

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
DEPARTAMENTO DE TELEMÁTICA

**CONSTELAÇÕES DE SINAIS E ANÁLISE DE
DESEMPENHO NO PLANO HIPERBÓLICO**

Eduardo Brandani da Silva

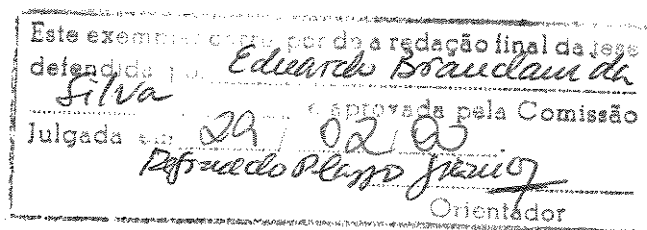
Orientador: **Prof. Dr. Reginaldo Palazzo Jr.**
Co-orientadores: **Prof. Dr. Marcelo Firer**
Prof. Dr. Sueli I. R. Costa

Banca examinadora: **Prof. Dr. Reginaldo Palazzo Jr.**
Prof. Dr. Dalton Soares Arantes
Prof. Dr. Jaime Portugheis
Prof. Dr. Dicesar Lass Fernandez
Prof. Dr. Sergio Antônio Tozoni

Tese apresentada à Faculdade de
Engenharia Elétrica e de Computação,
FEEC-UNICAMP, como requisito
parcial para obtenção do título de
DOUTOR EM ENGENHARIA
ELÉTRICA.

Fevereiro-2000

Campinas S.P.



2000191036

UNIDADE 30
N.º CHAMADA:
T/Unicom
Si38c
V. _____ Ex. _____
TOMBO BC/ 43203
PROC. 16-278100
C D
PREC@ R\$ 41,00
DATA 15/12/00
N.º CPD _____



CM-00154345-6

Si38c

Silva, Eduardo Brandani da
Constelações de sinais e análise de desempenho no
plano hiperbólico / Eduardo Brandani da Silva.--
Campinas, SP: [s.n.], 2000.

Orientador: Reginaldo Palazzo Jr.
Co-orientadores: Marcelo Firer, Sueli I. R. Costa

Tese (doutorado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Teoria dos sinais (Telecomunicações). 2.
Modulação (Eletrônica). 3. Geometria hiperbólica. 4.
Desempenho - Sistemas de comunicação. I. Palazzo Jr.,
Reginaldo. II. Firer, Marcelo. III. Costa, Sueli I. R. IV.
Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. V. Título.

AGRADEÇO

Ao professor Reginaldo Palazzo Jr., por ter acreditado em mim, dando todo apoio que estive ao seu alcance para a realização e divulgação deste trabalho.

Ao professor Marcelo Firer, por seu valioso auxílio, sempre encontrando tempo para discutirmos este trabalho.

À professora Sueli Costa, por seu auxílio e valiosas sugestões.

Aos professores Dalton Soares Arantes, Jaime Portugheis, Dicesar Lass Fernandez e Sergio Antônio Tozoni pela participação como membros da banca examinadora desta tese e por sugestões.

Aos meus amigos João Gerônimo e Henrique Lazari, pelas proveitosas discussões.

Ao meu amigo “Marquinhos”, pelas conversas e companherismo.

À Claudete, que revisou cuidadosamente o texto.

Aos funcionários da Faculdade de Engenharia Elétrica e de Computação, que de alguma forma contribuíram para a realização deste trabalho.

À Capes pela bolsa e auxílios concedidos.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

À minha esposa Claudete, que esteve ao meu lado em todos os momentos, bons ou ruins, e que através de seu amor e dedicação me deram toda a inspiração e força para concluir este trabalho.

Ao meu pequeno Eduardo, que com sua beleza, graça e inocência deu um novo sentido à minha vida.

Aos meus pais, que sempre me apoiaram e que através de muito esforço, possibilitaram meus estudos. Sem eles nada aqui seria realidade.

A Emerson e Camila que me receberam carinhosamente e que passaram a fazer parte de minha vida.

Em memória de minha querida avó Amélia, que deixou tanta saudade em mim.

Dedico

Resumo

Neste trabalho construímos constelações de sinais no plano hiperbólico. Analisamos o desempenho de constelações PAM, PSK e QAM-circular no plano hiperbólico em relação às constelações equivalentes do plano euclidiano. Para estabelecermos estas constelações introduzimos diversos conceitos de geometria hiperbólica, sendo o principal deles, o conceito de tesselação do plano. Para podermos fazer decisões em relação à escolha de quais tesselações fornecem constelações de interesse, obtivemos funções enumeradoras, que nos permitem contar o número de pontos em subconjuntos finitos das tesselações. Para podermos calcular o desempenho de constelações de interesse, obtivemos uma função densidade de probabilidade gaussiana para o plano hiperbólico e apresentamos suas principais propriedades. Partindo do conceito de função de probabilidade gaussiana hiperbólica, caracterizamos o ruído de um canal gaussiano hiperbólico, utilizando as isometrias do plano hiperbólico.

Abstract

In this work we design signal constellations in the hyperbolic plane. We analyse the performance of PAM, PSK and QAM-circular constellations in the hyperbolic plane in connection with similar constellations in the Euclidean plane. To set up these constellations, we introduce several concepts of hyperbolic geometry, among them is the concept of tessellations in the plane. To make decisions with respect to the choice of tessellations providing constellations of interest, we develop enumerating functions, in order to count the number of points in finite subsets of tessellations. To calculate the performance of signal constellations, we derive a gaussian probability density function in the hyperbolic plane and show its main properties. From the concept of hyperbolic gaussian probability density function, we characterize the hyperbolic gaussian channel noise, using isometries of the hyperbolic plane.

Lista de Figuras

2.1	Dados pontos z e w , através de isometrias do plano hiperbólico, movimentasse os dois pontos sobre o eixo y	16
2.2	Construção da reta perpendicular a uma reta dada r_h , através de um ponto dado de r_h , em H^2	26
2.3	Construção da reta perpendicular a uma reta dada, através de um ponto, em Δ	27
3.1	Eixos coordenados para o caso hiperbólico.	46
4.1	Determinação das possíveis tesselações nas geometrias consideradas.	57
4.2	Exemplo de tesselação $\{8, 4\}$ construída no plano hiperbólico.	59
4.3	Em cada vértice do polígono inicial são inseridas $q - 2$ novas arestas.	61
4.4	Em cada um dos $q - 3$ pares, originados das arestas inseridas em v_i , inserimos $p - 3$ pontos	62
4.5	Entre dois vértices consecutivos v_i e v_{i+1} inserimos $p - 4$ pontos.	63
4.6	Em cada par de arestas inseridas em pontos do tipo 2 de L_1 inserimos novos $p - 3$ pontos.	64
4.7	Primeiros três níveis da tesselação $\{4, 4\}$ no plano hiperbólico.	67

4.8	Em cada vértice de L_0 inserimos $q - 2$ arestas. Entre dois vértices consecutivos não são inseridos novos pontos.	68
4.9	Tesselação $\{3, 12\}$	70
4.10	Em um ponto do tipo 3 não são inseridas novas arestas, mas a cada um desses pontos está relacionado um conjunto de $p - 5$ pontos do próximo nível.	71
4.11	Tesselação $\{12, 3\}$	75
4.12	Alguns exemplos da região inicial de subconjuntos de tesselações, originados pelo método das arestas.	76
4.13	Um artifício para poder contar os pontos do tipo 2 e do tipo 3 neste caso.	81
5.1	Constelação 2-PAM em H	85
5.2	Região de Voronoi para um sinal z_i transmitido.	96
5.3	Regiões de integração para obtenção das probabilidades de acerto, para a constelação 16-QAM circular hiperbólico.	106

Índice

1	Introdução	5
2	Introdução à Geometria Hiperbólica	9
2.1	Revisão de Geometria Riemanniana	10
2.2	Geometria Hiperbólica	12
2.2.1	Modelo do semi-plano superior	12
2.2.2	Modelo do disco aberto unitário	17
2.2.3	Propriedades Geométricas em \mathbb{H}^2 e em Δ	19
3	Determinação da Densidade de Probabilidade Gaussiana Hiperbólica	33
3.1	Densidade de Probabilidade Gaussiana Hiperbólica em \mathbb{H}	34
3.2	Densidade de Probabilidade Gaussiana Hiperbólica em \mathbb{H}^2	37
3.3	As Probabilidades Marginais	43
3.4	Modelo do Canal Gaussiano no Espaço Hiperbólico.	51
4	Tesselações dos Planos Euclidiano e Hiperbólico e as Correspon-	

dentes Funções Enumeradoras	55
4.1 Tesselações nos Planos Euclidiano e Hiperbólico	56
4.2 Subconjuntos Originados por um Polígono Inicial	59
4.2.1 CASO 1: $p \geq 4$ e $q \geq 4$	61
4.2.2 CASO 2: $p = 3$	67
4.2.3 CASO 3: $p \geq 6, q = 3$	70
4.3 Subconjuntos Originados por q Arestas	75
4.3.1 CASO 1: $p \geq 4, q \geq 4$	76
4.3.2 CASO 2: $p = 3$	78
4.3.3 CASO 3: $p \geq 6, q = 3$	80
5 Receptor de Máxima Verossimilhança e Análise de Desempenho	83
5.1 Constelações de Sinais e Receptor PAM Hiperbólico	84
5.2 Receptor de Máxima Verossimilhança	90
5.3 Constelações PSK	92
5.4 QAM Circular Hiperbólico	101
5.5 Outras Constelações	110
6 Conclusões	115
6.1 Desenvolvimento	115
6.2 Trabalhos Futuros	117
7 Referências	119

Capítulo 1

Introdução

Um dos principais objetivos na proposta de novos sistemas de comunicações, é que estes sistemas apresentem ganhos de codificação, e menor complexidade quando comparados com os sistemas já conhecidos.

Em qualquer sistema de comunicações, a informação sendo transmitida sempre estará sujeita a um conjunto de perturbações provenientes do canal de transmissão. Esse conjunto é denominado ruído. Apesar do ruído ter origens físicas, o seu tratamento depende de um modelo probabilístico. Assim, caracterizar o processo estocástico é equivalente a especificar a função distribuição ou a função densidade de probabilidade. Este é um passo importante em direção a obtenção do receptor ótimo, e conseqüentemente a realização do sistema. Uma vez caracterizado o processo estocástico via a função densidade de probabilidade, o passo seguinte é o processamento do sinal a ser utilizado na transmissão, de tal forma que a ação do ruído seja melhor controlada. Isto pode ser realizado de várias formas, como por exemplo: via modulação, codificação, etc.

Motivações e objetivos

Tradicionalmente em comunicações digitais utiliza-se de espaços vetoriais na descrição das constelações de sinais, bem como de uma métrica associada. Em geral, o espaço de sinais apresenta uma estrutura geométrica euclidiana, independente da métrica em consideração.

Neste ponto vem naturalmente a pergunta: Por que estudar constelações de sinais em um espaço hiperbólico? Uma das principais razões é que no espaço hiperbólico é possível construir arranjos regulares de pontos que não são possíveis no espaço euclidiano. Com relação a este ponto, lembramos que os únicos reticulados totalmente regulares do plano euclidiano são aqueles formados somente por triângulos equiláteros, quadrados ou hexágonos. No espaço hiperbólico infinitas são as possibilidades de reticulados regulares (tesselações) as quais são chamadas de tesselações hiperbólicas. Desse modo, o projetista de um sistema de comunicações terá a seu dispor uma infinidade de tesselações hiperbólicas, de onde este poderá selecionar a constelação de sinais mais apropriada para a aplicação que está sendo considerada. Uma outra razão é, a de que ganhos de codificação são obtidos quando se utiliza constelações de sinais hiperbólicos, em comparação com as constelações de sinais no espaço euclidiano. Além disso, o processamento necessário na demodulação na geometria hiperbólica é similar ao da geometria euclidiana.

Um dos objetivos desta pesquisa é realizar uma análise comparativa do desempenho de um sistema de comunicações, cuja constelação de sinais esteja contida no espaço euclidiano, em relação a um sistema de comunicações, cuja constelação de

sinais seja uma "equivalente", contida no espaço hiperbólico. Mostramos que as constelações de sinais do tipo *M-PAM* hiperbólico, apresentam um ganho de codificação de pelo menos 3.61 dB, quando comparadas com as correspondentes constelações *M-PAM* euclidianas. Este comportamento se estende às constelações de sinais do tipo *M-PSK* e *M-QAM*, porém com valores de ganhos diferentes.

Desenvolvimento.

A tese consiste de seis capítulos. No Capítulo 2 introduzimos as noções básicas da geometria hiperbólica, tratamos de alguns modelos para esta geometria, suas propriedades bem como o conceito de geodésica. Aqui gostaríamos de chamar a atenção de que a geometria hiperbólica não é obtida da euclidiana, e que a mesma apresenta propriedades que a tornam bastante diferente da euclidiana. Duas das principais diferenças são: (i) o quinto axioma de Euclides não é mais satisfeito (Axioma das Paralelas), assim dada uma "reta" e um ponto fora desta, existem mais de uma reta paralela à reta dada; (ii) a soma dos ângulos interno de triângulo é menor que π .

Como colocado nos objetivos acima, queremos fazer comparações de desempenho entre constelações euclidianas e hiperbólicas, e essas comparações são feitas através da probabilidade de erro. Assim, no caso hiperbólico necessitamos de uma medida de probabilidade que permita realizar essas comparações. Com isso, no Capítulo 3 apresentamos o desenvolvimento da função densidade gaussiana hiperbólica, como um passo intermediário no estabelecimento de uma medida de probabilidade para o plano hiperbólico. Neste capítulo mostramos a sua construção e suas principais propriedades. Ainda neste capítulo, apresentamos o modelo do canal gaussiano

hiperbólico "aditivo" e também desenvolvemos a idéia de ruído gaussiano hiperbólico, como uma modelagem para a transmissão de um sinal hiperbólico.

Para realizarmos as comparações entre as constelações de sinais euclidianas e hiperbólicas, certas restrições serão impostas. Dentre estas restrições, encontramos aquela em que o número de sinais devem ser iguais, bem como a energia média. Assim, dada uma constelação euclidiana que provenha de um reticulado (tesselação), precisamos saber como escolher dentre as hiperbólicas, subconjuntos finitos das tesselações que tenham o mesmo número de sinais que a euclidiana. No Capítulo 4 introduzimos formalmente o conceito de tesselação regular dos planos euclidiano e hiperbólico. Desenvolvemos também as correspondentes funções enumeradoras, com o objetivo de permitir a realização da contagem do número de sinais no projeto de construção de constelações de sinais.

No Capítulo 5 apresentamos o desenvolvimento do receptor de máxima verossimilhança a ser utilizado. Além disso, realizamos a análise de desempenho das constelações PAM, PSK e QAM euclidianas com equivalentes hiperbólicas. Chamamos a atenção para o fato de que as constelações PAM e PSK podem ser definidas no plano hiperbólico, exatamente da mesma maneira em que estas são definidas no plano euclidiano.

Capítulo 2

Introdução à Geometria

Hiperbólica

Neste capítulo introduzimos os conceitos de geometria hiperbólica, visando as definições e resultados que são necessários para os propósitos desta tese. Todavia, referimos ao leitor alguns excelentes textos neste assunto, [1], [2] e [3]. A geometria hiperbólica pode ser introduzida de diferentes formas. A forma clássica é a axiomática, onde alguns axiomas são fixados e através destes as propriedades geométricas do espaço são deduzidas, como é comumente feito na geometria euclidiana. Sob o ponto de vista axiomático, a diferença para com o caso euclidiano é que o axioma das paralelas não é satisfeito, isto é, "dada uma reta e um ponto fora desta, existe mais de uma reta contendo o ponto e paralela à reta dada". A forma mais geral utiliza os conceitos de geometria Riemanniana.

Iremos dar preferência à forma via os conceitos de geometria Riemanniana, pois a introdução dos modelos a serem utilizados bem como a métrica hiperbólica, podem

ser colocados de forma bastante concisa. Devemos observar que existe uma outra forma, que é a apresentada em [1], onde a construção da geometria hiperbólica ocorre através de funções de Moebius e suas propriedades.

2.1 Revisão de Geometria Riemanniana

Iniciamos com o estabelecimento de algumas definições. Sem perda de generalidade, iremos considerar o espaço euclidiano bidimensional, isto é, o \mathbb{R}^2 . Todavia, podemos estender estas definições para o caso do espaço euclidiano n -dimensional, \mathbb{R}^n . Lembremos que uma matriz quadrada M é chamada positiva definida se $uMu^t \geq 0$ para todo vetor u e $uMu^t = 0$ se, e somente se $u = 0$.

Definição 2.1 *Seja Ω um subconjunto aberto e denso em \mathbb{R}^2 . Uma métrica Riemanniana em Ω é uma lei que faz corresponder a cada ponto p de Ω um produto interno (ou seja, uma forma bilinear simétrica, positiva definida), denotado por $\langle \cdot, \cdot \rangle_p$.*

Definição 2.2 *Sejam $g_{11} = g_{11}(x, y)$, $g_{12} = g_{12}(x, y)$, $g_{21} = g_{21}(x, y)$ e $g_{22} = g_{22}(x, y)$ funções diferenciáveis de \mathbb{R}^2 em \mathbb{R} . Uma métrica Riemanniana em \mathbb{R}^2 é dada pela matriz*

$$G(x, y) = \begin{bmatrix} g_{11}(x, y) & g_{12}(x, y) \\ g_{21}(x, y) & g_{22}(x, y) \end{bmatrix} \equiv \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

se $G(x, y)$ for uma matriz positiva definida para qualquer ponto fixado $(x, y) \in \Omega \subset \mathbb{R}^2$. Sejam $U(x, y)$ e $V(x, y)$ dois campos de vetores quaisquer. Definimos o produto interno de $u = U(x, y)$ e $v = V(x, y)$ relativo ao ponto $P = (x, y)$, através da matriz

G por

$$\langle U(P), V(P) \rangle = \langle u, v \rangle_P = uG(P)v^t.$$

Assim,

$$\|V(P)\|_P = \langle v, v \rangle^{1/2} = [vG(P)v^t]^{1/2}$$

é a norma de $V(P)$.

Deve ser observado que essa é uma definição local, ou seja a matriz G pode assumir diferentes valores conforme os movimentos em \mathbb{R}^2 . Através desta definição, apresentamos a seguir uma definição mais geral de comprimento de arco de uma curva.

Definição 2.3 *Seja $\alpha(t) = (x(t), y(t))$, $t \in (a, b)$, uma curva parametrizada em \mathbb{R}^2 diferenciável por partes. Denotando por C a curva determinada por $\alpha(t)$, o comprimento de arco de C , em relação a norma dada por $G(x, y)$, é definido por*

$$\begin{aligned} \|\alpha\|_G &= \int_C ds = \int_a^b \|\alpha'(t)\|_{G(\alpha(t))} dt \\ &= \int_a^b \|(x'(t), y'(t))\|_{G(x(t), y(t))} dt \\ &= \int_a^b \sqrt{g_{11}\left(\frac{dx}{dt}\right)^2 + \frac{dx}{dt}\frac{dy}{dt}(g_{12} + g_{21}) + g_{22}\left(\frac{dy}{dt}\right)^2} dt \\ &= \int_a^b \sqrt{g_{11}dx^2 + dx dy(g_{12} + g_{21}) + g_{22}dy^2}. \end{aligned}$$

onde $\alpha'(t) = \frac{d}{dt}\alpha(t)$.

Assim, dados dois pontos quaisquer P e Q em \mathbb{R}^2 , a função distância entre eles é

definida por

$$d(P, Q) = \inf_{\gamma} \int_{\gamma} |\gamma'|_G dt$$

onde o ínfimo é tomado sobre todas as curvas γ diferenciáveis por partes conectando P a Q e o símbolo $|\cdot|$ denota sempre a norma euclidiana.

Exemplo. Se $G(x, y) \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ para todo $x, y \in \mathbb{R}$, então $\|V(P)\|_p$, a métrica

resultante, é a métrica euclidiana. Mais ainda, tomando $G \equiv \begin{bmatrix} k^2 & 0 \\ 0 & k^2 \end{bmatrix}$, com $k \in \mathbb{R}$, obtemos ainda uma métrica euclidiana.

2.2 Geometria Hiperbólica

Existem vários modelos para a geometria hiperbólica. Os mais clássicos são: o modelo do semi-plano superior, o modelo do disco aberto unitário, o modelo de Klein e o modelo do hiperbolóide. A seguir apresentaremos com detalhes o modelo do semi-plano superior e o modelo do disco aberto unitário para a geometria hiperbólica. O motivo de trabalharmos somente com estes dois modelos está no fato de que os cálculos de área e volume, que aparecerão nos próximos capítulos são mais simples nestes dois modelos.

2.2.1 Modelo do semi-plano superior

Seja $\mathbb{H}^2 = \{z = x + iy \in \mathbb{R}^2 : y > 0\}$ o semi-plano superior. Através das Definições 2.2 e 2.3 iremos introduzir uma métrica neste conjunto. Considere a matriz $G(x, y)$

dada por

$$G(x, y) = \begin{bmatrix} 1/y^2 & 0 \\ 0 & 1/y^2 \end{bmatrix}, \quad y > 0.$$

Primeiramente, iremos obter uma métrica na semi-reta $\mathbb{H} = \{y \in \mathbb{R} : y > 0\}$. Podemos indentificar \mathbb{H} com a curva $\alpha(t) = (x_0, t)$, $t > 0$ e x_0 fixo, logo $\alpha(t)$ está contida em \mathbb{H}^2 . Sejam $z = (x_0, y_0)$ e $w = (x_0, y_1)$ dois pontos arbitrários em $\alpha(t)$ tal que $\alpha(t)$ esteja restrita ao intervalo $0 < y_1 \leq t \leq y_0$. Então,

$$d(z, w) = \int_{y_1}^{y_0} \|\alpha'\|_G dt,$$

$$G(\alpha(t)) = G((x_0, t)) = \begin{bmatrix} 1/t^2 & 0 \\ 0 & 1/t^2 \end{bmatrix}.$$

Assim,

$$\begin{aligned} d(z, w) &= \int_{y_1}^{y_0} \langle (0, 1), (0, 1) \rangle_{G(x(t), y(t))}^{1/2} dt = \int_{y_1}^{y_0} \left(\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1/t^2 & 0 \\ 0 & 1/t^2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)^{1/2} dt \\ &= \int_{y_1}^{y_0} (1/t) dt = \ln(y_0/y_1). \end{aligned}$$

Se $y_0 < y_1$ então o sinal da integral acima é negativo. Com isso,

$$d(z, w) = |\ln(y_0/y_1)|. \quad (2.1)$$

é uma métrica em \mathbb{H} . Agora, em relação a \mathbb{H}^2 resulta que $d(z, w) \leq |\ln(y_0/y_1)|$, pois temos outras curvas em \mathbb{H}^2 ligando z a w . Seja $\gamma(t) = (x(t), y(t))$, $a < t < b$, uma curva qualquer fixada ligando z a w e consideremos também a curva $\beta(t) = (x_0, t + y_1)$ com $0 \leq t \leq y_0 - y_1$. Podemos assumir que a função $y(t)$ é da forma $g(t) = t + y_1$, pois

seja $\psi : [0, y_0 - y_1] \rightarrow [a, b]$ a bijeção natural entre os dois intervalos, então $g = y \circ \psi$,

logo $y = g \circ \psi^{-1}$. Assim,

$$\begin{aligned} d(z, w) &= \int_a^b \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)} dt \geq \int_a^b \frac{\sqrt{y'(t)^2}}{y(t)} dt \\ &= \int_0^{y_1 - y_0} \frac{1}{t + y_0} dt = \ln(y_0/y_1). \end{aligned}$$

Portanto, determinamos a distância entre dois pontos z e w em \mathbb{H}^2 quando $\operatorname{Re}(z) = \operatorname{Re}(w)$.

A semi-reta \mathbb{H} junto com a métrica dada em (2.1) é chamada de *reta hiperbólica*.

Agora de maneira geral, seja $\gamma(t) = (x(t), y(t))$, $a < t < b$, uma curva arbitrária C^1 (diferenciável). Utilizando a matriz

$$G(x, y) = \begin{bmatrix} 1/y(t)^2 & 0 \\ 0 & 1/y(t)^2 \end{bmatrix},$$

o comprimento de $\gamma(t)$ é dado por

$$\begin{aligned} \|\gamma\| &= \int_a^b \|\gamma'(t)\|_G dt \\ &= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 \frac{1}{y(t)^2} + \left(\frac{dy}{dt}\right)^2 \frac{1}{y(t)^2}} dt \\ &= \int_a^b \frac{1}{y(t)} \sqrt{dx^2 + dy^2}. \end{aligned}$$

A matriz $G(x, y)$ dá origem ao diferencial $ds = |dz|/|\operatorname{Im}[z]|$, onde $|dz| = \sqrt{dx^2 + dy^2}$.

Desse modo, dada uma curva $\gamma : [a, b] \rightarrow \mathbb{H}^2$, seu comprimento de arco é dado por

$$\|\gamma\| = \int_a^b \frac{|\gamma'|}{\text{Im}[\gamma(t)]} dt.$$

Portanto, dados z e w em \mathbb{H}^2 definimos a distância entre eles por

$$d(z, w) = \inf_{\gamma} \|\gamma\|, \quad (2.2)$$

onde o ínfimo é tomado sobre todas as curvas conectando z a w em \mathbb{H}^2 . Para obtermos uma expressão da distância em termos de funções elementares (funções simples, tais como as trigonométricas, polinomiais, exponenciais, etc), vamos estudar as isometrias de \mathbb{H}^2 .

