizes de números inteiros, possibilitando assim a determinação automática por meios computacionais da estrutura algébrica de partições arbitrárias de reticulados, utilizada na construção de códigos lineares de classes laterais (Capítulo 4) e de reticulados (Capítulo 5).

No Capítulo 4 foi desenvolvida uma formulação geral para a avaliação e comparação de codificadores para canais AWGN limitados em banda, com estimativas explícitas para o ganho nominal e a perda pelo número de vizinhos mais próximos, exibindo a necessidade de normalização dos parâmetros de desempenho e complexidade de codificadores específicos. Foi também explicitada a analogia entre reticulados e códigos de classes laterais, na constução de codificadores baseados em reticulados, relacionando seus parâmetros aos destes codificadores.

No Capítulo 5 foi proposta uma construção multicamada de reticulados, baseada em um refinamento de uma partição binária elementar de reticulados e em códigos binários lineares, com expressões explícitas para os parâmetros dos reticulados obtidos. Foram dados alguns exemplos, mostrando novas formas de construção de reticulados conhecidos, bem como a obtenção de novos reticulados interessantes para utilização em codificadores para o canal AWGN limitado em banda.

No Capítulo 6 foi proposto um algoritmo de decodificação sub-ótimo por estágios para os reticulados obtidos pela construção multinível proposta no Capítulo 5. Foi mostrado que o algoritmo implementa decodificação com distância limitada, com raio efetivo de correção igual à metade da distância mínima do reticulado, de modo que a degradação de desempenho sofrida na utilização do algoritmo é devida exclusivamente ao decorrente aumento no coeficiente efetivo de erro do reticulado, para o qual uma expressão explícita foi obtida. Foi mostrado que cada estágio da decodificação poderia ser realizado utilizando um algoritmo de decodificação com decisão suave ótimo do código binário linear correspondente, com uma adequada definição da métrica de cada componente binária, obtendo-se então uma estimativa para a redução de complexidade auferida pelo algoritmo de decodificação por estágios proposto. Verificou-se que, para a maioria dos reticulados obtidos no Capítulo 5, a relação degradação de desempenho (em dB) por redução de complexidade (em oitavas) ficou próxima daquela observada

(≈ 0.4 dB/oitava) ao longo da seqüência dos melhores codificadores conhecidos para o canal AWGN limitado em banda utilizando algoritmos de decodificação ótima, evidenciando a efetiva vantagem do algoritmo sub-ótimo em relação ao algoritmo ótimo mais eficiente conhecido para aqueles reticulados.

7.2 Sugestões para futuras investigações

Como direções alternativas para posterior prosseguimento deste trabalho, podemos sugerir algumas extensões e aplicações dos resultados obtidos como se segue.

Uma constelação (bidimensional) de sinais m-PSK pode ser representada pelo conjunto de m pontos do plano complexo dados por $\mathrm{Exp}[2\pi mj/M]$, m=0,...,M-1, formando um grupo abeliano sob multiplicação complexa, com $z_0=1$ como elemento identidade. Assim, os conceitos e a notação introduzidos no Capítulo 2 para grupos abelianos genéricos se aplicam, bem como as várias técnicas dos capítulos posteriores; em particular, técnicas de construção e decodificação por estágios fornecem novas alternativas de construção e decodificação de codificadores de energia constante (códigos esféricos) para o canal AWGN limitado em banda.

A estrutura de uma dada partição de reticulados e as normas mínimas das suas classes laterais (obtidas pela decodificação da origem em relação a cada uma delas) são os ingredientes necessários e suficientes para a busca dos denominados "códigos convolucionais generalizados", idealizados por Calderbank e Sloane ([05]) para a construção de codificadores em treliça para o canal AWGN limitado em banda; a sistemática desenvolvida no Capítulo 3 torna computacionalmente mais simples a busca automática desses códigos, em especial para partições de reticulados em altas dimensões.

A complexidade de implementação do seletor de pontos (e do correspondente mapeamento inverso na recepção, após a decodificação), em um codificador reticulado ou baseado em um código de classes laterais, não foi considerada na avaliação da complexidade do codificador; como foi mencionado no Capítulo 4, para ganhos da forma relativamente altos (constelações bastante esféricas), estas complexidades passam a ser frações consideráveis da complexidade total, de modo que seria interessante buscar compromissos compensadores para esses objetivos conflitantes.

As expressões para a distância quadrática mínima e o número de vizinhos mais próximos, dadas no Capítulo 5, para a construção multicamada de reticulados proposta naquele capítulo, são válidas para o caso em que os códigos binários utilizados são encaixados; fora deste caso, as expressões dadas são, em geral, apenas limitantes (inferior para a distância e superior para o número de vizinhos), de modo que tem-se a chance de melhorar estes parâmetros, com uma escolha judiciosa dos representantes de classes laterais utilizados na construção; infelizmente, não existem ainda procedimentos gerais para essa escolha, devendo ser realizados estudos nesse sentido.

Finalmente, observe-se que o algoritmo de decodificação por estágios, proposto no Capítulo 6, pode ser ainda mais simplificado, utilizando algoritmos sub-ótimos com distância limitada para a decodificação das partições Γ_{i+1}/Γ_i e dos códigos binários C_i às custas,

naturalmente, de uma degradação adicional de desempenho devido ao crescimento do número de vizinhos que então resultaria; estudos sobre as possíveis compensações resultantes dessas alternativas seriam úteis e elucidativas.