Definição 2.4 *Seja (X, d) um espaço métrico. Uma isometria de X é uma função $f : X \rightarrow X$ que preserva distância, ou seja,*

$$d(f(x), f(y)) = d(x, y)$$

para todo $x, y \in X$.

Consideremos as aplicações $g : \mathbb{H}^2 \rightarrow \mathbb{H}^2$ da forma

$$g(z) = \frac{az + b}{cz + d}, \quad (2.3)$$

onde a, b, c e d são números reais tais que $ad - bc = 1$. Podemos verificar que

$$\frac{|g'(z)|}{\text{Im}[g(z)]} = \frac{1}{\text{Im}[z]}.$$

Isto permite concluir que $\|g(\gamma)\| = \|\gamma\|$, ou seja, o comprimento de uma curva em \mathbb{H}^2 é invariante pela aplicação de g . Devido a esta invariância, resulta que

$$d(g(z), g(w)) = d(z, w).$$

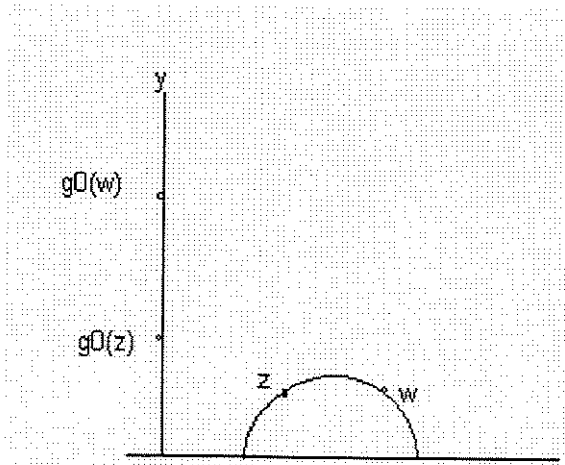


Figura 2.1: Dados pontos z e w , através de isometrias do plano hiperbólico, movimenta-se os dois pontos sobre o eixo y .

Portanto, as funções g são isometrias no espaço métrico (\mathbb{H}^2, d) . Conhecendo como são as isometrias em \mathbb{H}^2 , podemos determinar a expressão geral para a distância entre dois pontos de \mathbb{H}^2 .

Dados dois pontos quaisquer fixados z e w de \mathbb{H}^2 , seja g_0 a isometria que aplica os pontos z e w sobre o eixo euclidiano y . Esse eixo pode ser identificado com \mathbb{H} , onde já temos a função distância determinada. A isometria g_0 é uma composição de três outras isometrias g_1 , g_2 e g_3 que agem da seguinte forma: g_1 aplica z sobre o eixo y , g_2 é uma isometria que age sobre $g_1(z)$, deslocando o ponto $g_1(z)$ sobre o eixo y de forma que este ponto coincida com o ponto i , e finalmente g_3 é uma rotação com centro em i e que gira o ponto $g_2(g_1(w))$ de um ângulo θ , para que $g_2(g_1(w))$ seja um ponto do eixo y .

Assim, através destas isometrias pode-se demonstrar o seguinte teorema, que

fornece expressão de d que estavamos procurando. Para mais detalhes ver [6].

Teorema 2.1 ([2] e [6]). *Sejam $z, w \in \mathbb{H}^2$ e $d(\cdot, \cdot)$ definida como em (2.2). Então*

$$(i) \quad d(z, w) = \ln \left(\frac{|z - \bar{w}| + |z - w|}{|z - \bar{w}| - |z - w|} \right), \quad (2.4)$$

$$(ii) \quad \cosh(d(z, w)) = 1 + \frac{|z - w|^2}{2 \operatorname{Im}(z) \operatorname{Im}(w)}.$$

O conjunto \mathbb{H}^2 com a métrica definida em (2.2) constitui o espaço métrico referente ao modelo do semi-plano superior. Antes de apresentarmos as propriedades deste espaço métrico, iremos apresentar o próximo modelo a ser utilizado.

2.2.2 Modelo do disco aberto unitário

Consideremos, agora, o segundo modelo para o plano hiperbólico. Seja $\Delta = \{z = x + iy : |z| < 1\}$ o disco aberto de raio unitário. De maneira semelhante ao desenvolvimento do modelo do semi-plano superior, a matriz

$$G(z) = \begin{bmatrix} \frac{1}{(1-|z|^2)^2} & 0 \\ 0 & \frac{1}{(1-|z|^2)^2} \end{bmatrix},$$

dá origem ao diferencial

$$ds = \frac{2|dz|}{1-|z|^2},$$

onde $|dz| = \sqrt{dx^2 + dy^2}$.

Definimos uma métrica em Δ como sendo dada por

$$d^*(z, w) = \inf \|\gamma\|, \quad (2.5)$$

onde o ínfimo é tomado sobre todas as curvas C^1 por partes $\gamma : \mathbb{R} \rightarrow \Delta$ conectando z a w . Assim,

Teorema 2.2 ([2] e [6]). *Sejam $z, w \in \Delta$ e d^* definida como em (2.5). Então*

$$(i) \quad d^*(z, w) = \ln \left(\frac{|1 - z \bar{w}| + |z - w|}{|1 - z \bar{w}| - |z - w|} \right), \quad (2.6)$$

$$(ii) \quad \cosh(d^*(z, w)) = \frac{|1 - z \bar{w}|^2 + |z - w|^2}{|1 - z \bar{w}|^2 - |z - w|^2}. \quad (2.7)$$

Apesar de aparentemente distintos, os resultados dos Teoremas 2.1 e 2.2 implicam que os dois modelos são isométricos, pois a aplicação

$$F(z) = \frac{z - i}{z + i},$$

é uma isometria de \mathbb{H}^2 em Δ . A existência de uma isometria entre os dois modelos implica que os correspondentes espaços são essencialmente os mesmos, porém estão representados de formas diferentes. Conseqüentemente, os resultados obtidos em um modelo podem ser levados para o outro e vice-versa.

Observe que

$$F^{-1}(z) = \frac{(1 + z)i}{1 - z},$$

onde $z \in \Delta$ é uma isometria de Δ em \mathbb{H}^2 . Assim sendo, denotaremos as métricas d^* e d por d_h , a menos que o contexto não seja claro.

Algumas vezes é necessário utilizar as coordenadas reais de F . Para tal, seja $z = x + iy \in \mathbb{H}^2$, então

$$F(x, y) = \left(\frac{x^2 + y^2 - 1}{x^2 + (y + 1)^2}, \frac{-2x}{x^2 + (y + 1)^2} \right).$$

Por outro lado, dado $w = x + iy \in \Delta$, então

$$F^{-1}(x, y) = \left(\frac{-2y}{(1 - x)^2 + y^2}, \frac{1 - x^2 - y^2}{(1 - x)^2 + y^2} \right).$$

Além disso, se R é uma região qualquer contida em Δ , a área hiperbólica de Ω é dada por (ver [1])

$$A_h = \iint_R \left[\frac{2}{1 - |z|^2} \right]^2 dx dy.$$

Entretanto, se R está contida em \mathbb{H}^2 o integrando é trocado por $1/y^2$.

O conceito de área hiperbólica é importante, pois através dele podemos calcular o volume de superfícies geradas por funções que tem como domínio o plano hiperbólico. É através do cálculo destes volumes que obtemos os resultados apresentados nos Capítulos 3 e 5.

2.2.3 Propriedades Geométricas em \mathbb{H}^2 e em Δ

Nesta subseção introduzimos os conceitos geométricos que permitirão generalizarmos os correspondentes conceitos euclidianos. Esses conceitos são necessários para a construção das regiões sob as quais iremos efetuar o cálculo de desempenho de um sistema de comunicações.

Em espaços métricos mais gerais que o espaço euclidiano, o conceito de reta é generalizado para o conceito de geodésica, definida como sendo o caminho com a menor distância entre dois pontos distintos. Assim, uma geodésica γ em Δ é uma curva com a propriedade de que dados dois pontos distintos z e w em γ , o arco de γ ligando z a w tem comprimento $d(z, w)$, ou seja, γ é o caminho mais curto entre z e w (no caso euclidiano as geodésicas são as retas). Portanto, denominamos as geodésicas em Δ (ou \mathbb{H}^2) de retas hiperbólicas. No modelo do disco unitário Δ , pode-se mostrar que estas são arcos de círculos que interceptam ortogonalmente o bordo de Δ ou elas

são os diâmetros de Δ (círculos com raio infinito). Do mesmo modo, pode-se mostrar que em \mathbb{H}^2 as retas hiperbólicas são os semi-círculos ou as semi-retas ortogonais ao eixo x .

Sabendo como são as retas em \mathbb{H}^2 , então dados uma reta hiperbólica e um ponto fora desta, existe mais de uma reta passando pelo ponto e que são paralelas à reta dada. Dessa forma, a geometria construída no semi-plano superior é não-euclidiana, pois não satisfaz o quinto axioma de Euclides ou axioma das paralelas.

Uma outra forma de se estabelecer essa diferença é a seguinte: pode-se mostrar que a soma dos ângulos internos de um triângulo em Δ é estritamente menor que π . Este fato será de suma importância neste trabalho, pois como veremos no Capítulo 4, ele nos permitirá construir tesselações em Δ diferentes das que existem em \mathbb{R}^2 . Acrescentamos as seguintes observações:

- (i) Não existe isometria entre \mathbb{H}^2 munido com a métrica dada em (2.4), e \mathbb{R}^2 munido com a métrica usual.
- (ii) Em qualquer um dos modelos considerados anteriormente, a função distância d_h é também uma métrica, logo o espaço hiperbólico é um espaço métrico, mas não é um espaço normado.
- (iii) De (i) temos que $\mathbb{H}^2 \neq \mathbb{H} \times \mathbb{H}$, ou seja, com a métrica de \mathbb{H} , a métrica do espaço produto $\mathbb{H} \times \mathbb{H}$ não é equivalente a métrica de \mathbb{H}^2 dada em (2.4). Como a função exponencial

$$\text{Exp}(x) = e^x \tag{2.8}$$

é uma isometria entre \mathbb{R} e \mathbb{H} , temos que $\mathbb{H} \times \mathbb{H}$ é isométrico a \mathbb{R}^2 .

Chamamos a atenção para o fato de que o uso do símbolo " \mathbb{H} " para denotar a parte positiva de \mathbb{R} , é apenas para manter a consistência notacional. Como pode ser visto de (iii) a reta \mathbb{H} é a imagem de \mathbb{R} pela exponencial (2.8).

As observações (i), (ii) e (iii) são importantes, pois evidenciam as dificuldades em se trabalhar no plano hiperbólico. A afirmação (i) nos diz que resultados válidos em \mathbb{R}^2 não podem ser trivialmente considerados no caso hiperbólico. Em (ii) temos uma grande restrição, pois diz que o espaço hiperbólico não possui uma estrutura vetorial própria, o que gera grandes problemas quando pensamos em sistemas de comunicações, onde o conceito de espaço vetorial é a base dos modelos utilizados. Finalmente (iii) aponta outra grande dificuldade, pois diz que resultados obtidos em duas dimensões não podem ser generalizados diretamente para dimensões maiores.

Como no caso euclidiano, podemos definir círculos em Δ (ou em \mathbb{H}^2). Assim, um círculo C com centro $w \in \Delta$ e raio hiperbólico $r > 0$ é definido por

$$C = \{z \in \Delta : d_h(z, w) = r\}.$$

Para os círculos hiperbólicos temos os seguintes resultados, que valem para ambos os modelos da geometria hiperbólica em consideração,

Teorema 2.3 [1]. (i) *A classe dos círculos hiperbólicos em Δ (ou em \mathbb{H}^2) coincide com a classe dos círculos euclidianos em Δ (ou em \mathbb{H}^2);*

(ii) *A classe dos círculos hiperbólicos abertos em Δ (ou em \mathbb{H}^2) coincide com a classe dos círculos euclidianos em Δ (ou em \mathbb{H}^2) e dado $0 < r < a$, temos que a bola euclidiana de centro $z = (b, a)$ e raio r é uma bola hiperbólica com centro $z' = (b, \sqrt{a^2 - r^2})$ e raio $r_h = 1/2 \ln \left(\frac{a+r}{a-r} \right)$.*

A seguir apresentamos as expressões analíticas de alguns elementos hiperbólicos que serão necessários na análise do projeto de sinais em \mathbb{H}^2 ou em Δ . Estes elementos são importantes na caracterização das regiões de Voronoi associadas às constelações de sinais.

Reta através de dois pontos

Sejam $z = (a, b)$ e $w = (c, d)$ dois pontos de \mathbb{H}^2 , queremos determinar a equação da reta hiperbólica passando por z e w .

Se $a = c$, a reta hiperbólica é um círculo degenerado de raio infinito e a equação é $y = a$. Considere então $a \neq c$; a reta r passando através de z e w tem equação

$$y = \frac{d-b}{c-a}x + \frac{bc-da}{c-a}.$$

O ponto médio v do segmento hiperbólico \overline{zw} tem coordenadas

$$v = \left(\frac{a+c}{2}, \frac{b+d}{2} \right).$$

A reta r_1 passando por v e perpendicular a r tem equação

$$(c-a)x + (b-d)y + \frac{(b^2-d^2)^2}{2} - \frac{(c^2-a^2)}{2} = 0.$$

A intersecção de r_1 com o eixo $y = 0$ é o ponto m dado por

$$m = \left(\frac{b^2-d^2-c^2+a^2}{2(a-c)}, 0 \right),$$

onde m é o centro do semi-círculo procurado. O raio deste semi-círculo é a distância entre z e m dada por

$$r = \frac{1}{2} \sqrt{4a^2 + \frac{(b^2-d^2-c^2+a^2)^2}{(a-c)^2} + 4b^2 - \frac{4a(b^2-d^2-c^2+a^2)}{a-c}}.$$

Logo, o semi-círculo de raio r e centro m tem equação

$$x^2 + y^2 - \frac{(b^2 - d^2 - c^2 + a^2)x}{a - c} - a^2 - b^2 + \frac{a(b^2 - d^2 - c^2 + a^2)}{a - c} = 0, \quad y > 0.$$

Se z e w estão em Δ , a reta hiperbólica determinada por estes pontos é o arco de círculo passando por z e w sendo que este arco de círculo é ortogonal ao bordo de Δ . Para a determinação do arco de círculo passando por z e w , introduzimos o seguinte conceito.

Definição 2.5 *Seja $C(z, r)$ o círculo euclidiano de raio r e centro z . A inversão (ou reflexão) em $C(z, r)$ é a função ϕ dada por*

$$\phi(x) = z + \left(\frac{r}{|x - z|} \right)^2 (x - z).$$

Observe que $|x - z||\phi(x) - z| = r^2$ e que se x está em C , então $\phi(x) = x$. Note que a inversão não está definida em $x = a$. Para defini-la, denotamos por ∞ um ponto que não esteja em \mathbb{H}^2 (chamado de ponto ideal) e definimos $\phi(a) = \infty$. Finalmente, podemos provar que todo arco de círculo que passa através de x e $\phi(x)$, estes distintos, é ortogonal a C . Agora, considere $\partial(\Delta) = C(0, 1) = \{z : |z| = 1\}$, a fronteira de Δ . O inverso de $z = (a, b)$ em relação a $\partial(\Delta)$ é o ponto $\phi(z)$ de coordenadas

$$\phi(z) = \left(\frac{a}{a^2 + b^2}, \frac{b}{a^2 + b^2} \right).$$

Com isso, determinamos o círculo passando pelos pontos z, w e $\phi(z)$. Seja m o ponto médio do segmento $\overline{z\phi(z)}$ cujas coordenadas são

$$m = \left(\frac{(a^2 + b^2 + 1)a}{2(a^2 + b^2)}, \frac{(a^2 + b^2 + 1)b}{2(a^2 + b^2)} \right).$$

Seja n o ponto médio de $\overline{z\bar{w}}$ cujas coordenadas são $n = \left(\frac{a+c}{2}, \frac{b+d}{2} \right)$. Agora, as retas através de $\overline{z\bar{w}}$ e $\overline{z\phi(z)}$ tem equações $y - \frac{bx}{a} = 0$ e $(b - d)x + (c - a)y + da - bc = 0$,

respectivamente. Denote estas retas por r_1 e r_2 . As retas perpendiculares a r_1 e r_2 , através dos pontos m e n , respectivamente, são as retas p_1 e p_2 , cujas equações são

$$(c-a)x - (b-d)y + \frac{b^2-d^2}{2} + \frac{c^2-a^2}{2} = 0,$$

e

$$x + \frac{by}{a} - \frac{\frac{b}{2} \left(b + \frac{b}{a^2+b^2} \right)}{a} - \frac{1}{2}a - \frac{1}{2} \frac{a}{a^2+b^2} = 0.$$

O ponto q , intersecção de p_1 com p_2 , é dado por

$$q = \left(\frac{1}{2} \frac{b^2d - db^2 - bc^2 - b + d + a^2d}{bc - da}, \frac{1}{2} \frac{cb^2 + ca^2 - a + c - d^2a - c^2a}{bc - da} \right).$$

Logo, o raio do círculo procurado é a distância entre q e z dada por

$$r = \frac{1}{2} \left(\alpha_0 + 4a^2 + \alpha_1 + 4b^2 + 4\alpha_2 - 4\alpha_3 \right)^{1/2},$$

onde

$$\alpha_0 = \frac{(b^2d - bd^2 - bc^2 - b + d + a^2d)^2}{(bc - da)^2},$$

$$\alpha_1 = \frac{(cb^2 + ca^2 - a + c - d^2a - c^2a)^2}{(bc - da)^2},$$

$$\alpha_2 = \frac{(b^2d - bd^2 - bc^2 - b + d + a^2d) a}{bc - da},$$

$$\alpha_3 = \frac{(cb^2 + ca^2 - a + c - d^2a - c^2a) b}{bc - da}.$$

Finalmente, a equação do círculo é dada por

$$x^2 + y^2 - \beta_0x + \beta_1y + a^2 - b^2 + \beta_2 - \beta_3 = 0, \quad (2.9)$$

onde

$$\beta_0 = \frac{(-b^2d + bd^2 + bc^2 + b - d - a^2d)}{bc - da}$$

$$\beta_1 = \frac{(-cb^2 - ca^2 + a - c + d^2a + c^2a)}{bc - da}$$

$$\beta_2 = \frac{(-b^2d + bd^2 + bc^2 + b - d - a^2d) a}{bc - da}$$

$$\beta_3 = \frac{(-cb^2 - ca^2 + a - c + d^2a + c^2a) b}{bc - da}$$

Para determinar os extremos do arco do círculo, (2.9), contido em Δ basta calcular a intersecção dos dois círculos. A expressão geral não é apresentada por ser extremamente extensa. Finalmente, observamos que se z, w e $O = (0, 0)$ são colineares, a reta hiperbólica é um diâmetro de Δ .

Ângulo entre duas retas hiperbólicas

Dadas duas retas hiperbólicas em \mathbb{H}^2 , isto é, $\alpha(t)$ e $\beta(t)$, que se encontram em um ponto z , definimos o ângulo entre elas como sendo o ângulo entre as correspondentes tangentes. Este ângulo é determinado pela matriz G . Agora; se parametrizarmos as equações das retas, $\alpha(t), \beta(t)$, tal que $\alpha(0) = \beta(0) = z$, então o ângulo θ é definido como

$$\begin{aligned} \cos \theta &= \frac{\langle \alpha'(0), \beta'(0) \rangle_{G(\alpha(0))}}{\|\alpha'(0)\|_{G(\alpha(0))} \|\beta'(0)\|_{G(\alpha(0))}} \\ &= \frac{\frac{\langle \alpha'(0), \beta'(0) \rangle}{y_0^2}}{\left(\frac{\langle \alpha'(0), \alpha'(0) \rangle}{y_0^2} \right)^{1/2} \left(\frac{\langle \beta'(0), \beta'(0) \rangle}{y_0^2} \right)^{1/2}} \\ &= \frac{\langle \alpha'(0), \beta'(0) \rangle}{\|\alpha'(0)\| \|\beta'(0)\|}, \end{aligned}$$

onde $\alpha'(0)$ denota o valor da derivada da função no ponto 0.

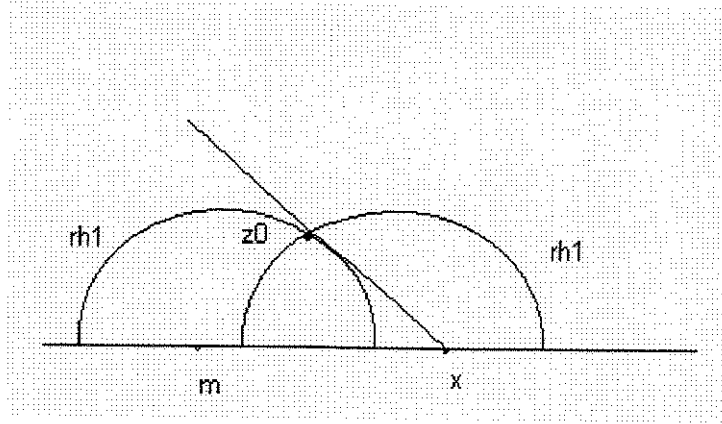


Figura 2.2: Construção da reta perpendicular a uma reta dada r_h , através de um ponto dado de r_h , em H^2 .

Logo, constatamos que este ângulo coincide com o ângulo euclidiano. Portanto, dadas retas hiperbólicas r_1 e r_2 , cujas correspondentes equações são $(x-m)^2 + y^2 = r_1^2$ e $(x-n)^2 + y^2 = r_2^2$, o ângulo α entre r_1 e r_2 é dado por

$$\alpha = \pi - \arccos \left(\frac{(m-n)^2 - r_1^2 - r_2^2}{2r_1 r_2} \right).$$

Reta perpendicular a uma reta dada.

Seja $(x-m)^2 + y^2 = r^2$ uma reta em \mathbb{H}^2 . Denotêmo-a por r_h e seja $z_0 = (a, b)$ um ponto de c . Queremos determinar r_{h_1} passando por z_0 , de forma que r_{h_1} seja perpendicular a r_h . Para isto fazemos a seguinte construção (ver Figura 2.2): seja t a reta euclidiana tangente a r_h passando por z_0 . A reta t encontra o eixo $y = 0$ em um ponto x . Seja $R = |z_0 - x|$, e consideremos o semi-círculo $|z - x| = R, z \in \mathbb{H}^2$. Então, z_0 pertence ao semi-círculo e o ângulo entre $\overline{z_0 m}$ e $\overline{z_0 x}$ é $\pi/2$. Assim, esse semi-círculo é a reta hiperbólica c_1 procurada.

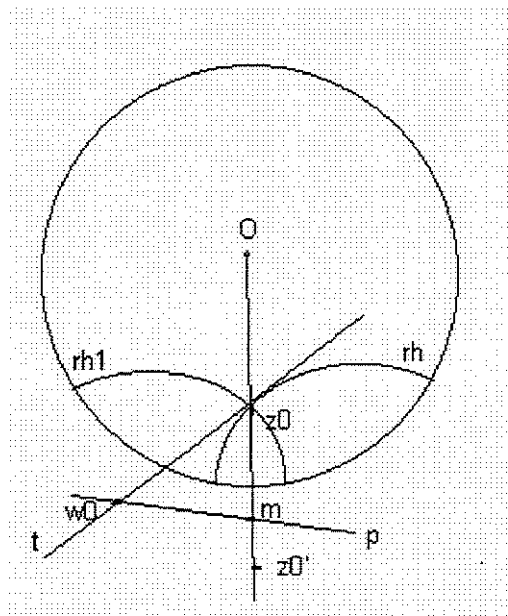


Figura 2.3: Construção da reta perpendicular a uma reta dada, através de um ponto, em Δ .

Se considerarmos a construção da reta perpendicular a uma reta dada, porém em Δ , então o procedimento a ser seguido é o seguinte. Seja t a reta tangente a r_h em z_0 e seja z'_0 o inverso de z_0 em relação a $\partial(\Delta)$ (ver Figura 2.3). Através do ponto médio m do segmento $\overline{z_0 z'_0}$ traçamos a perpendicular p . Esta perpendicular encontra a reta t no ponto w_0 . Agora, o arco do círculo r_{h_1} de centro w_0 e tendo como raio a distância euclidiana entre w_0 e z_0 , é a reta hiperbólica procurada, pois $\overline{w_0 z_0}$ é tangente a r_h . Logo, r_{h_1} é perpendicular a r_h em z_0 . Como z_0 e z'_0 estão em r_{h_1} , então r_{h_1} é perpendicular a Δ .

Ponto médio e bissetriz.

Sejam $z_0 = (a, b)$ e $w_0 = (c, d)$ dois pontos de \mathbb{H}^2 . O objetivo é determinar o ponto médio hiperbólico m de $\overline{z_0 w_0}$. Para isto iremos primeiro determinar a equação da bissetriz c determinada por z_0 e w_0 , ou seja, iremos determinar o conjunto dos pontos $z = x + iy$ de \mathbb{H}^2 tal que $d(z, z_0) = d(z, w_0)$. Utilizando a igualdade (ii) do Teorema 2.1 obtemos

$$\frac{|z - z_0|^2}{\operatorname{Im}[z_0]} = \frac{|z - w_0|^2}{\operatorname{Im}[w_0]}.$$

Desenvolvendo esta igualdade, resulta em

$$x^2 + y^2 + \frac{(-2bc + 2ad)x}{b - d} + \frac{b(c^2 + d^2) - d(a^2 + b^2)}{b - d} = 0 \quad y > 0.$$

Esta é a equação da reta c bissetriz ao segmento determinado por z_0 e w_0 .

Agora, a reta hiperbólica c_1 , que passa por z_0 e w_0 , tem equação

$$x^2 + y^2 - \frac{(b^2 - d^2 - c^2 + a^2)x}{a - c} - a^2 - b^2 + \frac{a(b^2 - d^2 - c^2 + a^2)}{a - c} = 0.$$

Da intersecção de c com c_1 , o ponto médio m é dado por

$$M = \left(\frac{ad + bc}{b + d}, \sqrt{\frac{bd(a^2 - 2ac + d^2 + b^2 + c^2 + 2bd)}{(b + d)^2}} \right).$$

Área de um polígono.

Para construirmos um polígono hiperbólico regular de p lados podemos proceder de forma semelhante ao caso euclidiano. Seja $C(0, r)$ um círculo hiperbólico em Δ centrado na origem de raio hiperbólico r . Primeiro dividimos o círculo em p partes iguais, através de p raios. O ângulo entre dois raios é $2\pi/p$. Portanto, temos p pontos igualmente distribuídos sobre o círculo. Unindo estes pontos pelas retas

hiperbólicas, que cada par determina, temos um polígono hiperbólico de p lados. Pode-se demonstrar que o ângulo θ formado por duas arestas adjacentes é uma função do raio r , e pode assumir qualquer valor entre $0 \leq \theta < \pi$, dependendo do valor de r . Essa construção pode ser adaptada para a construção de polígonos não-regulares, bastando para isto considerarmos raios de valores diferentes.

Seja P um polígono com n lados construído no plano hiperbólico e sejam $\theta_1 \dots \theta_n$ os n ângulos internos formados pelas arestas, então a área do polígono é dada pela fórmula, [1],

$$\text{Area}(P) = (n - 2)\pi - (\theta_1 + \dots + \theta_n). \quad (2.10)$$

No caso particular de um triângulo, temos que a área será dada por $\pi - (\alpha + \beta + \gamma)$, onde α , β e γ são os ângulos internos do triângulo.

Não apresentamos aqui a demonstração deste resultado, pois seria necessário a introdução de uma série de conceitos de geometria Riemanniana, os quais fogem ao escopo deste tema, porém indicamos a referência [1].

Classificação das Isometrias no Plano Hiperbólico.

Para concluirmos esta seção apresentamos uma classificação das isometrias de \mathbb{H}^2 . Esta classificação é importante quanto ao modelo de canal a ser utilizado no Capítulo 5. De forma geral as isometrias de \mathbb{H}^2 são obtidas através de (2.3). Seja M a matriz

$$M = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

A função traço de M é definida como

$$\text{traço}(M) = a_{11} + a_{22},$$

e denotamos por $Tr(M)$ o módulo da função traço.

Definição 2.6 *Dada uma matriz*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

onde a, b, c e d são números reais tais que $ad - bc = 1$, diremos que A e a transformação (isometria) $g_A(z) = \frac{az+b}{cz+d}$ são:

1. *Elíptica se $Tr(A) < 2$.*
2. *Parabólica se $Tr(A) = 2$.*
3. *Hiperbólica se $Tr(A) > 2$.*

Um ponto ideal de \mathbb{H}^2 ou é um ponto do eixo x ou $+\infty$. De maneira genérica denotaremos qualquer um desses pontos por ∞ simplesmente. Um ponto ordinário de \mathbb{H}^2 é um ponto que não é ideal.

Definição 2.7 *Um ponto $z \in \mathbb{H}^2 \cup \{\infty\}$ é um ponto fixo de uma isometria g se $g(z) = z$.*

Através dos pontos fixos de uma isometria do plano hiperbólico podemos caracterizá-la geometricamente através do seguinte resultado.

Teorema 2.4 ([2] e [6]). *São equivalentes as seguintes afirmações, entre itens do mesmo número:*

1i *Uma isometria g é elíptica.*

1ii *g possui um ponto fixo ordinário p .*

1iii *Existe ponto ordinário p tal que as esferas centradas em p são invariantes por*

g .

2i *Uma isometria g é parabólica.*

2ii *g possui um único ponto fixo ideal ξ .*

3i *Uma isometria g é hiperbólica.*

3ii *g possui dois pontos fixos ideais.*

3iii *Existe geodésica invariante por g .*

A família das isometrias de \mathbb{H}^2 é maior que a família das euclidianas, mas é possível fazer um paralelo. As isometrias elípticas são rotações hiperbólicas em torno de um ponto, que é exatamente o ponto fixo da isometria. Isometrias hiperbólicas transladam pontos que estejam em retas euclidianas paralelas ao eixo x e isometrias parabólicas transladam pontos nas retas euclidianas (também hiperbólicas) paralelas ao eixo y .

Citamos aqui que as isometrias no modelo do disco aberto unitário Δ são funções de Moebius com a forma

$$g(z) = \frac{az + \bar{c}}{cz + \bar{a}},$$

onde $z \in \Delta$ e a e c são números complexos tais que $|a|^2 - |c|^2 = 1$. A mesma classificação e resultados apresentados no Teorema 2.4 são válidas para as isometrias de Δ .

Capítulo 3

Determinação da Densidade de Probabilidade Gaussiana Hiperbólica

Neste capítulo iremos determinar a função densidade de probabilidade gaussiana a ser utilizada no plano hiperbólico. A necessidade de se ter um modelo para o canal gaussiano na geometria hiperbólica, se torna clara quando pensamos no modelo tradicional utilizado em sistemas de comunicações na geometria euclidiana. Neste modelo, é assumido que os sinais a serem transmitidos são perturbados por diversos agentes externos, e que são comumente denominados "ruído". Como é de conhecimento geral, a caracterização estatística do ruído vem através do estabelecimento de uma distribuição de probabilidade. Em sistemas de comunicações é usual caracterizarmos o ruído como sendo uma amostra de um processo estocástico gaussiano. Como con-

seqüência, a função densidade de probabilidade associada é gaussiana. O Teorema Central do Limite, no caso euclidiano, fundamenta de maneira precisa o modelo aditivo e gaussiano. No caso hiperbólico, a não existência de tal teorema para casos gerais, implica que a determinação e/ou caracterização da densidade de probabilidade tenha que fazer uso de propriedades geométricas inerentes ao processo gaussiano. Portanto, a função densidade de probabilidade do ruído gaussiano que estabeleceremos no plano hiperbólico apresenta uma característica importante, qual seja a de preservar todas as propriedades geométricas do caso euclidiano.

3.1 Densidade de Probabilidade Gaussiana Hiperbólica em \mathbb{H}

Nesta seção desenvolvemos a função densidade de probabilidade do ruído gaussiano na semi-reta (geodésica) $\mathbb{H} = \{y \in \mathbb{R} : y > 0\}$. Como visto no Capítulo 1, \mathbb{H} é a imagem de \mathbb{R} através da função exponencial, isto é

$$\begin{aligned} r &: \mathbb{R} \rightarrow \mathbb{H} \\ &: x \rightarrow e^x. \end{aligned}$$

Considere X como sendo uma variável aleatória gaussiana euclidiana unidimensional com média a e variância σ^2 . A função densidade de probabilidade de X é dada por

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-a)^2\right).$$

A função distribuição de probabilidade de X é então

$$P(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(t-a)^2\right) dt. \quad (3.1)$$

Dessa forma, os mapeamentos entre \mathbb{H} , \mathbb{R} e o intervalo $[0, 1]$ ficam explicitados por

$$\mathbb{H} \xrightarrow{r^{-1}} \mathbb{R} \xrightarrow{P} [0, 1],$$

logo a composição $P \circ r^{-1}$ nos fornece a função distribuição de probabilidade da variável aleatória $Y = Exp(X)$ em \mathbb{H} . Fazendo a mudança de variável $t = \log(u)$ em $P(x)$, temos

$$F(y) = \int_0^y \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2} \log^2(u/b)\right) \frac{1}{u} du,$$

onde $y = e^x$ e $b = e^a$. Portanto, a função densidade de probabilidade em \mathbb{H} da variável aleatória Y é dada por

$$f(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2} \log^2(y/b)\right). \quad (3.2)$$

O próximo passo é estabelecer a média e a variância de uma variável aleatória em \mathbb{H} . Para isso, seja $\mathcal{F}_{\mathbb{R}} = \{f : \mathbb{R} \rightarrow \mathbb{R} : \int_{-\infty}^{\infty} f(x) dx = 1\}$ o conjunto de todas as funções densidade de probabilidade em \mathbb{R} ; da mesma forma definimos $\mathcal{F}_{\mathbb{H}}$. Esses conjuntos são equivalentes, pois a função $r^* : \mathcal{F}_{\mathbb{H}} \rightarrow \mathcal{F}_{\mathbb{R}}$ definida por

$$r^*(g(x)) = g(e^x),$$

onde $g \in \mathcal{F}_{\mathbb{H}}$, é um difeomorfismo (bijeção diferenciável) entre $\mathcal{F}_{\mathbb{H}}$ e $\mathcal{F}_{\mathbb{R}}$. Temos que se $g \in \mathcal{F}_{\mathbb{H}}$, então $\int_{-\infty}^{\infty} g(e^x) dx = \int_0^{\infty} g(y) \frac{1}{y} dy = 1$. Logo, $g(e^x) \in \mathcal{F}_{\mathbb{R}}$. A média em $\mathcal{F}_{\mathbb{R}}$ é a função $\mu : \mathcal{F}_{\mathbb{R}} \rightarrow \mathbb{R}$ dada por

$$\mu(f) = \int_{-\infty}^{\infty} xf(x) dx.$$

Portanto, para estabelecermos a média em $\mathcal{F}_{\mathbb{H}}$ consideramos as composições

$$\mathcal{F}_{\mathbb{H}} \xrightarrow{r^*} \mathcal{F}_{\mathbb{R}} \xrightarrow{\mu} \mathbb{R} \xrightarrow{r} \mathbb{H}. \quad (3.3)$$

Logo, a média em $\mathcal{F}_{\mathbb{H}}$ é a função μ_h dada por

$$\mu_h(g) = \exp \left(\int_0^{\infty} \log(y) g(y) \frac{1}{y} dy \right).$$

Seguindo o mesmo raciocínio, temos que a variância em $\mathcal{F}_{\mathbb{R}}$ é a função σ^2 dada por $\sigma^2(f) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$, onde μ é a média da variável aleatória com densidade de probabilidade f . Utilizando composições como em (3.3) obtemos que a variância σ_h^2 tal que a densidade de probabilidade é g , isto é, $g \in \mathcal{F}_{\mathbb{H}}$, é dada por

$$\sigma_h^2(g) = \exp \left(\int_0^{\infty} (\log(y) - \mu_h)^2 g(y) \frac{1}{y} dy \right),$$

onde μ_h é a média da variável aleatória com densidade de probabilidade g em $\mathcal{F}_{\mathbb{H}}$.

Em suma, temos a seguinte definição:

Definição 3.1 *Seja $g(y)$ função densidade de probabilidade de uma variável aleatória Y definida em \mathbb{H} . Então, a média de Y é dada por*

$$\mu_h = E(Y) = \exp \left(\int_0^{\infty} \ln(y) g(y) \frac{1}{y} dy \right). \quad (3.4)$$

Se Y tem média μ , então a variância de Y é dada por

$$\sigma_h^2 = E(Y - E(Y))^2 = \exp \left(\int_0^{\infty} [\ln(y) - \mu_h]^2 g(y) \frac{1}{y} dy \right). \quad (3.5)$$

Observe que em comparação com a definição usual de média, isto é, $\int xp(x)dx$, o termo $\ln(y)$ nos fornece, a menos do sinal positivo ou negativo, exatamente a distância de y até o ponto 1. Fazendo uso desta interpretação associada a função $f(y)$, equação (3.2), obtemos $\mu = 0$, pois $f(y)$ está centrada em 1.

Na definição euclidiana de variância, $\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx$, temos que o termo $(x - \mu)^2$ é exatamente o quadrado da distância euclidiana do ponto x ao ponto fixo μ . Com isso, podemos pensar em μ como um ponto de \mathbb{R} , que é o domínio de $p(x)$. O mesmo deve ocorrer no caso hiperbólico. E pela definição acima, o resultado da função de média é um valor com significado hiperbólico, ou seja, está em \mathbb{H} . Se a função tem média hiperbólica μ , o termo que aparece dentro da integral que define a variância hiperbólica deve ser a distância hiperbólica ao quadrado, entre y e o ponto euclidiano cuja distância hiperbólica com sinal (sinal negativo se o ponto é menor que 1 ou positivo se é maior ou igual a 1) ao ponto 1 é μ . Como a função $r(x) = \exp(x)$ é uma bijeção entre \mathbb{R} e \mathbb{H} , o ponto procurado é exatamente $\exp(\mu)$. Com isso, podemos definir a variância como $\int_0^{\infty} \ln^2(y/\exp(\mu)) f(y) (1/y) dy$ onde o termo $\ln(y/\exp(\mu))^2$ é exatamente o quadrado da distância hiperbólica entre y e $\exp(\mu)$. Usando propriedades usuais do logaritmo, reduzimos essa expressão à aquela que aparece na Definição 3.1.

3.2 Densidade de Probabilidade Gaussiana

Hiperbólica em \mathbb{H}^2

Como no caso unidimensional, iremos agora obter a função densidade de probabilidade de uma variável aleatória definida no plano hiperbólico. O modelo escolhido é o do disco unitário Δ , devido ao fato dos cálculos serem mais simples neste modelo. O primeiro problema que encontramos é que, como foi citado no Capítulo 2, não

existe uma isometria entre \mathbb{R}^2 e Δ . Desse modo, não podemos aqui, reproduzir o procedimento utilizado na seção anterior. Primeiramente, devemos lembrar que no estabelecimento da densidade de probabilidade multivariável no plano euclidiano, implicitamente estamos fazendo uso de uma série de vantagens que o \mathbb{R}^2 possui, pelo fato deste ser um espaço vetorial normado. Por exemplo, o conceito de forma quadrática, que é utilizado na definição geral da função densidade de probabilidade gaussiana, aparece de forma bastante natural em \mathbb{R}^n . No caso hiperbólico um cuidado especial deve ser tomado, pois este não é um espaço vetorial normado como é o espaço euclidiano. Assim, conceitos como os de base hiperbólica ou de vetores hiperbólicos não surgem naturalmente como no caso euclidiano. Com isso, deveremos sempre que possível fazer uso das propriedades geométricas existentes em \mathbb{R}^2 e que sejam aplicáveis no caso hiperbólico.

A função densidade de probabilidade de duas variáveis aleatórias independentes e gaussianas X e Y com média zero e variâncias σ^2 é dada por

$$p(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}(x^2 + y^2)\right). \quad (3.6)$$

Esta é a função densidade de probabilidade mais utilizada para representar o ruído em sistemas de comunicações. Esta função será denominada por função densidade de probabilidade gaussiana circular ou simplesmente *gaussiana euclidiana circular*, devido ao fato de que as curvas de nível no plano xy serem círculos. O termo $x^2 + y^2$ que aparece na exponencial é o quadrado da distância euclidiana do ponto (x, y) a origem. Assim, no plano euclidiano a função densidade de probabilidade gaussiana circular possui círculos como curvas de nível e é uma exponencial da distância no

espaço considerado. Com esta interpretação, queremos obter em Δ a função densidade de probabilidade de uma variável aleatória Z em Δ , $f(z)$, dada por

$$f(z) = B \exp(-A d_h^2(z, 0)),$$

onde $z = x + iy = (x, y) \in \Delta$ e que suas curvas de nível sejam círculos em Δ .

As constantes A e B não são independentes pois f é uma densidade de probabilidade em Δ , então o volume sob a superfície gerada por f em Δ deve valer 1, ou seja, devemos ter

$$\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} B \exp\left(-A \ln^2\left(\frac{1+|z|}{1-|z|}\right)\right) \left(\frac{2}{1-|z|^2}\right)^2 dy dx = 1$$

Primeiro fazemos uma mudança para coordenadas polares, ou seja, $x = r \cos \theta$, $y = r \sin \theta$. Com essa mudança de coordenadas conseguimos reduzir a integral dupla acima na seguinte expressão:

$$8\pi B \int_0^1 \exp\left(-A \ln^2\left(\frac{1+r}{1-r}\right)\right) \frac{r}{(1-r^2)^2} dr = 1$$

Seja M dado por

$$M = \int_0^1 \exp\left(-A \ln^2\left(\frac{1+r}{1-r}\right)\right) \frac{r}{(1-r^2)^2} dr.$$

Façamos a mudança de variável

$$y = \ln\left(\frac{1+r}{1-r}\right)$$

em M . Isto conduz a

$$dr = \frac{1-r^2}{2} dy$$

e

$$r = \frac{e^y - 1}{e^y + 1}.$$

Substituindo esses termos em M obtemos

$$M = \frac{1}{8} \int_0^\infty \frac{e^{-Ay^2}(e^{2y} - 1)}{e^y} dy. \quad (3.7)$$

A integral, equação (3.7), pode ser escrita em termos da função erro $\operatorname{erf}(x)$ definida por

$$\operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-y^2/2} dy.$$

Assim,

$$\int_0^\infty \frac{e^{-Ay^2}(e^{2y} - 1)}{e^y} dy = \frac{e^{\frac{1}{4A}} \sqrt{\pi} \operatorname{erf}\left(\frac{1}{2\sqrt{A}}\right)}{\sqrt{A}}.$$

Portanto,

$$M = \frac{\sqrt{\pi}}{8\sqrt{A}} e^{1/(4A)} \operatorname{erf}\left(\frac{1}{2\sqrt{A}}\right).$$

Substituindo o valor de M em $8\pi BM = 1$, chegamos a

$$B = \frac{\sqrt{A}}{\sqrt{\pi^3}} e^{-\frac{1}{4A}} \operatorname{erf}\left(\frac{1}{2\sqrt{A}}\right)^{-1}, \quad (3.8)$$

Finalmente, $f(z)$ é dada por

$$f(z) = \frac{\sqrt{A}}{\sqrt{\pi^3}} e^{-\frac{1}{4A}} \operatorname{erf}\left(\frac{1}{2\sqrt{A}}\right)^{-1} \exp\left(-A \ln^2\left(\frac{1+|z|}{1-|z|}\right)\right). \quad (3.9)$$

Chamamos a função $f(z)$ de gaussiana circular em Δ . Para obter a gaussiana em \mathbb{H}^2 , simplesmente trocamos a expressão da distância que aparece na exponencial pela distância de \mathbb{H}^2 ao quadrado, $d_h^2(z, i)$ onde $z \in \mathbb{H}^2$, conduzindo a

$$f(z) = \frac{\sqrt{A}}{\sqrt{\pi^3}} e^{-\frac{1}{4A}} \operatorname{erf}\left(\frac{1}{2\sqrt{A}}\right)^{-1} \exp\left(-A \ln^2\left(\frac{|z+i|+|z-i|}{|z+i|-|z-i|}\right)\right). \quad (3.10)$$

No disco, se o centro da gaussiana não é a origem, mas um ponto qualquer $z_0 = (a, b)$, no expoente então aparece o termo $d_h^2(z, z_0)$. Desse modo se $z = (x, y)$, então a função densidade de probabilidade gaussiana circular hiperbólica no disco é dada por

$$f(x, y) = B \exp \left(-A \ln^2 \left(\frac{\sqrt{(1 - ax - by)^2 + (bx - ay)^2} + \sqrt{(x - a)^2 + (y - b)^2}}{\sqrt{(1 - ax - by)^2 + (bx - ay)^2} - \sqrt{(x - a)^2 + (y - b)^2}} \right) \right). \quad (3.11)$$

Se $z_0 = (a, b) \in \mathbb{H}^2$ é um ponto dado, da mesma forma temos que aparece o termo $d_h^2(z, z_0)$ no expoente e a função densidade de probabilidade gaussiana hiperbólica circular no semi-plano superior é dada por

$$f(x, y) = B \exp \left(-A \ln^2 \left(\frac{\sqrt{(x - a)^2 + (y + b)^2} + \sqrt{(x - a)^2 + (y - b)^2}}{\sqrt{(x - a)^2 + (y + b)^2} - \sqrt{(x - a)^2 + (y - b)^2}} \right) \right) \quad (3.12)$$

onde B é a constante dada em 3.8. As equações 3.11 e 3.12 representam os casos mais gerais da função densidade de probabilidade gaussiana hiperbólica circular, e a escolha de qual deve ser utilizada pode variar em cada caso, mas sempre deve ser levado em conta a complexidade dos cálculos que deverão ser feitos. A seguir apresentamos dois resultados que caracterizam geometricamente as gaussianas hiperbólicas circulares.

Teorema 3.1 *Seja $f(z) = B \exp(-Ad_h^2(z, z_0))$ uma gaussiana hiperbólica em \mathbb{H}^2 e consideremos a superfície gerada por $f(z)$ sobre \mathbb{H}^2 . Então as curvas de nível $f(z) = m$ são círculos hiperbólicos centrados em z para $m < B$, e para $m \geq B$ são vazios.*

Demonstração: Seja $z_0 = (a, b)$, $z = (x, y)$ pontos de \mathbb{H}^2 e $f(z) = m$ onde

$$m = B \exp \left(-A \ln^2 \left(\frac{|z - \bar{z}_0| + |z - z_0|}{|z - \bar{z}_0| - |z - z_0|} \right) \right).$$

Isto implica que

$$\ln^2 \left(\frac{|z - \bar{z}_0| + |z - z_0|}{|z - \bar{z}_0| - |z - z_0|} \right) = -\frac{1}{A} \ln \left(\frac{m}{B} \right) > 0$$

ou seja,

$$\ln \left(\frac{|z - \bar{z}_0| + |z - z_0|}{|z - \bar{z}_0| - |z - z_0|} \right) = \sqrt{-\frac{1}{A} \ln \left(\frac{m}{B} \right)},$$

ou equivalentemente, $d_h(z, z_0) = r$ onde $r = \exp \left(\sqrt{-\frac{1}{A} \ln \left(\frac{m}{B} \right)} \right)$. Isto nos fornece $\cosh(d_h(z, z_0)) = \cosh(r)$ e utilizando a igualdade (ii) do Teorema 2.2.1 obtemos

$$1 + \frac{|z - z_0|^2}{2 \operatorname{Im}[z] \operatorname{Im}[z_0]} = \cosh(r).$$

Substituindo $z = (x, y)$ e $z_0 = (a, b)$ temos $\frac{|(x-a)+(y-b)i|^2}{2yb} = \cosh(r) - 1$. Desenvolvendo essa expressão e completando o quadrado nos termos em y chegamos a

$$(x - a)^2 + (y - b \cosh(r))^2 = b^2(\cosh^2(r) - 1),$$

que é a equação de um círculo.

O Teorema 3.1 justifica o termo gaussianas hiperbólicas circular utilizado.

Teorema 3.2 *Sejam $f_1(z) = B \exp(-Ad_h^2(z, z_0))$ e $f_2(z) = B \exp(-Ad_h^2(z, z_1))$ duas gaussianas hiperbólicas com centros respectivamente em $z_0 = (a, b)$ e $z_1 = (c, d)$, onde os pontos são tomados em \mathbb{H}^2 . Sejam S_1 e S_2 os gráficos de f_1 e f_2 , então o projeção da intersecção de S_1 com S_2 em \mathbb{H}^2 é uma geodésica.*

Demonstração: Consideremos a igualdade $f_1(z) = f_2(z)$. É evidente que a igualdade vale se, e somente se, $d_h^2(z, z_0) = d_h^2(z, z_1)$, ou seja, $\cosh(d(z, z_0)) = \cosh(d(z, z_1))$.

Agora, utilizando a igualdade (ii) do Teorema 2.1 obtemos

$$1 + \frac{|z - z_0|^2}{2 \operatorname{Im}[z] \operatorname{Im}[z_0]} = 1 + \frac{|z - z_1|^2}{2 \operatorname{Im}[z] \operatorname{Im}[z_1]}. \quad (3.13)$$

Primeiro, se $\operatorname{Im}[z_0] = \operatorname{Im}[z_1]$, então cancelando os termos comuns em (3.8), conduz a

$$(x - a)^2 + (y - b)^2 = (x - c)^2 + (y - d)^2,$$

onde $b = d$. Após algumas manipulações algébricas obtemos

$$x = \frac{c^2 - a^2}{2c - 2a},$$

que é a equação da geodésica, que neste caso é uma reta euclidiana. Se $\text{Im}[z_0] \neq \text{Im}[z_1]$ em (3.13), então após algumas manipulações algébricas chegamos a

$$\frac{(x - a)^2 + (y - b)^2}{b} = \frac{(x - c)^2 + (y - d)^2}{d}. \quad (3.14)$$

Após o desenvolvimento dos termos quadráticos na equação (3.14), chegamos a equação da geodésica

$$x^2 - \frac{2x(ad - cb)}{d - b} + y^2 + \frac{(a^2 + b^2)d - (c^2 + d^2)b}{d - b} = 0.$$

Esta equação representa a equação de um círculo com centro no eixo x . Completando o quadrado em x a equação pode ser escrita na forma

$$\left(x - \frac{ad - cb}{d - b}\right)^2 + y^2 = \left(\frac{ad - cb}{d - b}\right)^2 - \frac{(a^2 + b^2)d - (c^2 + d^2)b}{d - b},$$

de onde obtemos as coordenadas do centro e o valor do raio deste semi-círculo.