Referências bibliográficas

[01] E. S. BARNES e N. J. A. SLOANE, "New lattice packings of spheres", *Can. J. Math.*, vol. 35, pp. 117-130, 1983

[02] Y. BE'ERY, B. SHAHAR e J. SNYDERS, "Fast decoding of the Leech lattice", *IEEE J. Select. Areas Commun.*, vol. SAC-7, pp. 959-967, 1989.

[03] A. BOS, J. H. CONWAY e N. J. A. SLOANE, "Further lattice packings in hight dimensions", *Mathematika*, vol. 29, pp. 171-180, 1982.

[04] A. R. CALDERBANK, "Multilevel codes and multistage decoding", *IEEE Trans. Inform. Theory*, vol IT-37, pp. 222-229, 1989.

[05] A. R. CALDERBANK e N. J. A. SLOANE, "New trellis codes based on lattices and cosets", *IEEE Trans. Inform. Theory*, vol IT-33, pp. 177-195, 1987.

[06] J. W. S. CASSELS, *An Introduction to the Geometry of Numbers*, Springer-Verlag, 1971.

[07] G. C. CLARK JR. e J. B. CAIN, *Error-Correction Coding for Digital Communication*, Plenum, 1981.

[08] J. H. CONWAY e N. J. A. SLOANE, "Fast quantizing and decoding algorithms for lattice quantizers and codes", *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1152-1187, 1988.

[09] J. H. CONWAY e N. J. A. SLOANE, "Soft decoding techniques for codes and lattices, including the Golay code and the Leech lattice", *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 41-50, 1986.

[10] J. H. CONWAY e N. J. A. SLOANE, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 1988.

[11] E. L. CUSAK, "Error control codes for QAM signalling", *Eletronics Letterss*, vol. 20, pp. 62-63, 1984.

[12] P. DELSARTE, "Four fundamental parameters of a code and their combinatorial significance", *Information and Control*, vol. 23, pp. 407-438, 1973.

[13] G. D. FORNEY, "Coset codes - Part I: Introduction and geometrical classification", *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1123-1151, 1988.

[14] G. D. FORNEY, "Coset codes - Part II: Binary lattices and related codes", *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1152-1187, 1988.

[15] G. D. FORNEY e L. -F. WEI, "Multidimensional constelations - Part I: Introduction, figures of merit, and generalized cross constelations", *IEEE J. Select. Areas Commun.*, vol. SAC-7, pp. 877-892, 1989.

[16] G. D. FORNEY, "Multidimensional constelations - Part II: Voronoi constellations", *IEEE J. Select. Areas Commun.*, vol. SAC-7, pp. 941-958, 1989.

[17] G. D. FORNEY, "A bounded-distance decoding algorithm for the Leech lattice, with generalizations", *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 906-909, 1989

[18] G. D. FORNEY, R. G. GALLAGER, G. R. LANG, F. M. LONGSTAFF e S. U. QURESHI, "Efficient modulation for band-limited channels", *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 632-647, 1984.

[19] J. B. FRALEIGH, *A First Course in Abstract Algebra*, Addison-Wesley, 1982

[20] R. G. GALLAGER, *Information Theory and Reliable Communication*, John Wiley and Sons, 1968.

[21] G. R. LANG e F. M. LONGSTAFF, "A Leech lattice modem", *J. Select. Areas Commun.*, vol. SAC-7, pp. 968-975, 1989.

[22] J. LEECH, "Notes on sphere packings", *Can. J. Math.*, vol. 13, pp. 251-267, 1967.

[23] J. LEECH e N. J. A. SLOANE, "Sphere packings and error-correcting codes", *Can. J. Math.*, vol. 23, pp. 718-745, 1971.

[24] S. LIN e D. J. COSTELO JR., *Error Control Coding*, Prentice-Hall, 1983.

[25] F. J. MACWILLIAMS e N. J. A. SLOANE, *The Theory of Error-Correction Codes*, North Holland, 1978.

[26] M. NEWMAN, *Integral Matrices*, Academic, 1972.

[27] C. A. ROGERS, *Packing and Covering*, Cambridge Univ. Press, 1964.

[28] S. I. SAYEGH, "A class of optimum block codes in signal space", *IEEE Trans. Commun.*, vol. COM-34, pp. 1043-1045, 1986.

[29] C. E. SHANNON e W. WEAVER, *A Mathematical Theory of Communication*, Univ. Illinois Press, 1949.

[30] R. A. SILVERMAN e M. BALSER, "Coding for constant-data-rate systems", *IRE trans. Inform. Theory*, vol. PGIT-4, pp. 50-63, 1954.

[31] C. C. SIMS, *Abstract Algebra, A Computational Approach*, Jonh Wiley and Sons, 1984.

[32] R. M. TANNER, "Algebraic construction of large Euclidean distance combined coding/modulation systems", Univ. Calif., Santa Cruz, Rel. Tec. UCSC-CRL-87-7, junho 1987.

[33] G. UNGERBOECK, "Channel coding with multilevel-phase signals", *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55-67, 1982.

[34] T. Verhoeff, "An updated table of minimum-distance bounds for binary linear codes", *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 665-680, 1987.

[35] A. J. VITERBI e J. K. OMURA, *Principles of Digital Communication and Coding*, McGraw-Hill, 1979.

[36] L. -F. WEI, "Trellis-coded modulation with multidimensional constellations", *IEEE Trans. Inform. Theory*, vol IT-33, pp. 483-501, 1987.

[37] J. K. WOLF, "Efficient maximum likelihood decoding of linear block codes using a trellis", *IEEE Trans. Inform. Theory*, vol IT-24, pp. 76-80, 1978.

[38] J. M. WOZENCRAFT e I. M. JACOBS, *Principles of Communication Engineering*, John Wiley and Sons, 1965.