A importância deste teorema está no fato que, através dele podemos definir as regiões de Voronoi de uma constelação de pontos, e estas regiões se comportam de forma semelhante ao caso euclidiano.

3.3 As Probabilidades Marginais

Para podermos fazer comparações de resultados obtidos utilizando uma função densidade de probabilidade hiperbólica em \mathbb{H}^2 ou em Δ , com resultados obtidos no caso euclidiano, precisamos de uma maneira para contornar o seguinte problema: dada uma

pdf gaussiana $p(x, y)$ euclidiana com variância σ^2 , qual a *pdf* gaussiana hiperbólica $f(z)$ que utilizamos no plano hiperbólico para fazermos uma análise de desempenho entre os dois espaços?

Para responder a essa questão, introduzimos o conceito de funções marginais de uma função densidade de probabilidade hiperbólica. Lembremos que dada uma *pdf* gaussiana euclidiana $p(x, y)$, associada às variáveis aleatórias X e Y , desta obtemos as densidades marginais, $p(x)$ e $p(y)$ definidas por

$$p(x) = \int_{-\infty}^{\infty} p(x, y) dy \quad (3.15)$$

e

$$p(y) = \int_{-\infty}^{\infty} p(x, y) dx, \quad (3.16)$$

e se $p(x, y)$ for uma *pdf* gaussiana euclidiana circular, as variâncias de suas marginais serão iguais. Portanto, o problema que queremos resolver pode ser posto da seguinte forma: dada uma *pdf* gaussiana euclidiana circular $p(x, y)$, onde suas marginais tem variância $\sigma_x^2 = \sigma_y^2 = \sigma^2$, queremos determinar uma *pdf* gaussiana hiperbólica circular $f(z)$ que tenha a mesma "variância" de $p(x, y)$. Para isto, precisamos definir no caso hiperbólico o conceito de "variância" e esta definição será proveniente de funções que farão no caso hiperbólico o papel das marginais do caso euclidiano.

As funções definidas em (3.15) e (3.16) podem ser interpretadas de uma outra forma. Sabemos que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1 \quad (3.17)$$

e que qualquer mudança de variável que fizermos nesta integral não alterará o resultado final. Além disso, temos que as definições de $p(x)$ e $p(y)$ estão contidas nesta

integral, pois $1 = \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} p(x, y) dy) dx = \int_{-\infty}^{\infty} p(x) dx$, o mesmo acontecendo com $p(y)$. Geometricamente na integral $1 = \int_{-\infty}^{\infty} (\int_{-\infty}^{\infty} p(x, y) dy) dx$, estamos considerando os eixos coordenados de \mathbb{R}^2 e em relação a estes construímos uma cobertura do plano e integramos sobre esta cobertura; neste caso especificamente a cobertura é o feixe de todas as retas perpendiculares ao eixo x .

Agora, seja Z uma variável aleatória em \mathbb{H}^2 . Como \mathbb{H}^2 não é um espaço vetorial, não é correto afirmar que Z é a soma de duas variáveis, ou seja, $Z = X + iY$. Em textos que tratam de probabilidade em \mathbb{H}^2 , como [9], uma variável aleatória neste espaço é simplesmente uma variável complexa em \mathbb{C} , o conjunto dos números complexos. Para contornarmos essa questão, introduziremos um sistema de eixos cartesianos hiperbólicos x_h e y . Portanto, através da equação (3.17) será possível definir em \mathbb{H}^2 as densidades marginais de uma variável aleatória Z .

Dada uma *pdf* gaussiana hiperbólica $f(z)$, consideremos no semi-plano as retas hiperbólicas $x = 0$ com $y > 0$, e $x^2 + y^2 = 1$, as quais denominamos de y e x_h respectivamente. A Figura 3.1 ilustra esta consideração.

Denominamos estas retas hiperbólicas de eixos coordenados de \mathbb{H}^2 . É possível obter um sistema de coordenadas hiperbólicas utilizando-se dessas retas. Assim, fixado um desses eixos coordenados, consideramos como cobertura de \mathbb{H}^2 o feixe de retas que é perpendicular ao eixo dado. Iniciemos com o eixo y .

Dado $r > 0$ seja c_r a reta hiperbólica que passa através do ponto ir no eixo y e é perpendicular a ele. Com isso, o conjunto $\{c_r : r > 0\}$ é uma cobertura de \mathbb{H}^2 . Logo,

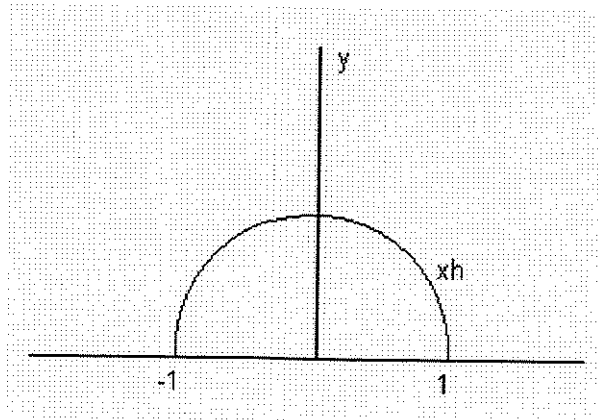


Figura 3.1: Eixos coordenados para o caso hiperbólico.

se na integral $1 = \int \int_{\mathbb{H}^2} f(z) \frac{1}{y^2} dz$ fizermos a mudança de variáveis

$$x = r \cos \theta$$

$$y = r \sin \theta$$

onde $\theta \in (0, \pi)$, após algumas manipulações algébricas chegamos a

$$\begin{aligned} 1 &= \int_0^\infty \int_0^\pi f(r \cos(\theta), r \sin(\theta)) \frac{1}{r^2 \sin(\theta)^2} r d\theta dr \\ &= \int_0^\infty \left[\int_0^\pi f(r \cos(\theta), r \sin(\theta)) \frac{1}{\sin(\theta)^2} d\theta \right] \frac{dr}{r}. \end{aligned}$$

Assim a função densidade de probabilidade marginal de $f(z)$ em relação ao eixo y (entendido a v.a. Y) é a função $f(r)$ dada por

$$f(r) = \int_0^\pi f(r \cos(\theta), r \sin(\theta)) \frac{1}{\sin(\theta)^2} d\theta,$$

Note que $f(r)$ é uma *pdf* em \mathbb{H} por construção.

Consideremos agora o eixo x_h . A cobertura considerada é o conjunto de todas as retas hiperbólicas que são perpendiculares a x_h . Todavia, neste caso um cuidado

especial deverá ser considerado, pois essa cobertura contém o eixo y , que euclidianamente não é da mesma forma que os outros elementos, que são semi-círculos. Para resolver este problema, dividiremos x_h em três partes: $x > 0$, $x = 0$ e $x < 0$ onde $x \in (-1, 1)$. Dado $0 < a < 1$ a esse ponto associamos o ponto $z = (a, \sqrt{1-a^2})$ de x_h . Seja c_z o círculo perpendicular a x_h passando por z . O conjunto $C_+ = \{c_z : z \in x_h \text{ e } \operatorname{Re}(z) > 0\}$ forma uma cobertura da parte de \mathbb{H}^2 que tem pontos com coordenada $x > 0$. Uma reta hiperbólica c_z em C_+ é representada por

$$\left(x - \frac{1}{a}\right)^2 + y^2 = \frac{1}{a^2} - 1.$$

Seja M a constante dada por

$$M = \int_0^\infty \int_0^\infty f(z) \frac{dx dy}{y^2}. \quad (3.18)$$

Logo, se em (3.18) fizermos a mudança de variáveis

$$x = \cos \theta \sqrt{\frac{1}{a^2} - 1} + \frac{1}{a}$$

$$y = \sin \theta \sqrt{\frac{1}{a^2} - 1}$$

onde $\theta \in (0, \pi)$, e calculando o Jacobiano desta transformação chegamos a

$$\frac{dx dy}{y^2} = \frac{|a \cos \theta \sqrt{1-a^2} + |a||}{a^2(1-a^2) \sin^2 \theta} da d\theta.$$

Substituindo em (3.18) temos

$$\begin{aligned} M &= \int_0^1 \int_0^\pi f \left(\cos \theta \sqrt{\frac{1}{a^2} - 1} + \frac{1}{a}, \sin \theta \sqrt{\frac{1}{a^2} - 1} \right) \frac{|a \cos \theta \sqrt{1-a^2} + |a||}{a^2(1-a^2) \sin^2 \theta} da d\theta \\ &= \int_0^1 \left[\int_0^\pi f \left(\cos \theta \sqrt{\frac{1}{a^2} - 1} + \frac{1}{a}, \sin \theta \sqrt{\frac{1}{a^2} - 1} \right) \frac{|a \cos \theta \sqrt{1-a^2} + |a||}{a^2 \sin^2 \theta} d\theta \right] \frac{da}{1-a^2} \end{aligned}$$

Portanto, para $0 < a < 1$ definimos a função densidade de probabilidade marginal $f(a)$ por

$$f(a) = \int_0^\pi f \left(\cos \theta \sqrt{\frac{1}{a^2} - 1} + \frac{1}{a}, \sin \theta \sqrt{\frac{1}{a^2} - 1} \right) \frac{|a \cos \theta \sqrt{1 - a^2} + |a||}{a^2 \sin^2 \theta} d\theta.$$

Agora seja $-1 < a < 0$ e aqui estamos considerando o feixe de retas hiperbólicas c_z , contidas em C_- , onde $C_- = \{c_z : z \in x_h \text{ e } \operatorname{Re}(z) < 0\}$. Como $\int_{\mathbb{H}^2} f(z) 1/y^2 dz = 1$, para $-1 < a < 0$ temos que $\int_{-\infty}^0 \int_0^\infty f(z) \frac{dx dy}{y^2} = 1 - M$. Fazendo a mesma mudança de variáveis que no caso $0 < a < 1$, concluímos que o Jacobiano desta transformação é o mesmo que o do caso anterior. Assim, $f(a)$ fica determinada para $-1 < a < 0$. Nos resta analisar o que acontece para $a = 0$. Neste caso a reta hiperbólica é exatamente o eixo y e como $1 = \int_{-\infty}^\infty (\int_0^\infty f(x, y) 1/y^2 dy) dx$ a integral $\int_0^\infty f(x, y) 1/y^2 dy$ existe para todo x . Conseqüentemente, quando $x = 0$ estamos considerando exatamente o eixo y . Então

$$f(0) = \int_0^\infty f(0, y) 1/y^2 dy.$$

Com essa definição pode-se demonstrar que

$$\lim_{a \rightarrow 0} f(a) = f(0).$$

Com isso, temos como calcular $f(a)$ para todo $-1 < a < 1$. Observe que se na integral $\int_{-1}^1 f(a) \frac{da}{1-a^2} = 1$ fizermos a mudança de variável $w = \frac{1+a}{1-a}$ obtemos

$$1 = \int_{-1}^1 f(a) \frac{da}{1-a^2} = \int_0^\infty \frac{1}{2} f \left(\frac{w-1}{w+1} \right) \frac{dw}{w}.$$

Definindo $g(w) = \frac{1}{2} f \left(\frac{w-1}{w+1} \right)$, temos que a função obtida em relação ao eixo x_h se comporta como a função $f(r)$. Aqui cabe observar que a função $f(a)$ está escrita na

variável a para facilitar sua expressão, pois na realidade é uma função do sistema de coordenadas associado a x_h , e para isto basta observar que todo ponto de x_h pode ser escrito em função de a .

As funções $f(r)$ e $g(w)$ são funções densidade de probabilidade hiperbólicas em \mathbb{H} . Logo podemos calcular a média e a variância de cada uma dessas variáveis aleatórias. Note que $f(r)$ e $g(w)$ têm um comportamento distinto daquele verificado no caso euclidiano no seguinte sentido: seja $p(x, y)$ uma *pdf* euclidiana, com marginais $p(x)$ e $p(y)$. Denotemos por $P = (x, y)$ um ponto qualquer de \mathbb{R}^2 e por μ_x a média da variável aleatória X . Sejam r_{μ_x} a reta perpendicular ao eixo x através do ponto com coordenada igual a μ_x e $d(P, r_{\mu_x})$ a distância do ponto P a reta r_{μ_x} . Temos para a variável aleatória X que

$$\begin{aligned}\sigma_x^2 &= \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x) dx = \int_{-\infty}^{\infty} (x - \mu_x)^2 \left(\int_{-\infty}^{\infty} p(x, y) dy \right) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d(P, r_{\mu_x})^2 p(x, y) dy dx.\end{aligned}$$

Devido ao paralelismo euclidiano entre r_{μ_x} e o eixo y , temos $d(P, r_{\mu_x}) = (x - \mu_x)^2$ e essa propriedade não se reproduz em \mathbb{H}^2 . Por exemplo, dada uma *pdf* $f(x, y)$ hiperbólica tal que $f(r)$ e $g(w)$ tenham médias hiperbólicas iguais a 1, temos que

$$\begin{aligned}\sigma_r^2 &= \exp \left(\int_0^{\infty} \ln^2(r) f(r) \frac{dr}{r} \right) \\ &= \exp \left(\int_0^{\infty} \ln^2(r) \int_0^{\pi} f(r \cos(\theta), r \sin(\theta)) \frac{1}{\sin^2(\theta)^2} \frac{dr d\theta}{r} \right) \\ &\neq \exp \left(\int \int_{\mathbb{H}^2} d^2(P, x_h) f(x, y) \frac{dx dy}{y^2} \right).\end{aligned}$$

Isso motiva a seguinte definição:

Definição 3.2 Dada uma *pdf* hiperbólica $f(z)$ com densidades marginais $f(r)$ e $g(w)$

com médias respectivamente μ_r e μ_w , então a pseudo-variância de $f(z)$ em relação ao eixo y é dada por

$$p\sigma_r^2 = \int \int_{\mathbb{H}^2} d^2(P, r_{\mu_r}) f(x, y) \frac{dx dy}{y^2}$$

onde r_{μ_r} é a reta hiperbólica perpendicular a y passando pelo ponto μ_r marcado como coordenada hiperbólica em y e $d(P, r_{\mu_r})$ é a distância hiperbólica do ponto $P = (x, y)$ de \mathbb{H}^2 à reta r_{μ_r} . Da mesma forma, definimos a pseudo-variância de $f(z)$ em relação ao eixo x_h por

$$p\sigma_w^2 = \int \int_{\mathbb{H}^2} d^2(P, r_{\mu_w}) f(x, y) \frac{dx dy}{y^2}$$

onde r_{μ_w} é a reta hiperbólica perpendicular ao eixo x_h passando por μ_w marcado hiperbolicamente em x_h , ou seja, é ponto de x_h que tem distância $\ln(\mu_w)$ até i , estando a esquerda de y se $\mu_w < 0$ e a direita se $\mu_w > 0$.

Aqui não utilizamos a exponencial, pois esta aparece nas definições em \mathbb{H} devido ao fato da função exponencial ser uma isometria entre \mathbb{R} e \mathbb{H} , e como mencionado, tal isometria não existe no caso bidimensional. Através do conceito de pseudo-variância, conhecendo-se as variâncias de uma *pdf* gaussiana euclidiana $p(x, y)$, podemos determinar uma *pdf* gaussiana hiperbólica que tenha as pseudo-variâncias iguais as variâncias da função euclidiana.

3.4 Modelo do Canal Gaussiano no Espaço

Hiperbólico.

Nas Seções 3.1 e 3.2 obtivemos as densidades de probabilidade para o ruído gaussiano no caso hiperbólico. Mas como esse ruído age sobre um ponto x que é transmitido?

No caso euclidiano o modelo utilizado para um canal gaussiano assume que o ruído é aditivo, ou seja dado um sinal x a ser transmitido, sobre este age um vetor ruído n de forma que o sinal recebido é dado por $y = x + n$. Note que a adição do ruído ao sinal pode ser interpretada como uma translação euclidiana $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$T(x) = x + n,$$

e o vetor y como a imagem de x por esta translação.

Dessa forma, interpretamos o ruído aditivo gaussiano euclidiano como uma família de isometrias euclidianas (translações) agindo sobre um sinal x . Assim assumir que no plano euclidiano age um ruído aditivo gaussiano, é equivalente a assumirmos em \mathbb{R}^2 a seguinte hipótese:

Hipótese do Ruído Gaussiano Euclidiano: Seja G_t o grupo das translações de \mathbb{R}^2 . Existe um espaço de probabilidade (Ω, ε, P) e uma variável aleatória $X : \Omega \rightarrow \mathbb{R}^2$, onde $\Omega = \mathbb{R}^2$, ε é a σ -álgebra em Ω dada pelos Borelianos e P é uma medida de probabilidade tal que $P(g(A)) = P(A)$ para todo $A \in \Omega$ e $g \in G_t$, tal que X tenha como função de distribuição de probabilidade uma gaussiana da forma (3.6).

Assim, seja x um sinal que queremos transmitir em \mathbb{H}^2 . Sabemos que no espaço hiperbólico não é possível definir uma estrutura vetorial como em \mathbb{R}^2 , portanto não

podemos assumir que o modelo de ruído gaussiano hiperbólico que temos é aditivo. Mas utilizando o conceito de isometria, de forma semelhante ao feito no caso euclidiano, assumir um ruído gaussiano em \mathbb{H}^2 significa que estamos assumindo a seguinte hipótese.

Hipótese do Ruído Gaussiano Hiperbólico: Seja G_h o conjunto das isometrias parabólicas e hiperbólicas de \mathbb{H}^2 . Existe um espaço de probabilidade (Ω, ε, P) e uma variável aleatória $Z : \Omega \rightarrow \mathbb{H}^2$ onde, $\Omega = \mathbb{H}^2$, ε é a σ -álgebra em Ω dada pelos Borelianos e P é uma medida de probabilidade tal que $P(g(A)) = P(A)$ para todo $A \in \Omega$ e $g \in G_h$, tal que Z tenha como função de distribuição de probabilidade uma gaussiana da forma (3.9).

As funções densidade de probabilidade $f(z)$ obtidas na Seção 3.2 são invariantes por qualquer isometria de \mathbb{H}^2 . Como citado na introdução, o fato de não termos disponível ainda um Teorema Central do Limite geral para o caso hiperbólico, não nos permite estabelecer exatamente qual tipo de isometria é a ideal para a caracterização do ruído hiperbólico. Como as isometrias elípticas possuem um ponto fixo ordinário, elas não servem para essa caracterização, pois se o ruído fosse caracterizado por uma dessas isometrias, o seu ponto fixo não sofreria nenhum tipo de interferência. Por outro lado as isometrias parabólicas e hiperbólicas não possuem pontos fixos ordinários. Também vale a seguinte propriedade para essas isometrias: dados dois pontos fixos z e w de \mathbb{H}^2 , existem pelo menos uma isometria parabólica e uma hiperbólica que leva z em w .

Aqui deve ser ressaltado que o conjunto das isometrias parabólicas e hiperbólicas não forma um grupo. Se quisermos que o conjunto de isometrias que caracteriza o

ruído seja um grupo, como no caso das translações no plano euclidiano, o conjunto G_t deve somente conter isometrias parabólicas ou hiperbólicas.

Citamos que existe um Teorema Central do Limite no caso hiperbólico ([12]), mas este somente vale para o caso em que as variáveis aleatórias tem densidades centradas na origem. Além disso, essas densidades devem ser invariantes por elementos de $SO(2)$, ou seja, pelas isometrias que possuem matrizes da forma
$$\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix},$$
 ou seja, as rotações euclidianas.

No caso de \mathbb{H} , temos que as isometrias são da forma $g(y) = e^k y$ com $y \in \mathbb{H}$ e $k \in \mathbb{R}$, portanto dado $x \in \mathbb{H}$, assumir que sobre x age um ruído gaussiano hiperbólico unidimensional é assumir o seguinte:

Hipótese do Ruído Gaussiano em \mathbb{H} : Seja G_{he} o conjunto das isometrias hiperbólicas de \mathbb{H} . Existe um espaço de probabilidade (Ω, ε, P) e uma variável aleatória $X : \Omega \rightarrow \mathbb{H}$ onde, $\Omega = \mathbb{H}$, ε é a σ -álgebra em Ω dada pelos Borelianos e P é uma medida de probabilidade tal que $P(g(A)) = P(A)$ para todo $A \in \Omega$ e $g \in G_{he}$, tal que X tenha como função distribuição de probabilidade uma gaussiana da forma (3.1).

Capítulo 4

Tesselações dos Planos Euclidiano e Hiperbólico e as Correspondentes Funções Enumeradoras

No Capítulo 2 vimos que no plano hiperbólico podem ser construídos polígonos da mesma maneira que são construídos polígonos no plano euclidiano. Iniciamos este capítulo mostrando que a construção de certas famílias de reticulados, que chamaremos de tesselações, é profundamente diferente nas duas geometrias. Estabelecidas estas diferenças, veremos que as constelações de sinais mais comumente utilizadas em sistemas de comunicações, são subconjuntos finitos de reticulados (tesselações). Dessa forma veremos que das tesselações no espaço hiperbólico, poderemos extrair constelações de sinais apropriadas para uso eficiente em sistemas de comunicações. Para podermos comparar o desempenho de uma constelação finita de sinais no plano

euclidiano com uma no plano hiperbólico, devemos manter o mesmo número de sinais em ambas as constelações, com a mesma energia média.

A partir da Seção 4.2, apresentamos um procedimento que nos permite enumerar o número de pontos de certos subconjuntos finitos de tesselações, tanto euclidianas quanto hiperbólicas. A relevância deste procedimento está relacionada com o projeto de constelações de sinais e às propriedades inerentes.

4.1 Tesselações nos Planos Euclidiano e Hiperbólico

Nesta seção introduzimos o conceito de tesselação regular dos planos euclidiano e hiperbólico. É através deste conceito que nos deparamos com mais uma grande diferença entre as duas geometrias. Estabelecido o conceito de tesselação, veremos que no caso euclidiano as constelações de sinais utilizadas em modulação, são subconjuntos finitos de tesselações neste plano. Os resultados a serem apresentados permitem estabelecer quais subconjuntos finitos das tesselações poderão ser utilizados como constelações de sinais.

Definição 4.1 *Uma tesselação regular $\{p, q\}$ do plano euclidiano (hiperbólico) é uma cobertura deste plano por polígonos regulares euclidianos (hiperbólicos) isométricos não sobrepostos. Os polígonos têm p lados (e são abreviados por p -ágonos), que se encontram somente em arestas completas ou em vértices, com q polígonos regulares se encontrando em cada vértice.*

Com esta definição, a primeira questão que se levanta é: Quais são as possíveis tesselações regulares que existem nas geometrias em consideração?

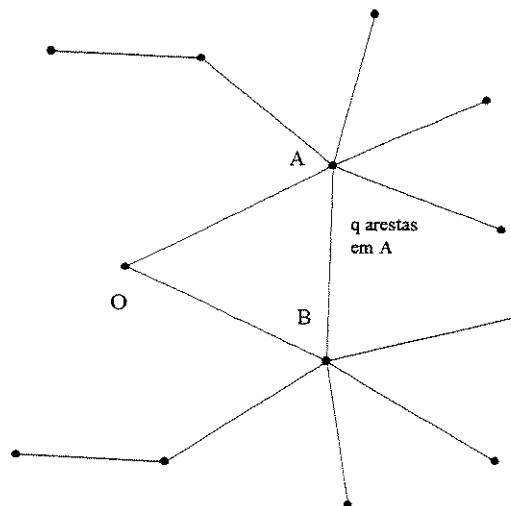


Figura 4.1: Determinação das possíveis tesselações nas geometrias consideradas.

Considere uma tesselação $\{p, q\}$ do plano euclidiano. Seja O o centro de um p -ágono e \overline{AB} uma de suas arestas, como mostra a Figura 4.1. O ângulo interior do encontro de q p -ágonos em um vértice é $2\pi/q$.

Da figura 4.1 concluímos que o ângulo \widehat{AOB} mede $2\pi/p$. Com isso, no caso euclidiano temos para o triângulo AOB , que

$$\begin{aligned} \pi - 2\pi/p &= \widehat{OAB} + \widehat{OBA} \\ &= 2\widehat{OAB} \\ &= \text{o ângulo entre duas arestas adjacentes do polígono.} \end{aligned}$$

Mas para uma tesselação regular, o ângulo entre arestas adjacentes é igual ao ângulo interior de um p -ágono em um vértice, logo $\pi - 2\pi/p = 2\pi/q$. Isto nos fornece a equação

$$(p-2)(q-2) = 4. \quad (4.1)$$

Esta equação tem três soluções inteiras positivas. Portanto existem somente três tesselações no plano euclidiano: $\{3, 6\}$, $\{4, 4\}$ e $\{6, 3\}$. Assim, podemos tesselar o plano euclidiano utilizando triângulos equiláteros, quadrados e hexágonos.

Considere uma tesselação $\{p, q\}$ no plano hiperbólico. Como vimos no Capítulo 2, a soma dos ângulos internos de um triângulo é menor que π , portanto o ângulo entre duas arestas adjacentes de um p -ágono hiperbólico é menor que $\pi - 2\pi/p$. Assim, $2\pi/q < \pi - 2\pi/p$ ou equivalentemente

$$(p - 2)(q - 2) > 4. \quad (4.2)$$

Como existem infinitas soluções inteiras para esta desigualdade, existe um número infinito de tesselações regulares do plano hiperbólico.

Aqui devemos observar a seguinte propriedade; no plano hiperbólico, dois triângulos que têm os mesmos ângulos internos são congruentes, logo possuindo a mesma área. Assim, a menos de movimentos rígidos (por isometrias), não existe o conceito de semelhança. Dados p e q que satisfazendo a desigualdade (4.2), a tesselação $\{p, q\}$ é única, a menos de movimentos por isometrias.

Consequentemente, no plano euclidiano somente três tipos de tesselações regulares são possíveis, porém tendo a liberdade de fixarmos qualquer medida positiva para o comprimento das arestas dos polígonos. No caso hiperbólico, existem infinitas tesselações regulares, mas em cada uma o comprimento das arestas dos polígonos é determinado por $\{p, q\}$. A Figura 4.2 mostra a aparência de uma tesselação hiperbólica no modelo do disco.

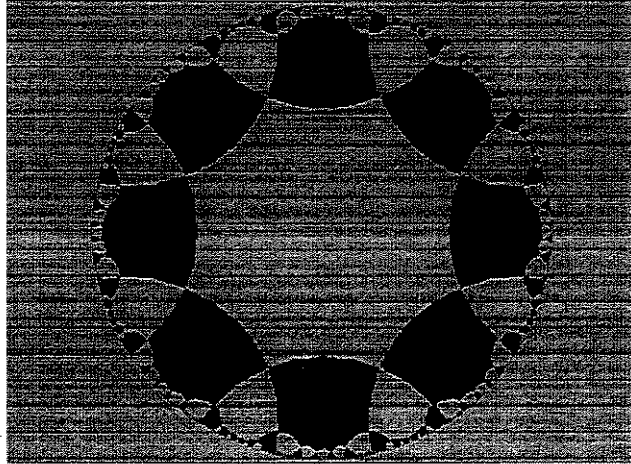


Figura 4.2: Exemplo de tesselação $\{8, 4\}$ construída no plano hiperbólico.

4.2 Subconjuntos Originados por um Polígono Inicial

As constelações de sinais usadas em sistemas de comunicações são subconjuntos finitos das tesselações euclidianas. Com isso as afirmações feitas na Seção 4.1 são extremamente importantes, pois mostram que dispomos de infinitas escolhas para novas constelações de sinais no plano hiperbólico. Diante disto, será necessário desenvolver um procedimento de contagem da evolução do número de pontos em subconjuntos finitos de uma tesselação.

A seguir, apresentamos os conceitos necessários para que possamos avaliar o número de pontos de certos subconjuntos das tesselações, tanto euclidianas quanto hiperbólicas.

Fixada uma tesselação $\{p, q\}$, escolhemos um polígono desta, que denotamos por

L_0 . Denominamos este polígono de região inicial ou nível zero. Os p vértices deste polígono serão denotados por $v_{0,1} \dots v_{0,p}$.

Consideremos agora a união de L_0 com todos os polígonos de $\{p, q\}$ que possuem um vértice em comum com L_0 , e denotemos esse conjunto por C_1 . Com isso, o primeiro nível da tesselação, L_1 , é o conjunto $L_1 = C_1 \setminus L_0$, isto é, o conjunto de polígonos tendo vértices em comum com L_0 , onde excluímos os vértices de L_0 . Os vértices de L_1 são denotados por $v_{1,1} \dots v_{1,k}$. Os próximos níveis L_2, L_3, \dots , são construídos da mesma forma. Os conjuntos C_j também são definidos do mesmo modo de C_1 , ou seja C_{j+1} é o conjunto dos polígonos que tem um vértice em C_j e $L_{j+1} = C_{j+1} \setminus C_j$.

Conhecendo o número de pontos (vértices e centros dos polígonos) no nível i , nosso objetivo é determinar quantos novos pontos (vértices ou centros de polígonos) existirão no nível $i + 1$.

Definição 4.2 *Consideremos o nível L_i de uma tesselação $\{p, q\}$ juntamente com seus níveis inferiores. Um vértice $v_{i,j}$ é do tipo k se em $v_{i,j}$ se encontram k arestas de C_i , onde k é sempre maior ou igual a 2.*

Dado L_i existem dois conjuntos de pontos que iremos quantificar. Um deles são os vértices de L_i e o outro são os centros dos polígonos que formam L_i . De modo a tornar sistemático o processo de contagem, iremos dividir o conjunto das tesselações $\{p, q\}$ em três casos. Em cada caso avaliaremos primeiro o número de vértices, e depois o número de polígonos que ocorre em cada nível. Observe que contar polígonos é o mesmo que contar o número de centros.

Esta divisão em três casos é feita em relação aos possíveis valores de $\{p, q\}$. A saber: Caso 1, com $p \geq 4$ e $q \geq 4$; Caso 2, com $p = 3$; e Caso 3, com $p \geq 6$ e $q = 3$.

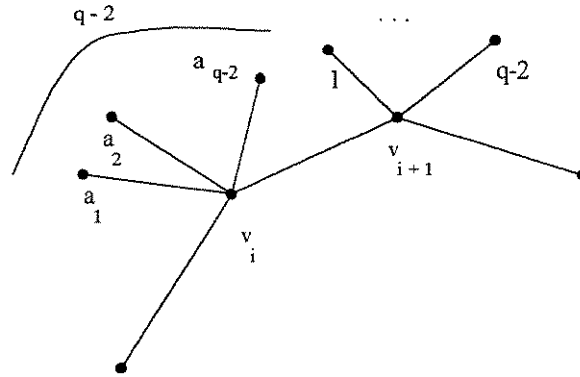


Figura 4.3: Em cada vértice do polígono inicial são inseridas $q - 2$ novas arestas.

Isto é necessário devido as particularidades de cada caso.

O processo é construtivo, iniciamos com o polígono inicial de uma tesselação, e a este iremos acrescentando arestas e contando o crescimento do número de vértices desta região. Esta análise é feita até o momento em que o processo se torna um processo de recorrência, ou seja, sabendo quantos pontos temos e a forma como estão distribuídos, então é possível saber quantos pontos terá o próximo nível.

4.2.1 CASO 1: $p \geq 4$ e $q \geq 4$.

Seja $\{p, q\}$ uma tesselação satisfazendo $p \geq 4$ e $q \geq 4$. Então em L_0 temos p vértices do tipo 2. Como cada vértice deve ter q arestas, de cada vértice em L_0 inserimos $q - 2$ novas arestas e de L_0 como um todo, inserimos um total de $p(q - 2)$ novas arestas. Sejam a_1, a_2, \dots, a_{q-2} as arestas inseridas em $v_{0,i}$ (Figura 4.3).

As arestas a_1 e a_2 são parte de um novo polígono de p lados e p vértices. Como a_1 e a_2 tem em comum o vértice $v_{0,i}$, já temos então 3 pontos do novo polígono,

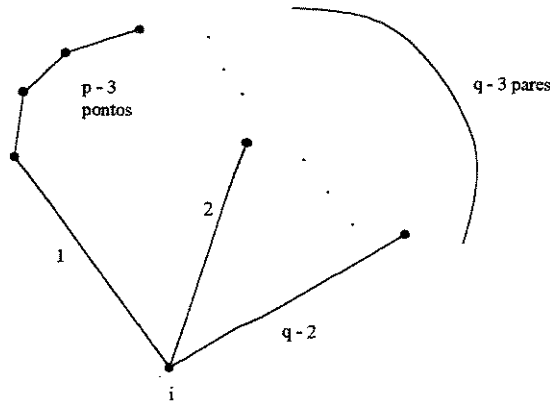


Figura 4.4: Em cada um dos $q - 3$ pares, originados das arestas inseridas em v_i , inserimos $p - 3$ pontos

portanto temos de acrescentar $p - 3$ novos pontos entre o par de arestas $a_1 a_2$ e tendo $q - 2$ arestas em $v_{0,i}$, temos então $q - 3$ pares de arestas como $a_1 a_2$, num total de $(q - 3)(p - 3)$ novos pontos obtidos a partir do vértice $v_{0,i}$ (Figura 4.4). Existem p vértices $v_{0,i}$, resultando em $p(p - 3)(q - 3)$ pontos, todos do tipo 2, pois cada ponto entre duas arestas está ligado por somente duas arestas.

Os vértices $v_{0,i}$ e $v_{0,i+1}$ estão ligados por uma aresta, e de cada um deles partem $q - 2$ arestas. Sendo $p \geq 4$, entre os dois vértices devemos ter um polígono com p lados, logo inserimos $p - 4$ novos pontos entre a última aresta de $v_{0,i}$ e a primeira de $v_{0,i+1}$ (Figura 4.5). Para a contagem adotamos um sentido de contagem horário. Em L_0 temos p pares como esse, então temos o total de $p(p - 4)$ novos pontos, todos do tipo 2.

Para concluirmos o primeiro nível, precisamos determinar qual o tipo dos $p(q - 2)$ pontos originados pelas arestas de L_0 . Sejam a_1, \dots, a_{q-2} as arestas de $v_{0,i}$. Como

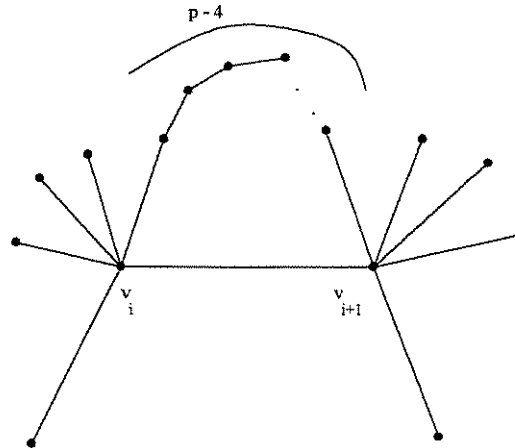


Figura 4.5: Entre dois vértices consecutivos v_i e v_{i+1} inserimos $p - 4$ pontos.

temos $p - 3 > 1$, da mesma forma que a construção anterior, entre cada par destas arestas existe pelo menos um ponto, determinando assim que os vértices das arestas a_2, \dots, a_{q-3} estejam ligados a exatamente mais duas novas arestas, portanto estes vértices são do tipo 3. Como entre dois vértices consecutivos de L_0 são colocados $p - 4$ pontos, os vértices das arestas a_1 e a_{q-2} também são do tipo 3. Assim, temos $p(q - 2)$ pontos do tipo 3 em L_1 . O total de pontos em L_1 é

$$N_1 = p(p - 4) + p(p - 3)(q - 3) + p(q - 2). \quad (4.3)$$

Ainda não é possível generalizar a contagem, pois em L_1 temos pontos do tipo 2 e do tipo 3. Procedemos da seguinte forma. Em L_1 temos $p[(p - 3)(q - 3) + (p - 4)]$ pontos do tipo 2 e $p(q - 2)$ pontos do tipo 3. Em um vértice do tipo 2, inserimos como antes, $q - 2$ novas arestas, e em um vértice do tipo 3 inserimos $q - 3$ arestas. Como no primeiro nível, essas novas arestas inseridas nos vértices de L_1 , originarão

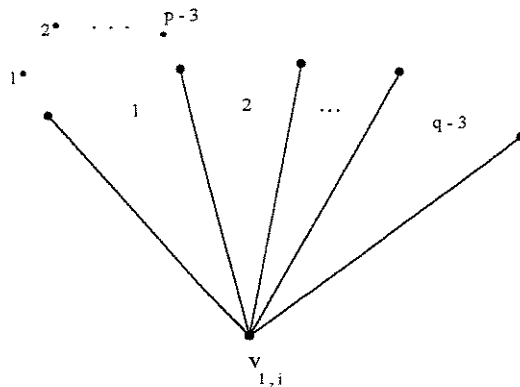


Figura 4.6: Em cada par de arestas inseridas em pontos do tipo 2 de L_1 inserimos novos $p - 3$ pontos.

vértices do tipo 3 num total de $p[(p-3)(q-3) + (p-4)](q-2) + p(q-2)(q-3)$. Por outro lado, seja $v_{1,i}$ um vértice do tipo 2 (Figura 4.6), onde inserimos $q - 2$ arestas, gerando $q - 3$ pares de arestas, que são partes dos novos polígonos.

Em cada par de arestas precisamos acrescentar $p - 3$ novos pontos para obtermos o polígono completo. Dessa forma, temos que $(p - 3)(q - 3)$ é o número total de pontos gerados pelos pares de arestas, sendo estes pontos do tipo 2. Em todo L_1 os $p[(p-3)(q-3) + (p-4)]$ pontos do tipo 2 geram $p[(p-3)(q-3) + (p-4)](p-3)(q-3)$ pontos do tipo 2.

Seja $v_{1,j}$ um vértice do tipo 3 em L_1 . Neste vértice inserimos $q - 3$ arestas, que geram $q - 4$ pares de arestas. Em cada par de arestas colocamos $p - 3$ pontos, resultando em $(q - 4)(p - 3)$ novos pontos do tipo 2 no vértice $v_{1,j}$ do tipo 3. Consequentemente, conduzindo a um total de $p(q - 2)(q - 4)(p - 3)$ pontos obtidos desta forma em todo L_1 .

Entre dois vértices consecutivos $v_{1,i}$ e $v_{1,i+1}$ de L_1 devemos inserir $p - 4$ pontos, e como L_1 tem N_1 pontos, temos N_1 pares de vértices, dando um total de $N_1(p - 4)$ pontos do tipo 2 (veja Figura 4.4). Portanto, o total de pontos N_2 em L_1 é dado pela soma dos seguintes termos:

$$[p(p - 4) + p(p - 3)(q - 3)](p - 3)(q - 3) \quad \text{tipo 2}$$

$$p(q - 2)(q - 4)(p - 3) \quad \text{tipo 2}$$

$$[p(p - 4) + p(p - 3)(q - 3) + (q - 2)](p - 4) \quad \text{tipo 2}$$

$$[p(p - 4) + p(p - 3)(q - 3)](q - 2) + p(q - 2)(q - 3) \quad \text{tipo 3.}$$

Agora, o total de pontos nos níveis superiores pode ser generalizado da seguinte forma:

Teorema 4.1 *Se no nível L_k temos $M_{2,k}$ pontos do tipo 2 e $M_{3,k}$ pontos do tipo 3, tendo no total $N_k = M_2 + M_3$ pontos, então em L_{k+1} temos a seguinte distribuição de pontos:*

$$M_{2,k}(q - 2) + M_{3,k}(q - 3) \quad \text{tipo 3}$$

$$M_{2,k}(q - 3)(p - 3) \quad \text{tipo 2}$$

$$M_{3,k}(q - 4)(p - 3) \quad \text{tipo 2}$$

$$(M_{2,k} + M_{3,k})(p - 4) \quad \text{tipo 2}$$

e N_{k+1} é a soma desses termos.

Utilizando os cálculos realizados acima, podemos agora determinar o número de polígonos que existe em cada nível, o que nos fornece diretamente o número de centros de polígonos. Seja NP_k o número de polígonos existentes no nível k . Logo, $NP_0 = 1$. Em $\{p, q\}$ a partir de L_0 colocamos $q - 2$ arestas em cada vértice, essas arestas geram $q - 3$ pares de arestas, logo temos $q - 3$ polígonos, num total de $p(q - 3)$. Seja

$v_{0,1}, \dots, v_{0,p}$ os vértices de L_0 . Entre $v_{0,i}$ e $v_{0,i+1}$ colocamos $p - 4$ pontos formando um novo polígono, e como temos p pares desses vértices em L_0 , obtemos p novos polígonos. O número total de polígonos em L_1 é

$$NP_1 = p + p(q - 3). \quad (4.4)$$

Em L_1 temos N_1 vértices, logo N_1 pares de vértices, sendo que cada par origina um novo polígono. Através da expressão (4.3) de N_1 sabemos a quantidade de pontos do tipo 2. Em um ponto deste tipo inserimos $q - 2$ arestas, logo $q - 3$ pares de arestas e então $q - 3$ polígonos, tendo um total de $p(p - 3)(q - 3)(q - 3)$ polígonos originados pelos pontos do tipo 2. Em cada ponto do tipo 3 inserimos $q - 3$ arestas, tendo assim $q - 4$ pares de arestas, originando $q - 4$ polígonos. Assim, o total de polígonos originados pelos pontos do tipo 3 é $p(q - 2)(q - 4)$. Podemos agora concluir que o número total de polígonos em L_2 é

$$NP_2 = N_1 + [p(p - 3)(q - 3)](q - 3) + p(q - 2)(q - 4).$$

De forma geral temos,

Teorema 4.2 *Se o nível k tem $M_{2,k}$ pontos do tipo 2 e $M_{3,k}$ pontos do tipo 3 com N_k sendo o total de seus pontos, então em L_{k+1} temos que*

$$NP_{k+1} = N_k + M_{2,k}(q - 3) + M_{3,k}(q - 4)$$

é o total de centros no novo nível.

Exemplo. A título de exemplo vamos calcular o número de pontos dos três primeiros níveis da tesselação $\{5, 4\}$ (Figura 4.7). Como $N_0 = p$ e temos $p = 5$ e $q = 4$, obtemos que $N_0 = 5$. Da expressão (4.3) obtemos que $N_1 = 25$. Para N_2 temos

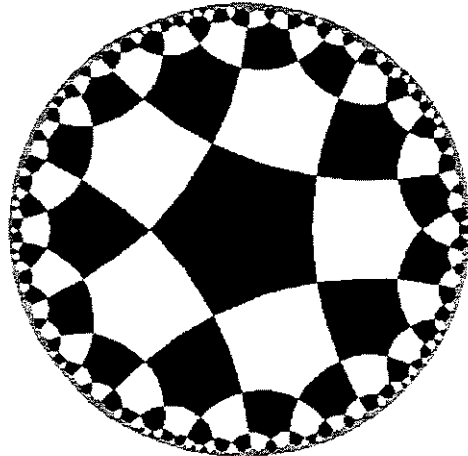


Figura 4.7: Primeiros três níveis da tesselação $\{4, 4\}$ no plano hiperbólico.

$M_{2,2} = 30 + 0 + 17 = 47$ e $M_{2,3} = 30 + 10 = 40$ e temos o total para o segundo nível dado por $N_2 = M_{2,2} + M_{2,3} = 87$. Como temos a distribuição de pontos do segundo nível, usamos o Teorema 4.1 para o número de pontos do terceiro nível. Assim, para L_3 temos $M_{3,2} = 94 + 0 + 87 = 181$ e $M_{3,3} = 94 + 40 = 134$, e o total é $N_3 = M_{3,2} + M_{3,3} = 181 + 134 = 315$. O número de pontos dos próximos níveis são obtidos da mesma forma, se utilizando o Teorema 4.1. Os próximos resultados são utilizados de forma semelhante.

4.2.2 CASO 2: $p = 3$.

Neste caso temos tesselações triangulares e por (4.1) e (4.2) temos que $q \geq 6$. Seja $v_{0,1}, v_{0,2}$ e $v_{0,3}$ os três vértices do triângulo do nível zero. Em $v_{0,i}$ inserimos $q - 2$ arestas, logo obtendo $q - 2$ novos vértices, denotados por $a_{i,1}, \dots, a_{i,q-2}$. Entre dois vértices consecutivos $v_{0,i}$ e $v_{0,i+1}$ não colocamos nenhum ponto, pois a aresta $a_{i,q-2}$ se

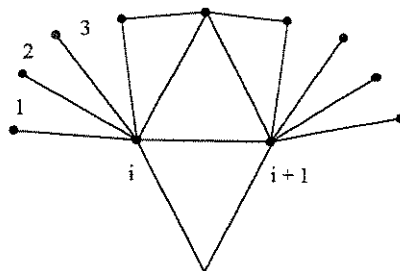


Figura 4.8: Em cada vértice de L_0 inserimos $q - 2$ arestas. Entre dois vértices consecutivos não são inseridos novos pontos.

junta à aresta $a_{i+1,1}$ (Figura 4.8). Entre os pares de arestas gerados pelas $q - 2$ arestas inseridas em $v_{0,i}$ também não aparecem novos pontos, pois o vértice da aresta $a_{i,k}$ está ligado ao vértice da aresta $a_{i,k+1}$ por uma nova aresta. Em particular temos que o vértice de $a_{i,q-3}$ está ligado ao vértice de $a_{i,q-2}$, o mesmo acontece com $a_{i+1,1}$ e $a_{i+1,2}$, implicando que cada par $v_{0,i}, v_{0,i+1}$ gera um ponto do tipo 4. Com isso, temos em L_1 três pontos do tipo 4. Das $q - 2$ arestas inseridas em $v_{0,i}$ sabemos que a primeira e a última geram pontos do tipo 4, logo sobram $q - 4$ arestas que, como observado anteriormente possuem seus vértices ligados, nos dando $q - 4$ pontos do tipo 3 (Figura 4.8).

Assim o número total de vértices em L_1 é

$$N_1 = p + p(q - 4). \quad (4.5)$$

Agora seja $v_{1,i}$ um ponto qualquer de L_1 . Dos pontos de L_1 partem as arestas que determinarão L_2 . Se $v_{1,i}$ é do tipo 3, inserimos $q - 3$ arestas, onde destas duas (a primeira e a última) originarão pontos do tipo 4 (entre $v_{1,i-1}, v_{1,i}$ e $v_{1,i+1}$), logo sobram $q - 5$ arestas, o que nos dá $p(q - 4)(q - 5)$ pontos do tipo 3 em L_2 . Se $v_{1,i}$

é do tipo 4, inserimos $q - 4$ arestas e seguindo o mesmo raciocínio do caso anterior, obtemos que os vértices do tipo 4 nos fornecem $p(q - 6)$ novos pontos do tipo 3. Entre $v_{1,i}$ e $v_{1,i+1}$, temos um ponto do tipo 4 originado pelas duas últimas arestas inseridas em $v_{1,i}$, e pelas duas primeiras arestas inseridas em $v_{1,i+1}$, temos assim N_1 pontos do tipo 4 em L_2 e o número de vértices é

$$N_2 = p + p(q - 4) + p(q - 4)(q - 5) + p(q - 6).$$

De forma geral

Teorema 4.3 *Se no nível L_k temos $M_{3,k}$ pontos do tipo 3 e $M_{4,k}$ do tipo 4, então em L_{k+1} temos*

$$(M_{3,k} + M_{4,k}) \quad \text{tipo 4}$$

$$M_{3,k}(q - 5) \quad \text{tipo 3}$$

$$M_{4,k}(q - 6) \quad \text{tipo 3}$$

pontos e N_{k+1} é a soma desses termos.

A figura 4.9 mostra a tesselação $\{3, 12\}$ no modelo do disco.

Mantendo as notações, queremos agora o número de polígonos. Em cada aresta do triângulo L_0 são inseridas $q - 2$ arestas, originando $q - 3$ pares de arestas, logo originando $q - 3$ novos triângulos num total de $p(q - 3)$. Entre $v_{0,i}$ e $v_{0,i+1}$ temos um novo triângulo, assim o total em L_1 é

$$NP_1 = p + p(q - 3).$$

Agora em L_1 temos p pontos do tipo 4 e $p(q - 4)$ pontos do tipo 3, gerando um total de $p + p(q - 4)$ pares de pontos, logo temos $p + p(q - 4)$ triângulos. Analisando a inserção de novas arestas em L_1 como feito anteriormente, observamos que os pontos

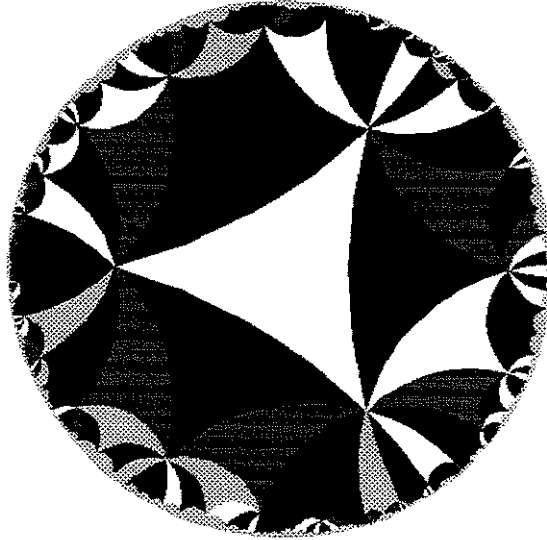


Figura 4.9: Tesselação $\{3, 12\}$.

do tipo 3 originarão $p(q-4)(q-4)$ novos triângulos, e os pontos do tipo 4 originarão $p(q-5)$ novos triângulos. O número de polígonos total em L_2 é

$$NP_2 = [p + p(q-4)] + p(q-4)(q-4) + p(q-5).$$

De forma geral

Teorema 4.4 *Se no nível k temos $M_{3,k}$ pontos do tipo 3 e $M_{4,k}$ pontos do tipo 4 com N_k sendo o total de seus pontos, então*

$$NP_{k+1} = N_k + M_{3,k}(q-4) + M_{4,k}(q-5)$$

é total de polígonos no novo nível.

4.2.3 CASO 3: $p \geq 6, q = 3$.

Esse é o caso mais complexo, e a soma dos pontos aparece de forma mais dividida que nos casos acima. Como $q = 3$, em cada vértice de L_0 inserimos somente uma

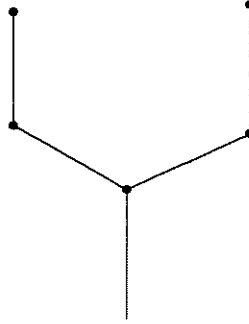


Figura 4.10: Em um ponto do tipo 3 não são inseridas novas arestas, mas a cada um desses pontos está relacionado um conjunto de $p - 5$ pontos do próximo nível.

aresta, e logo temos p novas arestas e também p pares destas arestas. Em cada um desses pares colocamos $p - 4$ pontos que são do tipo 2, pois são os vértices do novo polígono que aparece entre cada par. Agora, no vértice $v_{0,i}$ de L_0 temos a aresta $a_{0,i}$, e o novo vértice gerado por esta aresta é do tipo 3 pois recebe mais duas arestas, uma proveniente dos pontos colocados entre $v_{0,i-1}$ e $v_{0,i}$, e outra dos pontos entre $v_{0,i}$ e $v_{0,i+1}$, logo esses p pontos são todos do tipo 3. O número total de pontos é:

$$N_1 = p + p(p - 4).$$

Passemos para L_2 . Os p pontos do tipo 3 de L_1 não contribuem com novas arestas. Pela construção de L_1 , se $v_{1,i}$ é do tipo 3 então os vértices $v_{1,i-1}$ e $v_{1,i+1}$ são do tipo 2, e entre as arestas $a_{1,i-1}$ e $a_{1,i+1}$ colocamos $p - 5$ pontos, fornecendo um total de $p(p - 5)$ pontos do tipo 2. Assim, podemos dizer que a cada ponto do tipo 3 está associado um conjunto de $p - 5$ pontos no próximo nível (Figura 4.10).

Em L_1 temos $p(p - 4)$ pontos do tipo 2, ou seja, p conjuntos de $p - 4$ pontos do tipo 2 e em cada um desses pontos inserimos uma aresta. Sejam $a_{1,1}, \dots, a_{1,p-4}$ as

arestas de um desses conjuntos. Logo temos $p - 5$ pares de arestas, e em cada par colocamos novos $p - 4$ pontos, portanto um conjunto de $p - 4$ vértices (arestas) gera $(p - 5)(p - 4)$ pontos do tipo 2, e temos um total de $p(p - 5)(p - 4)$ pontos. Com isso concluímos que o total de pontos em L_2 é

$$\begin{array}{ll} p(p - 4) & \text{tipo 3} \\ p(p - 5) & \text{tipo 2} \\ p(p - 5)(p - 4) & \text{tipo 2} \end{array}$$

e N_2 é a soma desses termos.

Diferentemente dos outros dois casos, aqui precisamos analisar L_3 para chegarmos a uma expressão genérica. Para a contagem dos novos pontos em L_3 e nos níveis acima, necessitamos diferenciar os pontos do tipo 2. Temos aqueles que estão em conjuntos de $(p - 4)$ pontos, e aqueles que estão em conjuntos de $(p - 5)$ pontos. Em L_2 temos $p(p - 4)$ pontos do tipo 3, e a cada ponto do tipo 3, como feito acima, associamos a um conjunto de $p - 5$ pontos do tipo 2, e portanto temos $p(p - 4)(p - 5)$ pontos do tipo 2. Os pontos do tipo 2 de L_2 são os novos pontos do tipo 3 de L_3 , num total de $p(p - 5) + p(p - 5)(p - 4)$ pontos.

Consideremos os p conjuntos de $p - 5$ pontos em L_2 . Temos então $p - 6$ pares de pontos e em cada par inserimos $p - 4$ pontos, assim cada conjunto gera $(p - 6)(p - 4)$ pontos, num total de $p(p - 6)(p - 4)$ pontos ou $p(p - 6)$ novos conjuntos com $p - 4$ pontos. Por último temos $p(p - 5)$ conjuntos de $(p - 4)$ pontos em L_2 , e como acima, cada conjunto gera $(p - 5)(p - 4)$ novos pontos, nos fornecendo um total de $[p(p - 5)](p - 5)(p - 4) = p(p - 5)^2(p - 4)$ pontos do tipo 2, originados dos conjuntos

com $p - 4$ pontos. Assim o total de pontos N_3 em L_3 é a soma dos termos

$$p(p - 5) + p(p - 5)(p - 4) \quad \text{tipo 3}$$

$$[p(p - 4)](p - 5) \quad \text{tipo 2}$$

$$[p(p - 6) + p(p - 5)^2](p - 4) \quad \text{tipo 2.}$$

Assim de forma geral suponha que no nível k temos $N_k = M_{2,k} + M_{3,k}$ onde $M_{2,k}$ é o total de pontos do tipo 2, e $M_{3,k}$ o de pontos do tipo 3. Além disso temos que $M_{2,k} = M'_{2,k} + M''_{2,k}$ onde $M'_{2,k} = k_1(p - 5)$ são os pontos do tipo 2 em conjuntos de $(p - 5)$ pontos, e $M''_{2,k} = k_2(p - 4)$ são os pontos do tipo 2 em conjuntos de $(p - 4)$ pontos. Portanto, em L_{k+1} os pontos do tipo 3 geram conjuntos de $p - 5$ pontos, num total de $M_{3,k}(p - 5)$ pontos do tipo 2. O número de pontos do tipo 3 em L_{k+1} é o mesmo do de números do tipo 2 de L_k , ou seja $M'_{2,k} + M''_{2,k}$. Temos k_1 conjuntos de $p - 5$ pontos, que nos dá $k_1(p - 6)(p - 4)$ pontos. Temos k_2 conjuntos de $p - 4$ pontos, dando $k_2(p - 5)(p - 4)$ novos pontos. De forma geral:

Teorema 4.5 *Se no nível L_k temos $N_k = M_{2,k} + M_{3,k}$ pontos, onde $M_{2,k}$ é o número de pontos do tipo 2 em L_k e $M_{3,k}$ o número de pontos do tipo 3, e além disso $M_{2,k} = M'_{2,k} + M''_{2,k}$ onde $M'_{2,k} = k_1(p - 5)$ e $M''_{2,k} = k_2(p - 4)$, então a distribuição de pontos em L_{k+1} é dada por*

$$M'_{2,k} + M''_{2,k} \quad \text{tipo 3}$$

$$M_{3,k}(p - 5) \quad \text{tipo 2}$$

$$k_1(p - 6)(p - 4) + k_2(p - 5)(p - 4) \quad \text{tipo 2}$$

onde N_{k+1} é a soma desses termos.

Novamente, para a contagem dos polígonos em cada nível, mantemos a notação utilizada acima. Como $q = 3$, de cada um dos p vértices de L_0 inserimos somente

uma aresta, originando assim p pares de arestas, o que nos dá

$$NP_1 = p.$$

Em L_1 temos p pontos do tipo 3 e $p(p-4)$ pontos do tipo 2. A cada ponto do tipo 3 associamos um conjunto de $p-3$ pontos, e temos assim um polígono originado por cada ponto deste tipo, totalizando p polígonos. Para os pontos do tipo 2, temos p conjuntos de $p-4$ pontos. Em cada um desses conjuntos inserimos $p-4$ arestas originando $p-5$ pares de arestas, logo $p-5$ polígonos e no total $p(p-5)$ polígonos originados desta forma. Portanto

$$NP_2 = p + p(p-5).$$

Seguindo esse raciocínio e verificando a distribuição de pontos dada por N_2 obtemos

$$NP_3 = p(p-4) + p(p-6) + p(p-5)(p-5).$$

De forma geral,

Teorema 4.6 *Se no nível k temos $M_{3,k}$ pontos do tipo 3 e $M_{2,k} = k_1(p-5) + k_2(p-4)$ pontos do tipo 2, obtemos que*

$$NP_{k+1} = M_{3,k} + k_1(p-6) + k_2(p-5)$$

é o total de polígonos em L_{k+1} .

A figura 4.11 mostra a tesselação $\{12, 3\}$ no modelo do disco.

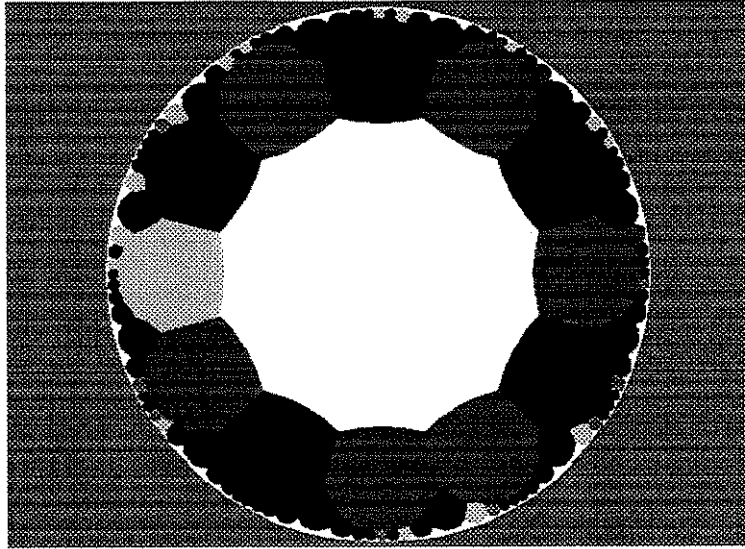


Figura 4.11: Tesselação $\{12, 3\}$.

4.3 Subconjuntos Originados por q Arestas

Nesta seção apresentamos uma outra forma ordenada de obter subconjuntos finitos de uma tesselação.

Seja $\{p, q\}$ uma tesselação qualquer fixada. Nesta tesselação fixamos um vértice v_0 e denotamos por $l_1 \dots l_q$ as q arestas que se interceptam em v_0 .

O primeiro nível do subconjunto que queremos, é formado pela união dos q polígonos de p lados que se interceptam em v_0 , excluindo v_0 . Assim no primeiro nível temos q arestas partindo de v_0 , logo q pares de arestas. Cada par origina um novo polígono, e temos assim q polígonos ou q centros, portanto

$$NP_1 = q.$$

Novamente, temos q arestas a partir de v_0 , logo q pontos a partir de v_0 . Cada par de arestas deve originar um novo polígono. Para isto, inserimos $p - 3$ pontos entre

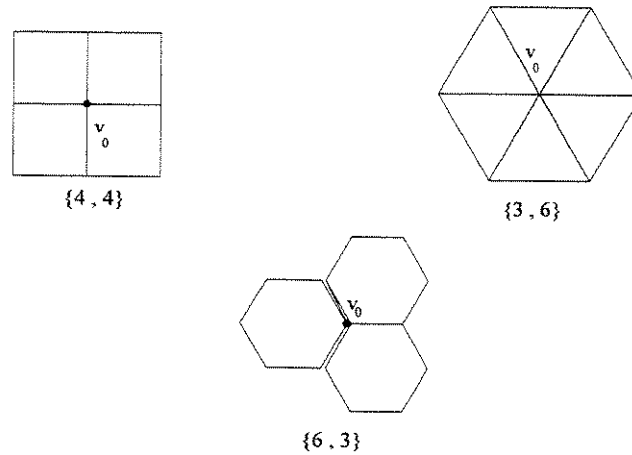


Figura 4.12: Alguns exemplos da região inicial de subconjuntos de tesselações, originados pelo método das arestas.

cada par de arestas. Esses são pontos do tipo 2, então os q pontos originados a partir das arestas em v_0 são do tipo 3. O número total de pontos é

$$N_1 = q + q(p - 3).$$

A Figura 4.12 mostra o primeiro nível obtido a partir das três tesselações euclidianas.

A seguir apresentamos as soluções gerais para os níveis superiores, seguindo a mesma divisão de casos feitos nas seções anteriores.

4.3.1 CASO 1: $p \geq 4$, $q \geq 4$.

Em cada ponto do tipo 2 colocamos $q - 2$ arestas, obtendo $q(p - 3)(q - 2)$ novos pontos do tipo 3. Em cada um dos $q(p - 3)$ conjuntos de $q - 2$ arestas temos $q - 3$ pares de arestas. Em cada par colocamos $p - 3$ novos pontos do tipo 2, num total de $[q(p - 3)](q - 3)(p - 3)$ pontos do tipo 2.

Cada ponto do tipo 2 de L_1 gera $q - 3$ pares, logo $q - 3$ polígonos, o que nos dá um total de $q(p - 3)(q - 3)$ polígonos. Em cada ponto do tipo 3 de L_1 inserimos $q - 3$ arestas, originando $q(q - 3)$ pontos do tipo 3. Em cada conjunto de $q - 3$ arestas temos $q - 4$ pares de arestas, em cada par adicionamos $p - 3$ pontos, tendo assim $q(q - 4)(p - 3)$ novos pontos do tipo 2. Como cada ponto do tipo 3 está associado a $q - 4$ pares, temos $q - 4$ polígonos, num total de $q(q - 4)$ polígonos.

Em L_1 temos N_1 vértices, logo N_1 pares de pontos e como $p \geq 4$, entre os vértices $v_{1,i}$ e $v_{1,i+1}$ colocamos $p - 4$ pontos, obtendo um total de $N_1(p - 4)$ pontos do tipo 2. Cada par $v_{1,i}v_{1,i+1}$ gera um polígono, temos assim $q + q(p - 3)$ novos polígonos. Portanto em L_2 temos

$$\begin{array}{ll} q(p - 3)(q - 2) + q(q - 3) & \text{tipo 3} \\ q(p - 3)(q - 3)(p - 3) & \text{tipo 2} \\ q(q - 4)(p - 3) & \text{tipo 2} \\ [q + q(p - 3)](p - 4) & \text{tipo 2} \end{array}$$

e N_2 é a soma destes termos. Para o número de polígonos temos

$$NP_2 = q(p - 3)(q - 3) + q(q - 4) + [q + q(p - 3)].$$

Assim de forma geral

Teorema 4.7 *Suponhamos que no nível L_k temos $M_{2,k}$ pontos do tipo 2 e $M_{3,k}$*

pontos do tipo 3. Então os pontos em L_{k+1} estão distribuídos da seguinte forma

$$M_{2,k}(q-2) + M_{3,k}(q-3) \quad \text{tipo 3}$$

$$M_{2,k}(q-3)(p-3) \quad \text{tipo 2}$$

$$M_{3,k}(q-4)(p-3) \quad \text{tipo 2}$$

$$(M_{2,k} + M_{3,k})(p-4) \quad \text{tipo 2}$$

e N_{k+1} é dado pela soma desses termos. O total de polígonos em L_{k+1} é dado por

$$NP_{k+1} = (M_{2,k} + M_{3,k}) + M_{2,k}(q-3) + M_{3,k}(q-4).$$

4.3.2 CASO 2: $p = 3$.

Como $p = 3$ temos que $N_1 = q \geq 6$ e $NP_1 = q$ em L_1 , e os q pontos são do tipo 3. Em cada um desses pontos acrescentamos $q-3$ arestas, originando $q-4$ pares de arestas, logo $q-4$ novos polígonos tendo um total de $q(q-4)$ polígonos. Como $p = 3$, entre dois vértices consecutivos não acrescentamos novos pontos, e como vimos no caso equivalente da seção anterior esses vértices originam pontos do tipo 4, assim temos q pontos do tipo 4 e também mais q polígonos.

Em cada ponto de L_1 são acrescentados $q-3$ arestas a_1, \dots, a_{q-3} , mas as arestas a_1 e a_{q-3} dão origem a pontos do tipo 4 de L_2 , logo sobram $q-5$ arestas num total parcial de $q(q-5)$ pontos do tipo 3. O total é

$$N_2 = q + q(q-5),$$

onde q é o número de pontos do tipo 4 e $q(q-5)$ do tipo 3. O número de polígonos é dado por

$$NP_2 = q + q(q-4).$$

Analisemos agora o terceiro nível. Temos em L_2 , N_2 pontos ou N_2 pares de pontos, logo no caso anterior, temos N_2 novos pontos do tipo 4 e também N_2 novos polígonos. Em pontos do tipo 3 inserimos $q - 3$ arestas e como vimos, duas são partes de pontos do tipo 4 de L_3 sobrando $q - 5$, dando assim um total parcial de $q(q - 5)(q - 5)$ pontos do tipo 3 e $q(q - 5)(q - 4)$ polígonos. Em um ponto do tipo 4 de L_2 inserimos $q - 4$ arestas, mas de novo duas já foram contadas para os pontos do tipo 4, assim sobram $q - 6$ arestas, tendo um total parcial de $q(q - 6)$ pontos do tipo 3 e $q(q - 5)$ polígonos. Portanto temos que N_3 é a soma dos seguintes termos

$$q + q(q - 5) \quad \text{tipo 4}$$

$$q(q - 5)(q - 5) \quad \text{tipo 3}$$

$$q(q - 6) \quad \text{tipo 3}$$

e

$$NP_3 = [q + q(q - 5)] + q(q - 5)(q - 4) + q(q - 5).$$

Assim de forma geral

Teorema 4.8 *Se no nível L_k , $k \geq 2$ temos $M_{3,k}$ pontos do tipo 3 e $M_{4,k}$ pontos do tipo 4, então em L_{k+1} obtemos que o número de pontos tem a distribuição*

$$M_{3,k} + M_{4,k} \quad \text{tipo 4}$$

$$M_{3,k}(q - 5) \quad \text{tipo 3}$$

$$M_{4,k}(q - 6) \quad \text{tipo 3}$$

e N_{k+1} é dado pela soma dos termos. O número de polígonos é dado por

$$NP_{k+1} = (M_3 + M_4) + M_3(q - 4) + M_4(q - 5).$$

4.3.3 CASO 3: $p \geq 6$, $q = 3$.

Temos $q = 3$, logo $NP_1 = 3$, temos 3 pares de arestas onde em cada par precisamos de $p - 3$ pontos para completar um polígono, assim em cada par temos $p - 3$ pontos do tipo 2, nos dando um total parcial de $3(p - 3)$. Como $p \geq 6$, as 3 arestas iniciais dão origem a três pontos do tipo 4, e o total é

$$N_1 = 3 + 3(p - 3)$$

onde temos 3 pontos do tipo 3 e $3(p - 3)$ pontos do tipo 2. Para obtermos o segundo nível, cada ponto do tipo 2 de L_1 origina um ponto do tipo 3 em L_2 , logo temos $3(p - 3)$ pontos do tipo 3 em L_2 .

Cada par de pontos do tipo 2 de L_1 gera um par de arestas, onde inserimos $p - 4$ pontos do tipo 2 e cada par gera um novo polígono, mas quantos pares de arestas temos? Para respondermos a esta pergunta, usaremos um raciocínio um pouco diferente daquele que estamos utilizando até o momento.

Temos N_1 pontos em L_1 . Seja $v_{1,i}$ um vértice qualquer fixado de L_1 , escolhemos um sentido (horário ou anti-horário) para caminhar de $v_{1,i}$ para os próximos vértices. Assim podemos imaginar que esses N_1 pontos estão distribuídos ao redor de um círculo (Figura 4.13), e queremos determinar quando dois pontos consecutivos podem ser ligados por arestas originadas destes.

Como $p \geq 6$ e $q = 3$, não podemos ter dois pontos consecutivos do tipo 3. A princípio temos um máximo de N_1 pares de pontos. Se $v_{1,i}$ é do tipo 3, os pontos $v_{1,i-1}$ e $v_{1,i+1}$ são do tipo 2 e não originarão arestas que se conectam com $v_{1,i}$, logo dois pares devem ser excluídos para cada ponto do tipo 3; no total são 6 pares de pontos.

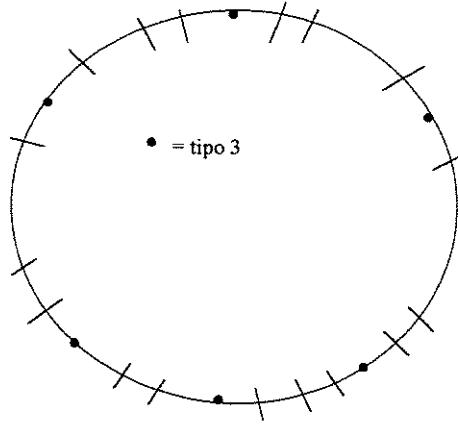


Figura 4.13: Um artifício para poder contar os pontos do tipo 2 e do tipo 3 neste caso.

Assim sobram $N_1 - 6 = 3(p - 3) - 3$ pares de arestas e em cada par acrescentamos $p - 4$ pontos, num total parcial de $[3(p - 3) - 3](p - 4)$ novos pontos do tipo 2 e também $3(p - 3) - 3$ novos polígonos. Cada ponto do tipo 3 de L_1 origina um conjunto de 5 pontos e 4 arestas de um novo polígono de L_2 , logo devemos acrescentar $p - 5$ pontos num total de $3(p - 5)$ novos pontos do tipo 2, e 3 novos polígonos. Assim em L_2 temos a seguinte distribuição de pontos

$$\begin{array}{ll} 3(p - 3) & \text{tipo 3} \\ [3(p - 3) - 3](p - 4) & \text{tipo 2} \\ 3(p - 5) & \text{tipo 2} \end{array}$$

o que determina N_2 . O número de polígonos é dado por

$$NP_2 = [3(p - 3) - 3] + 3.$$

Portanto de forma geral

Teorema 4.9 Se no nível k temos $M_{3,k}$ pontos do tipo 3 e $M_{2,k} = M'_{2,k} + M''_{2,k}$, onde

$M'_{2,k}$ é o número de pontos do tipo 2 em conjuntos de $(p-4)$ pontos e $M''_{2,k}$ o número em conjuntos de $(p-5)$ pontos, então temos

$$M_{2,k} \quad \text{tipo 3}$$

$$\frac{M'_{2,k}}{p-4}(p-5)(p-4) \quad \text{tipo 2}$$

$$\frac{M''_{2,k}}{p-5}(p-6)(p-5) \quad \text{tipo 2}$$

e

$$NP_{k+1} = M_{3,k} + \frac{M'_{2,k}}{p-4}(p-5) + \frac{M''_{2,k}}{p-5}(p-6).$$

Observemos por exemplo, que em $\frac{M'_{2,k}}{p-4}(p-5)(p-4)$ temos que $\frac{M'_{2,k}}{p-4}$ é o número de conjuntos com $p-4$ pontos, $p-5$ é o número de pares de pontos e $p-4$ o número de pontos colocados em cada par. Os termos no enunciado do Teorema 4.9 não estão simplificados para explicitar a forma de distribuição dos pontos.

Capítulo 5

Receptor de Máxima

Verossimilhança e Análise de

Desempenho

Neste capítulo apresentamos o receptor de máxima verossimilhança para um sistema de modulação no plano hiperbólico. Veremos que este receptor apresenta semelhanças com o do caso euclidiano e não é muito mais complexo.

Também realizamos uma análise comparativa em relação ao desempenho, entre algumas constelações de sinais euclidianas com equivalentes hiperbólicas. Além da importância por si própria destas análises, estes exemplos mostram o processo de construção das constelações e os resultados do Capítulo 4.

5.1 Constelações de Sinais e Receptor PAM Hiperbólico

Constelações do tipo PAM são conjuntos discretos de pontos fixados na reta euclidiana \mathbb{R} . Uma constelação PAM é dita simétrica se a distância entre quaisquer dois pontos consecutivos é sempre igual. Agora, se considerarmos conjuntos discretos na reta \mathbb{H} , temos o PAM hiperbólico. O objetivo aqui é determinar o desempenho do sistema de comunicações, fazendo uso do critério da probabilidade de erro associada e realizar comparações entre constelações semelhantes construídas em cada uma das retas.

Primeiramente consideremos uma constelação PAM simétrica em \mathbb{R} dada pelos pontos $\{x_1, \dots, x_M\}$ tal que $d_e(x_i, x_{i+1}) = k$. Como existe uma isometria entre \mathbb{R} e \mathbb{H} , dada pela exponencial $r(X) = e^x$, esta dará origem em \mathbb{H} a uma constelação $\{y_1, \dots, y_M\}$, onde $y_i = e^{x_i}$ e $d_h(y_i, y_{i+1}) = |\ln(y_{i+1}/y_i)| =$

$$\begin{aligned} d_h(y_i, y_{i+1}) &= |\ln(y_{i+1}/y_i)| \\ &= |\ln(e^{x_{i+1}}/e^{x_i})| \\ &= |x_{i+1} - x_i| \\ &= k. \end{aligned}$$

Logo temos que uma constelação PAM simétrica em \mathbb{R} origina uma constelação PAM simétrica em \mathbb{H} , ambas com a mesma distância entre pontos consecutivos. Mas isso não é um bom resultado, pois como as distâncias são as mesmas e uma gaussiana em \mathbb{H} é imagem por r de uma gaussianana Euclidiana, não existe nenhuma diferença a nível de desempenho, em se utilizar a constelação em \mathbb{R} ou em \mathbb{H} .



Figura 5.1: Constelação 2-PAM em H .

Então consideremos uma situação diferente. Seja $\{y_1, \dots, y_M\}$ uma constelação em \mathbb{H} . Como conjunto de pontos temos que $\mathbb{H} \subset \mathbb{R}$, logo uma constelação de \mathbb{H} é também uma constelação de \mathbb{R} . Assim, sem utilizar a isometria, estamos interessados em constelações que sejam assimétricas em \mathbb{R} , mas que preservem simetria em \mathbb{H} .

Primeiramente consideremos a constelação 2-PAM assimétrica em \mathbb{R} , dada pelos pontos e e e^{-1} . Esses pontos também são pontos de \mathbb{H} , e com a distância introduzida no Capítulo 2, temos que essa constelação é simétrica em \mathbb{H} , pois

$$d_h(e^{-1}, 1) = d_h(e, 1) = 1.$$

O próximo passo é determinar a regra de decisão a ser utilizada no receptor ótimo. Para tal iremos considerar que a constelação de sinais seja do tipo 2-PAM. Mais especificamente, iremos considerar que a constelação 2-PAM seja assimétrica no modelo euclidiano, como mostrado na Fig 5.1. A constelação 2-PAM no modelo hiperbólico é também mostrada na Fig. 5.1.

A seguir iremos assumir que os sinais 2-PAM hiperbólico são utilizados na transmissão da informação. Iremos assumir que o ruído no canal de transmissão seja uma amostra do ruído gaussiano em ambos os modelos. Com isso, de acordo com a hipótese de ruído hiperbólico introduzida no Capítulo 3, temos o seguinte teste de hipóteses a considerar

$$H_0 : y_0 = e^k x_0$$

$$H_1 : y_1 = e^k x_1$$

onde $x_0 = e^{-1}$ e $x_1 = e$ e e^k é uma amostra de um processo gaussiano hiperbólico com média zero e variância σ^2 .

Como é usual em sistemas de comunicações digitais, iremos assumir que o processo de demodulação é o de máxima verossimilhança.

Diante destes fatos, temos que o teste de hipóteses a ser considerado é reduzido a

$$H_0 \longleftrightarrow f_0(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{\ln^2(y/e^{-1})}{2\sigma^2} \right\}$$

$$H_1 \longleftrightarrow f_1(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{\ln^2(y/e)}{2\sigma^2} \right\}.$$

Então, uma decisão a favor da hipótese H_1 será feita se

$$\frac{f_1(y)}{f_0(y)} \geq \lambda \longrightarrow -\frac{(\ln y - 1)^2}{2\sigma^2} + \frac{(\ln y + 1)^2}{2\sigma^2} \geq \ln \lambda.$$

Após algumas manipulações algébricas chegamos a

$$\ln y \geq \frac{\sigma^2 \ln \lambda}{2}.$$

Se os sinais são igualmente prováveis temos que

$$\lambda = \frac{P(H_0)}{P(H_1)} = \frac{1/2}{1/2} = 1.$$

Portanto, $\ln \lambda = 0$ e conseqüentemente, o limiar de decisão será 1. Se $y \geq 1$ então decide-se pela hipótese H_1 , caso contrário pela hipótese H_0 . Dessa forma,

o receptor ótimo colherá uma amostra do sinal $y(t)$, $y(t = t_k) = y_k$ na saída do canal, e comparará com o valor do limiar, se $y_k \geq 1$ então decidirá pela hipótese H_1 , caso contrário pela hipótese H_0 . Este procedimento é o mesmo que será utilizado no receptor ótimo quando do modelo euclidiano, porém com valor de limiar zero.

Agora consideramos os pontos e e e^{-1} como uma constelação euclidiana, e usamos o mesmo limiar de decisão que o hiperbólico, $\lambda = 1$, assim em relação a este limiar temos uma constelação euclidiana assimétrica, e nesta constelação somamos -1 aos dois pontos de forma a ter uma constelação equivalente, mas com a menor energia média possível. Com isso temos no caso euclidiano o seguinte teste de hipótese

$$H_0 : \quad y(t) = x_0(t) + n(t)$$

$$H_1 : \quad y(t) = x_1(t) + n(t)$$

onde $x_0(t) = -1 + e^{-1}$ e $x_1(t) = -1 + e$ e $n(t)$ é uma amostra do processo gaussiano com média zero e variância σ^2 como no caso hiperbólico.

Com esses elementos em mãos, fica fácil mostrar que a probabilidade de erro média no modelo hiperbólico é dada por

$$P_e = P_{e,0} = P_{e,1} = Q\left(\frac{1}{\sigma}\right),$$

e a probabilidade de erro média no modelo euclidiano é dada por

$$P_e = \frac{1}{2}Q\left(\frac{1.71828}{\sigma}\right) + \frac{1}{2}Q\left(\frac{0.6322}{\sigma}\right),$$

onde

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx.$$

O ganho, G , é então dado por

$$G = 10 \log_{10} \left(\frac{1/\sigma^2}{(-0.6322)^2/\sigma^2} \right) = 10 \log_{10} \left(\frac{1}{0.39955} \right) = 3.98,$$

onde o valor $-0.6322 = e^{-1} - 1$ que aparece na expressão, é a distância do ponto da constelação que está mais próximo do limiar 0. Conseqüentemente, o 2-PAM simétrico hiperbólico é 3.98 dB melhor do que o 2-PAM assimétrico euclidiano correspondente. Chamamos a atenção para o fato de que a constelação 2-PAM hiperbólica foi derivada da constelação 2-PAM euclidiana.

Podemos realizar extensões para constelações M-PAM hiperbólicas. Por exemplo, consideramos o 4-PAM dado pelos pontos $\{e^{-3}, e^{-1}, e^1, e^3\}$. Temos que $d_h(e^{-3}, 1) = d_h(e^3, 1) = 3$ e $d_h(e^{-1}, 1) = d_h(e, 1) = 1$, o que mostra que estamos considerando uma constelação hiperbólica simétrica.

Assumimos agora que os sinais 4-PAM hiperbólico são utilizados na transmissão da informação. Assumimos que o ruído no canal de transmissão seja uma amostra do ruído gaussiano em ambos os modelos. Com isso, de acordo com a hipótese de ruído hiperbólico introduzida no Capítulo 3, temos o seguinte teste de hipóteses a considerar

$$H_0 : y_0 = e^k x_0$$

$$H_1 : y_1 = e^k x_1$$

$$H_2 : y_2 = e^k x_2$$

$$H_3 : y_3 = e^k x_3$$

onde $x_0 = e^{-3}$, $x_1 = e^{-1}$, $x_2 = e$, $x_3 = e^3$ e e^k é uma amostra de um processo gaussiano hiperbólico com média zero e variância σ^2 . Assumimos agora que o processo

de demodulação é o de máxima verossimilhança, e como feito para o 2-PAM, temos que o teste de hipóteses a ser considerado é reduzido a

$$H_j \longleftrightarrow f_j(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{\ln^2(y/e^{-3+2j})}{2\sigma^2} \right\}$$

onde $j = 0, 1, 2, 3$. Devemos agora obter os limiares de decisão entre as hipóteses H_0 e H_1 , H_1 e H_2 e H_2 e H_3 . Para o caso H_0 e H_1 , uma decisão a favor da hipótese H_1 será feita se

$$\frac{f_1(y)}{f_0(y)} \geq \lambda \longrightarrow \frac{(\ln y + 3)^2}{2\sigma^2} - \frac{(\ln y + 1)^2}{2\sigma^2} \geq \ln \lambda.$$

Após algumas manipulações algébricas chegamos a

$$4 \ln y + 8 \geq 2\sigma^2 \ln \lambda.$$

Se os sinais são igualmente prováveis temos que

$$\lambda = \frac{P(H_0)}{P(H_1)} = \frac{1/4}{1/4} = 1.$$

Portanto, $\ln \lambda = 0$, e conseqüentemente, o limiar de decisão l_1 será e^{-2} . Repetindo este procedimento obtemos que os outros dois limiares de decisão são $l_2 = 1$ e $l_3 = e^2$. Assim se $y \leq e^{-2}$ o demodulador decide pela hipótese H_0 , se $e^{-2} \leq y \leq 1$ decide pela hipótese H_1 e assim por diante. Portanto, o receptor ótimo colherá uma amostra do sinal $y(t)$, $y(t = t_k) = y_k$ na saída do canal, e comparará com o valor dos limiares, e tomará a decisão por uma das quatro hipóteses. Este procedimento é o mesmo que será utilizado no receptor ótimo quando do modelo euclidiano, porém com valor de limiar zero.

Considemos agora os pontos $\{e^{-3}, e^{-1}, e, e^3\}$ como uma constelação euclidiana. Fixando o mesmo limiar de decisão que o hiperbólico, $\lambda = 1$. Em relação a este limiar temos que a constelação euclidiana é assimétrica, e nesta constelação somamos -1 aos quatro pontos de forma a ter uma constelação equivalente, mas com a menor energia média possível. Para o caso euclidiano também existirão quatro hipóteses a serem consideradas. Pode-se mostrar que a probabilidade de erro média no modelo hiperbólico para o 4-PAM, conduz a $P_e \approx 2Q(1/\sigma)$, enquanto que a probabilidade de erro média para o modelo euclidiano para o 4-PAM correspondente, é dada por $P_e \approx Q(0.159/\sigma)$.

Portanto, o ganho G do 4-PAM hiperbólico é 15.97 dB melhor do que o correspondente 4-PAM aassimétrico euclidiano.

5.2 Receptor de Máxima Verossimilhança

Apresentamos a seguir o modelo que deve ser utilizado para o receptor de máxima verossimilhança para o plano hiperbólico.

Seja $\{s_j\}$, $1 \leq j \leq L$, uma constelação de sinais em \mathbb{H}^2 . Assumimos que o sinal recebido é dado pela forma

$$y = g(s_j)$$

para algum $1 \leq j \leq L$ e g é uma isometria hiperbólica que age como ruído gaussiano.

Consideremos os seguintes testes de hipóteses,

$$H_1 \leftrightarrow y = g(s_1)$$

$$H_2 \leftrightarrow y = g(s_2)$$

⋮

$$H_L \leftrightarrow y = g(s_L).$$

Assumindo que o canal é caracterizado por uma densidade de probabilidade gaussiana circular hiperbólica, e que os sinais são igualmente prováveis, então o critério de decisão a ser utilizado é o de máxima verossimilhança, isto é, o receptor decide pela hipótese H_j se

$$f_j(y) \geq f_k(y), \quad \forall k \neq j. \quad (5.1)$$

Do Capítulo 3, $f_j(y)$ é dada por

$$\begin{aligned} f_j(y) &= B \exp(-A d_h^2(y, s_j)) \\ &= B \exp(-A d_h^2(g(s_j), s_j)) \end{aligned} \quad (5.2)$$

Substituindo (5.2) em (5.1) chegamos a

$$d_h^2(g(s_j), s_j) \leq d_h^2(g(s_j), s_k).$$

Geometricamente, o receptor escolherá o sinal dentre os elementos da constelação $\{s_j, 1 \leq j \leq L\}$, aquele que minimiza $d_h(y, s_j)$. Mas minimizar $d_h(y, s_j)$ é o mesmo que minimizar $\cosh(d_h(y, s_j))$. Como

$$\cosh(d_h(z, w)) = 1 + \frac{|z - w|^2}{2 \operatorname{Im}(z) \operatorname{Im}(w)},$$

resulta que

$$\cosh(d_h(y, s_j)) = 1 + \frac{|y - s_j|^2}{2 \operatorname{Im}(y) \operatorname{Im}(s_j)}$$

$$\begin{aligned}
&= 1 + \frac{|y|^2 + |s_j|^2 - 2 \operatorname{Re}(\langle y, s_j \rangle)}{2 \operatorname{Im}(y) \operatorname{Im}(s_j)} \\
&= 1 + \frac{1}{\operatorname{Im}(y)} \left(\frac{|y|^2}{2 \operatorname{Im}(s_j)} + \frac{|s_j|^2}{2 \operatorname{Im}(s_j)} - \frac{\operatorname{Re}(\langle y, s_j \rangle)}{\operatorname{Im}(s_j)} \right).
\end{aligned}$$

onde $\langle \dots \rangle$ representa o produto interno euclidiano entre duas variáveis complexas.

Então minimizar $\cosh(d_h(y, s_j))$ é o mesmo que minimizar

$$\left\{ \frac{|y|^2}{2 \operatorname{Im}(s_j)} + \frac{|s_j|^2}{2 \operatorname{Im}(s_j)} - \frac{\operatorname{Re}(\langle y, s_j \rangle)}{\operatorname{Im}(s_j)} \right\},$$

que por sua vez é o mesmo que maximizar

$$\left\{ \frac{\operatorname{Re}(\langle y, s_j \rangle)}{\operatorname{Im}(s_j)} - \frac{|y|^2}{2 \operatorname{Im}(s_j)} - \frac{|s_j|^2}{2 \operatorname{Im}(s_j)} \right\}. \quad (5.3)$$

Os termos $|s_j|^2$ e $\operatorname{Im}(s_j)$ podem ser pré-calculados, e $|y|^2$ e $\operatorname{Im}(y)$ são conhecidos quando o sinal é recebido. Assim, o receptor calcula um conjunto de L variáveis de decisão $\operatorname{Re}(\langle y, s_j \rangle)$ e as utiliza para obter o máximo em (5.3). Aqui deve ser observado que utilizamos as expressões em \mathbb{H}^2 ao invés de em Δ . Todavia, como estes modelos são isométricos, não existe diferença aonde a decisão é tomada. Por outro lado, como as expressões em Δ são mais complexas, implica em uma maior complexidade na construção do receptor. Este fato reforça a grande liberdade de manipulação e possibilidades que temos pelo fato de existirem vários modelos do espaço hiperbólico.

5.3 Constelações PSK

Nesta seção comparamos o desempenho de constelações de sinais M -PSK no espaço hiperbólico, com a correspondente constelação de sinais no espaço euclidiano. Mostramos

que a constelação de sinais M -PSK no espaço hiperbólico oferece ganho de codificação assintótico, quando comparado com a constelação de sinais M -PSK no caso euclidiano.

Aqui consideramos o modelo do semi-plano superior \mathbb{H}^2 para o plano hiperbólico. Denotemos por $C_h(i, r_h)$ o círculo hiperbólico com centro i e raio r_h . Em C_h fixamos M pontos equidistantes $\{P_1 \dots P_M\}$, de modo que eles formem um polígono regular com M lados. Esse conjunto de pontos é a constelação de sinais M -PSK hiperbólica. No plano euclidiano consideramos da mesma forma, M pontos equidistantes $\{Q_1 \dots Q_M\}$ em um círculo de raio r_e e centro em 0. Sejam $P_i = (a_i, b_i)$ e $Q_i = (c_i, d_i)$ as coordenadas dos correspondentes pontos de sinais.

Suponhamos que $Q_i(P_i)$ seja o sinal a ser transmitido. Estamos interessados em determinar a probabilidade de erro quando o sinal $Q_i(P_i)$ é transmitido, e o canal interfere sobre o sinal transmitido via um processo aleatório gaussiano. Assumimos que o receptor utiliza o procedimento de máxima verossimilhança para a decisão. Do teste de hipóteses padrão, podemos determinar as regiões de Voronoi de cada sinal $Q_i(P_i)$. Utilizando as técnicas padrões de integração, a probabilidade de erro associada com as constelações de sinais euclidianas e hiperbólicas são determinadas. As funções densidade de probabilidades utilizadas nos cálculos das probabilidades de erro são dadas por

$$f_{ei}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x - c_i)^2 - \frac{1}{2}(y - d_i)^2\right)$$

no caso do plano euclidiano, e por

$$f_{hi}(z) = 0.184164 \exp\left(-\frac{1}{2}d_h^2(z, z_i)\right)$$

no plano hiperbólico, onde $z = (x, y) \in \mathbb{H}^2$. A constante que aparece em $f_{hi}(z)$ foi calculada via integração numérica, de forma que as pseudo-variâncias de f_{hi} sejam iguais a 1.

Consideremos agora uma constelação 4-PSK hiperbólica, dada pelos sinais z_0, z_1, z_2 e z_3 distribuídos em sentido anti-horário sobre o círculo. Assumimos que os sinais 4-PSK hiperbólicos são utilizados na transmissão da informação. Também assumimos que o ruído no canal de transmissão seja uma amostra do ruído gaussiano em ambos os modelos euclidiano e hiperbólico. Com isso, de acordo com a hipótese de ruído hiperbólico em \mathbb{H}^2 introduzida no Capítulo 3, temos o seguinte teste de hipóteses a considerar

$$H_0 : y_0 = g(z_0)$$

$$H_1 : y_1 = g(z_1)$$

$$H_2 : y_2 = g(z_2)$$

$$H_3 : y_3 = g(z_3)$$

onde z_i são os pontos das constelação PSK e $g(\cdot)$ é uma isometria que age como uma amostra de um processo gaussiano hiperbólico com média zero e pseudo-variância 1.

Como é usual em sistemas de comunicações digitais, iremos assumir que o processo de demodulação é o de máxima verossimilhança. Diante destas fatos, temos que o teste de hipóteses a ser considerado é

$$H_i \longleftrightarrow f_{hi}(z) = 0.184164 \exp\left(-\frac{1}{2}d_h^2(z, z_i)\right).$$

Analisemos a decisão em relação ao ponto z_0 . Os vizinhos de z_0 são os pontos z_1 e z_3 e começamos com z_1 (Figura 5.2). Então, uma decisão a favor da hipótese H_0

será feita se

$$\frac{f_0(y)}{f_1(y)} \leq \lambda \longrightarrow -\frac{1}{2}d_h^2(z, z_0) + \frac{1}{2}d_h^2(z, z_1) \leq \ln \lambda. \quad (5.4)$$

Se os sinais são igualmente prováveis temos que

$$\lambda = \frac{P(H_0)}{P(H_1)} = \frac{1/4}{1/4} = 1.$$

Portanto, $\ln \lambda = 0$, e conseqüentemente, o limiar de decisão será 1. Portanto, em (5.4) temos $d_h(z, z_1) \leq d_h(z, z_0)$. Como visto no Capítulo 2, os pontos $z \in \mathbb{H}^2$ que satisfazem a igualdade $d_h(z, z_1) = d_h(z, z_0)$ são exatamente os pontos que formam a bissetriz do segmento determinado por z_0 e z_1 , que tem equação

$$x^2 + y^2 + \frac{(-2bc + 2ad)x}{b-d} + \frac{b(c^2 + d^2) - d(a^2 + b^2)}{b-d} = 0 \quad y > 0.$$

onde $z_0 = (a, b)$ e $z_1 = (c, d)$ são as coordenadas dos pontos. Se $b = d$, então a bissetriz é a reta euclidiana de equação $x = \frac{a+c}{2}$. Assim, a desigualdade $d_h(z, z_1) \leq d_h(z, z_0)$ é equivalente a

$$x^2 + y^2 + \frac{(-2bc + 2ad)x}{b-d} + \frac{b(c^2 + d^2) - d(a^2 + b^2)}{b-d} \leq 0 \quad y > 0,$$

que é o conjunto dos pontos que se encontram dentro do semi-círculo (geodésica) da bissetriz de z_0 e z_1 , ou seja a região que contém o ponto z_0 .

Por esses cálculos podemos concluir: dado um ponto z_j de uma constelação, quando transmitido, o seu limiar de decisão em relação a qualquer outro ponto z_k da constelação, é dado pela geodésica bissetriz ao segmento formado por z_j e por z_k . É a região de decisão é sempre a região da geodésica que contém z_j . Por essas observações,

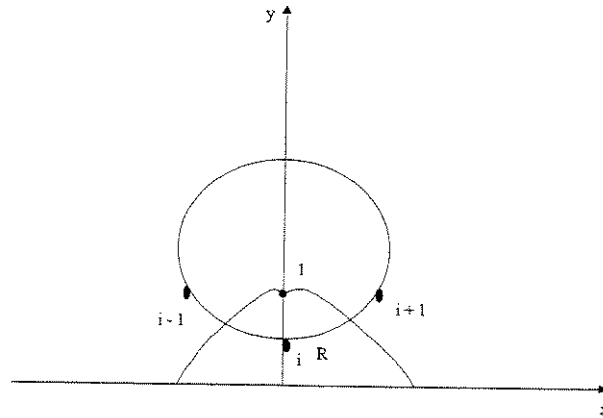


Figura 5.2: Região de Voronoi para um sinal z_i transmitido.

vemos que a construção das regiões de decisão se comporta geometricamente igual ao caso euclidiano.

No caso particular de uma constelação 4-PSK, teremos quatro regiões de decisão. Dessa forma, o receptor ótimo colherá uma amostra do sinal $y(t)$, $y(t = t_k) = y_k$, na saída do canal e verificará em qual região o y_k está contido, e decidirá pela hipótese H_j .

Para o cálculo de P_c^h , devemos calcular o volume de cada gaussiana hiperbólica centrada em cada um dos pontos da constelação e delimitada pela região de Voronoi. Mas como assumimos que os sinais são equiprováveis, basta calcularmos em uma das regiões. Além disso, como a probabilidade de acerto é invariante por rotação, podemos posicionar os sinais da forma mais conveniente sobre o círculo $C_h(i, r_h)$. A Figura 5.2 mostra uma das alternativas utilizadas para obter os valores mostrados nas Tabelas 5.1 e 5.2. Esta figura mostra um sinal z_i e sua região de Voronoi, quando este for utilizado em uma transmissão.

Assim temos que

$$P_c^h = \int \int_R f_h(x, y) \frac{1}{y^2} dy dx.$$

Na Tabela 5.1 são mostrados os ganhos de codificação associados com cada uma das constelações de sinais M -PSK em consideração. Para um valor fixado de probabilidade de acerto, P_c^h , é mostrada a energia média necessária e o valor ao quadrado da distância mínima entre sinais nas correspondentes constelações de sinais.

O ganho de codificação assintótico, GCA, é definido como sendo

$$G = 10 \log \left[\frac{d_h^2 E_M^e}{E_M^h d_e^2} \right],$$

onde d_h^2 e d_e^2 denotam a distância mínima ao quadrado entre pontos da constelação, e E_M^h e E_M^e denotam a energia média nos planos hiperbólico e euclidiano, respectivamente. Definimos a energia média como sendo a média do quadrado das distâncias dos pontos da constelação até a origem do sistema. Neste caso, como os pontos estão sobre um círculo, todas as distâncias até o centro são iguais. Como pode ser visto, os ganhos de codificação assintóticos obtidos pelo uso das constelações de sinais no plano hiperbólico, são de 2.3 dB para o 4-PSK, 6.5 dB para o 8-PSK, 12 dB para o 16-PSK, e acima de 16 dB para o 64-PSK quando comparados com os correspondentes no plano Euclidiano.

Modulação	$P_C^h = P_C^e$	E_M^h	E_M^e	d_h^2	d_e^2	Gan.(dB)
4-PSK	0.4700	0.25	0.4900	0.5201	0.9799	0.1716
	0.7207	1.00	2.1316	2.2901	4.2634	0.5880
	0.9992	9.00	23.5225	28.2142	47.0445	1.9520
	0.999995	16.00	42.25	53.4069	84.50	2.2245
8-PSK	0.2599	0.25	0.5102	0.1570	0.2988	0.3021
	0.4572	1.00	2.4273	0.7596	1.4219	1.1290
	0.9918	9.00	47.61	16.8685	27.8893	5.0509
	0.99988	16.00	100.00	37.0005	58.5786	5.9634
16-PSK	0.1332	0.25	0.5155	0.04117	0.07848	0.3411
	0.2481	1.00	2.5440	0.20666	0.3873	1.3273
	0.9377	9.00	90.250	8.1003	13.7397	7.7173
	0.99761	16.00	249.64	22.5530	38.0054	9.6655
64-PSK	0.0335	0.25	0.5329	0.00261	0.00513	0.3540
	0.0638	1.00	2.5600	0.01327	0.0247	1.3841
	0.4749	9.00	167.184	0.8972	1.6101	11.5701
	0.854813	16.00	>676.00	4.9284	>6.510	>15.049
	0.990274	25.00		16.0641		
	0.99984	36.00		35.7126		

Tabela 5.1: Desempenho mantendo a mesma probabilidade de acerto.

Na Tabela 5.2 são mostradas as probabilidades de acerto associadas com as constelações de sinais quando a energia média dos sinais é fixada. Como podemos observar, as constelações de sinais no plano hiperbólico melhoram consideravelmente o desempenho do sistema de comunicação.

Modulação	$E_M^e = E_M^h$	P_C^h	P_C^e
4-PSK	0.25	0.4700	0.4072
	1.00	0.7207	0.5779
	9.00	0.9992	0.9663
	16.00	0.999995	0.995328
	25.00	0.999999	0.9996
8-PSK	0.25	0.2599	0.2145
	1.00	0.4572	0.3247
	9.00	0.9918	0.7491
	16.00	0.99988	0.87416
	25.00	0.999996	0.94430
16-PSK	0.25	0.1332	0.1088
	1.00	0.2481	0.1678
	9.00	0.9377	0.4416
	16.00	0.997619	0.564823
	25.00	0.999977	0.670664
64-PSK	0.25	0.0335	0.0273
	1.00	0.0638	0.0423
	9.00	0.4749	0.1170
	16.00	0.854813	0.155602
	25.00	0.990274	0.193806
	36.00	0.999843	0.231552

Tabela 5.2. Desempenho mantendo a mesma energia média.

5.4 QAM Circular Hiperbólico

Nesta seção analisamos o desempenho de constelações *QAM*-circulares no plano hiperbólico.

As constelações *QAM* euclidianas são subconjuntos finitos das tesselações euclidianas $\{4, 4\}$, e esses subconjuntos não podem ser reproduzidos em \mathbb{H}^2 , pois como vimos no Capítulo 4, uma tesselação $\{p, q\}$ existe no plano hiperbólico se $(p-4)(q-4) > 4$. Como não podemos realizar no plano hiperbólico as constelações *QAM* euclidianas, para podermos fazer comparações, consideraremos a seguinte situação: fixado o número de pontos da constelação de \mathbb{R}^2 , escolheremos de uma tesselação hiperbólica um subconjunto com o mesmo número de pontos, para fazer o papel da constelação euclidiana. Essa é uma escolha complexa, pois devemos ter em mente o seguinte fato: existem infinitas tesselações no plano hiperbólico, mas fixado p e q , automaticamente fixamos a distância entre os pontos da constelação, pois a menos de movimentos rígidos, não existe em \mathbb{H}^2 o conceito de semelhança, assim uma tesselação $\{p, q\}$ no plano hiperbólico é essencialmente única.

Consideremos as tesselações $\{p, 3\}$, $p > 6$. Como feito no Capítulo 4, fixemos um polígono qualquer P_0 de $\{p, 3\}$, o qual chamamos de polígono inicial ou nível zero da tesselação. O primeiro nível de $\{p, 3\}$ é o conjunto formado por P_0 unido com os polígonos de $\{p, 3\}$ que possuem um vértice em comum com P_0 , denotamos esse conjunto por L_1 . Nesse primeiro nível, existem dois conjuntos de pontos que pode-

mos usar como constelação de sinais: o conjunto formado pelos vértices dos polígonos que constituem L_1 , e o conjunto formado pelos centros dos polígonos de L_1 . Temos respectivamente que o número de pontos no primeiro conjunto é $2p + p(p - 4)$, e no segundo é $p + 1$, logo o conjunto formado pelos centros dos polígonos é mais interessante, pois dado $n > 6$ sempre temos uma constelação hiperbólica com n pontos, para isto basta considerarmos o primeiro nível da tesselação $\{p, 3\}$ com $p = n - 1$, que denotaremos de *QAM-circular*. Como o nosso interesse são as constelações *QAM-circulares*, iremos considerar dois casos: 16-*QAM* e 64-*QAM*. Logo, as constelações de sinais *QAM-circulares* hiperbólicas, são provenientes respectivamente das tesselações $\{15, 3\}$ e $\{63, 3\}$. A construção do primeiro nível dessas tesselações é tão complexa quanto os casos onde $q > 3$.

Além da facilidade em obtermos uma tesselação com o número de pontos que desejamos, existe uma outra justificativa para a escolha das tesselações $\{p, 3\}$. Quanto maior for a área originada pela região de Voronoi em uma constelação, maior será a probabilidade de acerto. Suponhamos que nossa constelação tenha n pontos, originados da tesselação $\{n - 1, 3\}$. Dado um polígono desta tesselação, os seus ângulos internos medem $2\pi/3$, portanto pela equação (2.10) do Capítulo 2 a área A deste polígono é dada por

$$A = (n - 3) - \frac{2\pi(n - 1)}{3}.$$

Seja $\{p, q\}$ uma tesselação tal que o número total de pontos centrais até o nível i seja também n . Não podemos ter $p > n$, pois já no primeiro nível teríamos um número de centros maior que n . Assim, p terá que necessariamente ser menor do que

n . Logo a área é dada por

$$A_1 = (p - 2)\pi - \frac{2\pi p}{q}. \quad (5.5)$$

Como $p < n$, então $q \neq 3$ e das equações (4.3) e (4.5) do Capítulo 4, quando $q > 3$, o número de pontos cresce muito rápido nos níveis. Portanto para termos os mesmos n pontos devemos ter que p é muito menor que n . Com isso $A_1 < A$, implicando que a escolha de $\{n - 1, 3\}$ dará sempre um melhor desempenho.

Consideremos como exemplo a determinação de uma constelação 64-QAM-circular. Sabemos que o primeiro nível da tesselação $\{63, 3\}$ possui 64 pontos, mas quais as outras possibilidades de tesselações conduzindo a 64 pontos? Utilizando as funções enumeradoras desenvolvidas no Capítulo 4, no caso "construção por polígonos", e contando os centros, obtemos que os primeiros níveis das tesselações $\{7, 11\}$, $\{9, 9\}$, $\{21, 5\}$ e $\{3, 23\}$ também possuem 64 pontos. Na Tabela 5.3 apresentamos a área dos polígonos nestas tesselações,

$\{p, q\}$	$\{7, 11\}$	$\{9, 9\}$	$\{21, 5\}$	$\{3, 23\}$	$\{63, 3\}$
Área	11,709	15,708	33,300	2,322	59,690

Tabela 5.3: Tesselações com 64 pontos no primeiro nível.

Como mencionado, a área do polígono na tesselação $\{63, 3\}$ é maior do que a área das demais tesselações.

Nos cálculos das probabilidades de acerto, o modelo considerado para o plano hiperbólico será o do disco Δ , o polígono inicial de cada tesselação tem seu centro na origem e um de seus vértices está na parte positiva do eixo x . Para o cálculo das probabilidades, as funções de densidade de probabilidade gaussiana Euclidiana

escolhidas possuem variâncias iguais a 1. No caso hiperbólico as *pdf* gaussianas possuem a constante B , (ver equação (3.8) no Capítulo 3) calculada de forma que as pseudo-variâncias das *pdf* gaussianas hiperbólicas também sejam 1.

Iremos assumir que o processo de demodulação é o de máxima verossimilhança. Diante destes fatos, temos que o teste de hipóteses a ser considerado é

$$H_i \longleftrightarrow f_{hi}(z) = 0.184164 \exp\left(-\frac{1}{2}d_h^2(z, z_i)\right),$$

onde agora d_h representa a distância hiperbólica no círculo unitário. As regiões de decisão são obtidas também como no caso do PSK e do QAM Euclidiano. Portanto, dado um ponto z_0 da constelação, toma-se as intersecções das regiões obtidas pelas bissetrizes de z_0 e z_i , onde os z_i são os outros pontos da constelação.

No caso $\{15, 3\}$ temos 16 pontos, sendo um a origem e os outros 15 contidos em um círculo centrado em 0. A cada ponto da constelação associamos uma função densidade de probabilidade gaussiana hiperbólica dada por

$$f_i(z) = 0.184164 \exp\left(-\frac{1}{2} \ln^2\left(\frac{|1 - z\bar{w}_i| + |z - w_i|}{|1 - z\bar{w}_i| - |z - w_i|}\right)\right) \quad (5.6)$$

onde o ponto $z = x + yi$ e $w_i = (a_i, b_i)$ é o centro da gaussiana. Como na Seção 5.3, através do teste de hipóteses, essas *pdf* determinam as regiões de Voronoi da constelação. Neste caso temos quinze regiões infinitas e uma finita, que é exatamente o polígono inicial. Devido a simetria das quinze regiões as probabilidades de acertos devem ser iguais nessas regiões, então devemos somente calcular a probabilidade p_i em uma dessas regiões (R_i) e a probabilidade no polígono central p_c . Logo, a

probabilidade média é dada por

$$p_m = \frac{15p_i + p_c}{16}.$$

A seguir apresentamos os cálculos necessários para a obtenção dessas probabilidades. Primeiro, p_c é a probabilidade no polígono fundamental e a gaussiana está centrada na origem, assim pela simetria da gaussiana, giramos o polígono de forma a termos dois vértices simétricos em relação ao eixo x , e obtemos o triângulo hiperbólico OAB mostrado na Fig. 5.3.

Esse triângulo é dividido nos triângulos OAM e OMB , portanto a probabilidade de acerto em relação ao polígono principal é dado pela integral

$$p_c = 30 \left[\int_0^{l_1} \int_0^{\tan(\frac{\pi}{p})x} f_0(z) \left[\frac{2}{1-x^2-y^2} \right]^2 dydx + \int_{l_1}^{l_2} \int_{c_1}^{\tan(\frac{\pi}{p})x} f_0(z) \left[\frac{2}{1-x^2-y^2} \right]^2 dydx \right]$$

onde l_1 e l_2 são constantes dadas no eixo x , e c_1 é a equação da parte acima do eixo x do círculo Euclidiano que determina a reta hiperbólica AB . Então dividimos a integração no triângulo OAM em duas regiões, uma envolvendo somente retas e outra que envolve o círculo. Esse raciocínio pode ser estendido para toda tesselação $\{p, 3\}$, e em cada caso devemos recalculamos l_1 , l_2 e c_1 , sendo que a constante que multiplica as integrais é $2p$ (30 neste caso). No caso de $p = 15$ temos $l_1 = 0.642343$, $l_2 = 0.664722$ e $c_1 : y = \sqrt{1 - 2.199142x - x^2}$. As constantes aqui são dadas a nível de exemplo, pois nos cálculos numéricos elas foram utilizadas com pelo menos 20 casas decimais.

Para o cálculo da probabilidade (p_i) associada com a região R_i , procedemos de forma semelhante, usando parte da construção acima, onde um dos centros está ex-

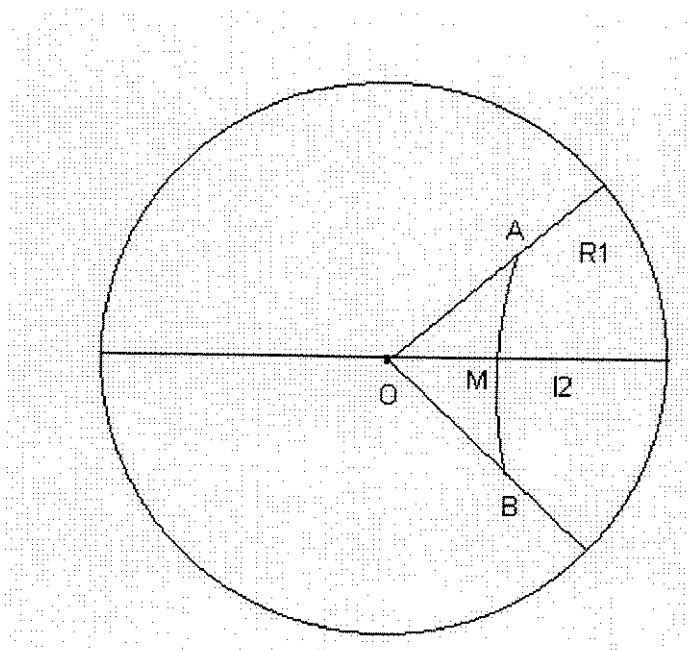


Figura 5.3: Regiões de integração para obtenção das probabilidades de acerto, para a constelação 16-QAM circular hiperbólico.

atamente sobre o eixo x , assim temos

$$p_i = 2 \left[\int_{l_1}^{l_2} \int_0^{c_1} f_0(z) \left[\frac{2}{1-x^2-y^2} \right]^2 dydx + \int_{l_2}^1 \int_0^{\tan(\frac{\pi}{p})x} f_0(z) \left[\frac{2}{1-x^2-y^2} \right]^2 dydx \right],$$

onde os parâmetros são os mesmos da integral acima e são relacionados com a mesma figura. Lembremos que a energia média de uma constelação de pontos $P_1 \dots P_k$, tanto euclidiana quanto hiperbólica, que tenha centro de massa em um ponto Q é dada por

$$EM = \frac{1}{k} \sum_{i=1}^k d(P_i, Q)^2$$

onde d é a função que fornece a distância entre dois pontos.

A constelação 16- QAM euclidiana que usamos para comparar com a hiperbólica, é construída de forma que possua a mesma energia média do caso hiperbólico e de forma que seu centro de massa seja a origem. No caso em questão, o conjunto 16- QAM euclidiano é constituído de quadrados com lado $l = 1.86688$. Para o cálculo das probabilidades do caso euclidiano usamos as gaussianas

$$g_j(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x - x_i)^2 - \frac{1}{2}(y - y_i)^2\right)$$

onde $P_i = (x_i, y_i)$ são os pontos que formam a constelação 16- QAM . Os procedimentos para a determinação das regiões de Voronoi bem como das probabilidades, são bem conhecidos. Assim obtemos que a probabilidade média na constelação hiperbólica é $P_c^h = 0.81799$ e a euclidiana é $P_c^e = 0.54325$. Temos os valores das probabilidades euclidiana e hiperbólica, mas qual é o ganho obtido quando utilizamos a constelação hiperbólica? Para obtermos a resposta consideremos o seguinte problema: Qual é a medida do lado da constelação 16- QAM euclidiana que nos fornece a mesma probabilidade média que a hiperbólica? No caso em questão, com um novo lado de aproxi-

madamente $l' = 3.02$, temos uma probabilidade média de 0.81309, bastante próxima da hiperbólica, assim o ganho é definido por

$$GCA = 10 \log \left[\frac{l_h^2 E_M^e}{E_M^h l_e^2} \right],$$

onde l_h é a distância entre os pontos da constelação *QAM*-circular hiperbólica, l_e é a distância entre dois pontos consecutivos da constelação *QAM* euclidiana, que possui a mesma probabilidade média da constelação hiperbólica, E_M^2 é a energia média hiperbólica e E_M^e é a energia média euclidiana. A energia média euclidiana no caso de uma constelação 16-QAM é dada por

$$E_M^e = \frac{5l_e^2}{2},$$

e no caso de uma constelação 64-QAM é dada por

$$E_M^e = \frac{672l_e^2}{64}.$$

Para os cálculos das probabilidades na constelação 64-QAM, estamos considerando o primeiro nível da constelação $\{63, 3\}$, repetindo os argumentos utilizados no caso da constelação 16-QAM circular. Na Tabela 5.4 temos compilados os resultados obtidos para as constelações 16 e 64-QAM, bem como os desempenhos em relação aos casos euclidianos. Na tabela temos as seguintes notações: distância origem nos fornece a distância dos pontos da constelação hiperbólica até a origem, no caso das tesselações $\{p, 3\}$ essa distância é a mesma da distância entre dois pontos consecutivos da constelação; novo lado e nova probabilidade média, que se referem aos valores da constelação euclidiana construída para alcançar a probabilidade hiperbólica.

	$n = 16$	$n = 64$
dist. origem	3.04861	5.99265
dist. entre sinais	3.04861	5.99265
energia média hip.	8.71317	35.3507
área polígono princ.	3π	19π
prob. polígono princ.	0.76301	0.99625
prob. região infinita	0.82165	0.99980
prob. média hip.	0.81799	0.99975
lado euclidiano	1.86688	1.83486
prob. média euclid.	0.54325	0.47050
área euclidiana	3.48527	3.3667
novo lado	$\simeq 3.02$	$\simeq 6.6$
nova prob. média	0.81309	0.99833
Ganho (dB)	4.26	10.28

Tabela 5.4: Desempenho das constelações 16 e 64-QAM circulares.

Analisando a Tabela 5.4 verificamos que as constelações de sinais $\{15, 3\}$ e $\{63, 3\}$ no plano hiperbólico possuem ganhos de codificação de 4.26 e 10.20 dBs quando comparadas com as constelações de sinais 16 e 64-QAM no plano euclidiano. Aqui citamos que estes ganhos de codificação também podem ser obtidos no espaço euclidiano pelo uso de códigos de Ungerboeck com memórias $\mu = 5$ e $\mu = 6$, respectivamente. A complexidade envolvida para a utilização dos códigos de Ungerboeck é superior à complexidade de um sistema que utiliza somente a constelação hiperbólica, justifi-

cando assim, a troca de constelações euclidianas por hiperbólicas, quando isto for possível.

5.5 Outras Constelações

Além das constelações *PSK* e *QAM*-circulares, qualquer subconjunto finito de uma tesselação hiperbólica pode ser utilizado como constelação de sinais em um sistema de comunicações. Nas tabelas a seguir, apresentamos um limitante inferior para a probabilidade de acerto de diversas tesselações $\{p, q\}$. Este limitante é a probabilidade de acerto quando consideramos uma constelação com infinitos pontos, ou equivalentemente quando o efeito de borda é desprezado. A função densidade de probabilidade gaussiana hiperbólica, utilizada para calcular a probabilidade de acerto em todos os casos a seguir, possui a mesma constante B da função (5.6) que aparece na seção anterior. O centro do polígono inicial de cada tesselação $\{p, q\}$, contruída em Δ é $(0, 0)$ que é o centro da *pdf* gaussiana que estamos considerando.

q	7	8	9	10	15	20	100
P_A	0,07747	0,12881	0.16470	0.19080	0.25447	0.27750	0.30933

Tabela 5.5: $p = 3$.

q	5	6	7	8	9	10	20	100
P_A	0.19913	0.30066	0.36003	0.39790	0.42359	0.44181	0.49912	0.516915

Tabela 5.6: $p = 4$.

q	4	5	6	7	8	9	30	100
P_A	0.24218	0.41064	0.49270	0.53925	0.56830	0.58769	0.65040	0.65568

Tabela 5.7: $p = 5$.

q	4	5	6	7	8	9	50	100
P_A	0.41418	0.55598	0.62266	0.65964	0.68237	0.69738	0.74775	0.74891

Tabela 5.8: $p = 6$.

q	3	4	5	6	7	10	100	200
P_A	0.17140	0.53981	0.65913	0.71362	0.74331	0.78133	0.81313	0.81335

Tabela 5.9: $p = 7$.

q	3	4	5	6	8	10	100	200
P_A	0.30733	0.63355	0.73414	0.77895	0.81748	0.83339	0.85840	0.85858

Tabela 5.10: $p = 8$.

q	3	4	5	6	8	9	100	200
P_A	0.41641	0.70473	0.78980	0.82687	0.85825	0.86585	0.89101	0.89115

Tabela 5.11: $p = 9$.

q	3	4	6	11	12	100	200
P_A	0,63692	0,83632	0.91074	0.93758	0.93909	0.94632	0.94639

Tabela 5.12: $p = 12$.

q	3	4	7	11	16	100
P_A	0.79262	0.91737	0.96400	0.97183	0.97419	0.97609

Tabela 5.13: $p = 16$.

q	3	4	5	30	64
P_A	0.99647	0.99927	0.99964	0.99988	0.99989

Tabela 5.14: $p = 64$.

Como podemos verificar, quando o número de lados p do polígono inicial cresce, também aumenta a probabilidade de acerto. Isso se deve a fórmula (5.5) para a área dos polígonos em uma tesselação $\{p, q\}$. Como podemos verificar, a área é função direta de p e crescerá com o aumento deste.

Aqui devemos citar que as tesselações euclidianas e hiperbólicas estão relacionadas com a teoria de superfícies da seguinte forma: as tesselações euclidianas podem ser orientadas de forma a gerarem um tóro. As tesselações hiperbólicas não podem ser orientadas de forma a gerarem um tóro, mas podem gerar bi-tóros, tri-tóros, etc. Portanto existe uma relação entre as tesselações e o gênero de uma superfície (o número de furos; a esfera tem gênero 0, o toro tem gênero 1 e assim por diante). Um importante subconjunto das tesselações são as chamadas auto-duais, que são as tesselações $\{p, q\}$ tais que $p = q$. Estas tesselações geram os n -tóros a partir de simples orientações. A nível de exemplo a partir das tabelas 5.5 a 5.14, construímos a Tabela 5.15 que mostra o desempenho de tesselações auto-duais e também o desempenho daquelas com $q = 3$. Em nosso trabalho, não exploramos esses aspectos de teoria de superfícies, e suas possíveis relações com sistemas de comunicações, mas fazemos esse comentário de forma a apontar para uma possível pesquisa futura.

$\{p, q\}$	P_c^h	$\{p, 3\}$	P_c^h
$\{5, 5\}$	0.41064	-	-
$\{6, 6\}$	0.62266	-	-
$\{7, 7\}$	0.74331	$\{7, 3\}$	0.17140
$\{8, 8\}$	0.81748	$\{8, 3\}$	0.30733
$\{9, 9\}$	0.86585	$\{9, 3\}$	0.41641
$\{12, 12\}$	0.93909	$\{12, 3\}$	0.63692
$\{16, 16\}$	0.97419	$\{16, 3\}$	0.79262
$\{64, 64\}$	0.99989	$\{64, 3\}$	0.99647

Tabela 5.15: Desempenho de constelações

auto-duais em relação às $\{p, 3\}$.

Capítulo 6

Conclusões

Este trabalho apresentou pela primeira vez, o uso de uma geometria não-euclidiana para modelar um sistema de comunicações. Desta forma ele é um trabalho pioneiro no contexto de Teoria de Comunicações, projeto de Sistemas de Comunicações. Assim, a preocupação principal foi de sermos bastante cuidadosos com a elaboração dos conceitos teóricos necessários para o embasamento do trabalho proposto, além de apresentarmos resultados que justifiquem o possível interesse prático futuro..

6.1 Desenvolvimento

O Capítulo 2 é introdutório, nele apresentamos de forma sucinta a geometria hiperbólica, utilizando conceitos de geometria Riemanniana. Neste capítulo apresentamos dois modelos para a geometria hiperbólica, que são aqueles utilizados ao longo de todo o trabalho. Apresentamos os principais teoremas relacionados à geometria hiperbólica, que são relevantes no desenvolvimento do trabalho.

Neste capítulo, as definições e construções de objetos geométricos hiperbólicos são apresentados com maiores detalhes, possibilitando maior clareza no desenvolvimento dos outros capítulos. Finalmente são feitos diversos comentários e observações, com o objetivo de deixar claro as semelhanças e as diferenças existentes entre as geometrias euclidiana e hiperbólica.

No Capítulo 3, um dos mais importantes deste trabalho, apresentamos o desenvolvimento da função densidade de probabilidade gaussiana no plano hiperbólico. Este resultado é fundamental, pois é através desta função que podemos calcular o desempenho de sistemas de comunicações que utilizem constelações hiperbólicas. As definições e cálculos desenvolvidos ilustram as dificuldades que o espaço hiperbólico apresenta, principalmente no desenvolvimento do "ferramental matemático", similar àquele do caso Euclidiano, no qual é utilizado fortemente as propriedades de espaço vetorial e de norma.

O Capítulo 4 introduz o conceito de tesselação do plano, seja euclidiano ou hiperbólico. É através das tesselações do plano que poderão ser projetadas as constelações de sinais para os sistemas de comunicações. Portanto, é necessário dispormos de funções enumeradoras que possibilitem efetuar a contagem do número de pontos de subconjuntos finitos das tesselações, de forma que possamos selecionar as tesselações de interesse. As funções enumeradoras, apresentadas neste capítulo, permitem um total controle sobre o crescimento do número de vértices e polígonos nos subconjuntos finitos considerados. Dessa forma, a partir do número de pontos é possível propor novas constelações de sinais para sistemas de comunicações.

No Capítulo 5, construímos constelações de sinais no plano hiperbólico. A con-

strução destas constelações tem como objetivo emprega-las em sistemas de comunicações, de forma semelhante ao que é feito no caso euclidiano. Verificou-se que constelações PAM e PSK podem ser construídas sem grandes dificuldades, mas as constelações QAM não podem ser reproduzidas no plano hiperbólico, pois o reticulado \mathbb{Z}^2 não está contido em \mathbb{H}^2 . Todavia as constelações QAM circulares [10], usuais em \mathbb{R}^2 , podem ser reproduzidas em \mathbb{H}^2 . Dessa forma, devemos procurar subconjuntos das tesselações hiperbólicas que tenham o mesmo número de pontos que a constelação QAM circular. Para isto, as fórmulas desenvolvidas no Capítulo 4 são fundamentais.

Paralelamente à construção das constelações, foram feitas análises de desempenho para algumas destas constelações de sinais no plano euclidianos como no plano hiperbólico. Em todos os casos verificou-se um melhor desempenho da constelação hiperbólica frente à equivalente euclidiana. Isto mostra que, quando o canal de transmissão tiver um comportamento descrito por gaussianas hiperbólicas, a utilização de constelações hiperbólicas fornecerá desempenhos bem melhores.

6.2 Trabalhos Futuros

O trabalho aqui desenvolvido apresenta diversas perspectivas de continuidade, visando diversos contextos. No contexto de modulação digital, um estudo experimental do modelo de canal gaussiano hiperbólico é fundamental, pois através de uma perfeita caracterização deste canal, ele poderá ser utilizado em sistemas comerciais.

O estudo de teoria de probabilidade em espaços hiperbólicos necessita ser completamente estabelecida e a obtenção de um Teorema Central do Limite para mul-

tivariáveis hiperbólicas mais geral do que aquele que aparece em [12] é fundamental para o estudo do canal gaussiano hiperbólico.

Os resultados obtidos neste trabalho são aplicados no espaço hiperbólico bidimensional. A extensão dos resultados para o espaço hiperbólico n -dimensional seria interessante e importante.

Num contexto próximo, mas não tratado neste trabalho, o estudo de códigos corretores de erros no contexto hiperbólico é um dos principais problemas que aqui propomos. A generalização de conceitos como G -linearidade, casamento de códigos e grupos, uniformidade geométrica e outros conceitos semelhantes, além daqueles apresentados em [8] é muito importante e devem fornecer novos resultados. Para isto, o estudo de grupos de isometrias do plano hiperbólico (grupos Fuchsianos) ou de dimensões superiores (grupos Kleinianos) é imprescindível.

Referências

- [1] Beardon, A. F., The Geometry of Discrete Groups, Springer-Verlag, 1983.
- [2] Beardon, A. F., An Introduction to Hyperbolic Geometry, on "Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces ", Oxford University Press, 1991.
- [3] Carmo, M. P., Geometria Riemanniana, Impa-Projeto Euclides, 1979.
- [4] Coxeter, H. S. M. e Moser, W. O. J., Generators and Relations for Discrete Groups, Terceira Edição, Springer-Verlag, 1972.
- [5] Firby, P. A. e Gardiner, C. F., Surface Topology, Segunda Edição, Ellis Horwood, 1991.
- [6] Firer, M., Grupos Fuchsianos, Notas de Aula, 1998.
- [7] Katok, S., Fuchsian Groups, The University of Chicago Press, 1992.
- [8] Lazari, H., Uma Contribuição Teoria dos Códigos Geométricamente Uniformes Hiperbólicos, Tese, FEEC-Unicamp, 2000.
- [9] Papoulis, A., Probability, Randon Variables and Stochastic Processes, McGraw-Hill, 1965.
- [10] Proakis, J. G., Digital Communications (segunda edição), McGraw-Hill, 1989.
- [11] Shannon, C. E., A Mathematical Theory of Communications, Bell Syst. Tech. J., vol. 27, 379-423, 623-656.
- [12] Terras, A., Harmonic Analysis on Symmetric Spaces and Applications I, Springer-Verlag, 1985.
- [13] Wozencraft, J. M. & Jacobs, I. M., Principles of Communication Engineering, John Wiley & Sons, 1965.