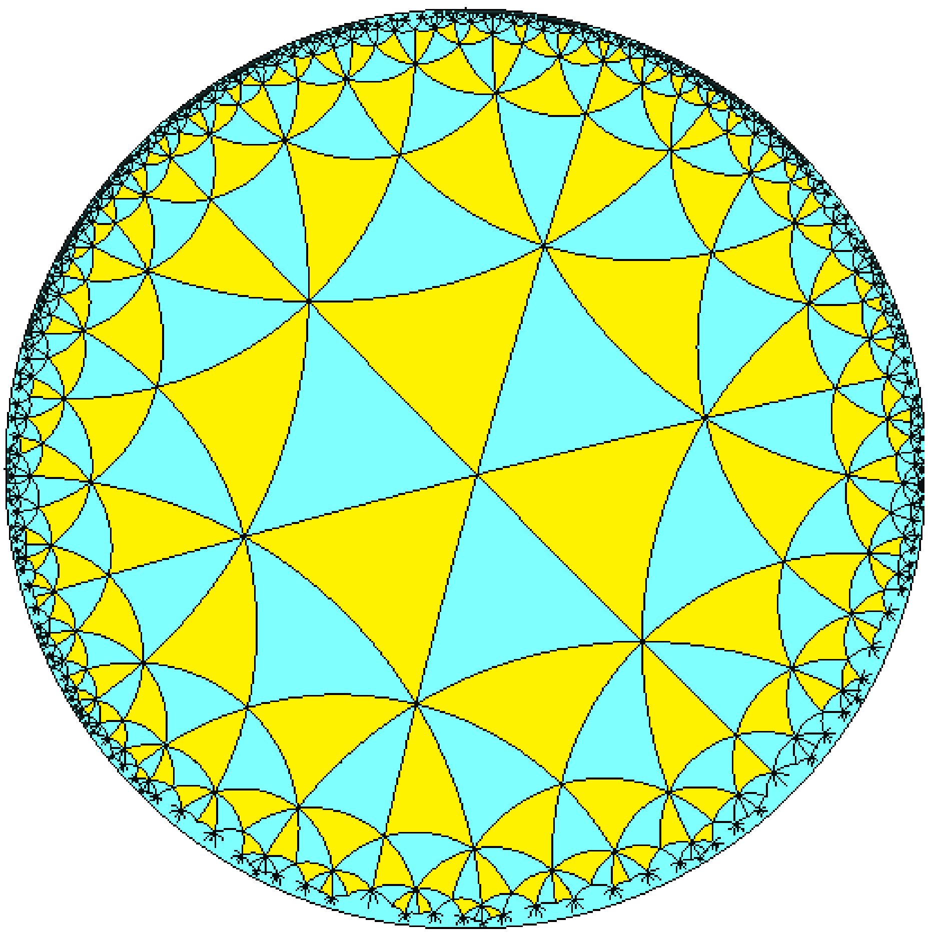
Medidas Conformes em Geometria Hiperbólica

KELVYN WELSCH

ORIENTADOR: RODRIGO BISSACOT



UNIVERSIDADE DE SÃO PAULO

Financiado pela FAPESP - processo 2019/26838-7

Conteúdo

Introdução

Este texto visa a introduzir o conceito e algumas propriedades dos grupos fuchsianos. Especificamente, nosso objetivo último é provar alguns resultados sobre a dimensão Hausdorff do conjunto limite de grupos fuchsianos de segundo tipo, resultado este devido a Patterson [Pat76], com ênfase na construção de medidas conformes que, apesar de ser um passo intermediário, abrirá caminho para estudar a teoria de Denker-Urbanski. Grupos fuchsianos são uma classe de grupos que possuem ligações com diversas áreas da matemática, tais quais geometria hiperbólica, geometria projetiva e análise complexa. Os grupos fuchsianos foram nomeado em homenagem a Lazarus Fuchs (1833-1902), sendo Poincaré o primeiro a se debruçar mais profundamente sobre o estudo deles.

Diferentemente de como é o caso de grande parte dos textos, este trabalho foi escrito ao mesmo tempo em que o autor aprendia os assuntos apresentados. Se, por um lado, isso fez com que fosse possível atentar para detalhes e aspectos que dificultaram a experiência de aprendizagem do autor, de modo que a didática ficasse mais adequada para quem nunca teve contato com o assunto, por outro, o esforço para que o texto apresentasse homogeneidade no estilo, fluidez e continuidade foram maiores que os usuais, dado que a experiência de aprendizagem é muitas vezes oscilante, interrupta não-homogênea e a longo prazo. Advirto desde já que, apesar de grandes, os esforços para suavizar o aspecto quimérico do texto podem não ter sido suficientes.

A fim de termos um melhor aproveitamento sobre o assunto, o texto será dividido em duas partes. Na primeira, iremos recordar alguns resultados básicos de teoria de grupos, topologia, grupos topológicos e análise complexa. Na segunda parte, voltada mais diretamente para os nossos propósitos, daremos um panorama geral sobre geometria hiperbólica antes de lidarmos com os grupos fuchsianos propriamente ditos. Naturalmente, leitores mais confortáveis com os requisitos podem pular esta parte introdutória. De modo geral, supôe-se que o leitor saiba, no mínimo, a definição de um grupo e de uma topologia.

Parte I

Requisitos

Capítulo 1

Teoria de Grupos

1.1 Homomorfismos

Definição 1.1. Sejam $(G, \cdot)$, $(H, *)$ grupos. Uma função $h : G \to H$ é dita um *homomorfismo de grupos* se $h(a \cdot b) = h(a) * h(b)$, $\forall a, b \in G$. Quando um homomorfismo é injetivo, sobrejetivo e bijetivo, dizemos que é um monomorfismo, epimorfismo e isomorfismo, respectivamente. Quando o domínio e o contradomínio são idênticos, dizemos que se trata de um endomorfismo. Um automorfismo é um endomorfismo que também é um isomorfismo. ♣

Dizemos que um homomorfismo de grupos é uma função que preserva a estrutura do grupo. A razão para isto é a que se segue. Por definição, um homomorfismo preserva a operação de grupo. Além disso, a proposição abaixo nos diz que o elemento neutro e a operação de elemento inverso também são preservadas. Em outras palavras, um homomorfismo mapeia elemento neutro em elemento neutro e o homomorfismo do inverso é o inverso do homomorfismo. De forma ainda mais clara, quando dizemos que homomorfismos são mapas que preservam a estrutura do grupo, queremos dizer que tanto faz realizar as operações de grupo no domínio ou na imagem.

A partir de agora iremos deixar de enfatizar e distinguir os símbolos das operações dos grupos, pelo bem da sanidade mental.

Proposição 1.1. *Sejam G, H grupos e $h : G \to H$ um homomorfismo. Se e_G e e_H são os elementos neutros de G e H, respectivamente, vale que $h(e_G) = e_H$. Além disso, se $g \in G$, então $h(g^{-1}) = h(g)^{-1}$.* □

Demonstração. Tome $g \in G$, qualquer. Teremos que $g = e_G \cdot g \implies h(g) = h(e_G) \cdot h(g) \implies h(g) \cdot h(g)^{-1} = h(e_G) \cdot h(g) \cdot h(g)^{-1} \implies e_H = h(e_G) \cdot e_H \implies e_H = h(e_G)$.

Usando isso, $e_G = g \cdot g^{-1} \implies e_H = h(g) \cdot h(g^{-1}) \implies h(g)^{-1} = h(g)^{-1} \cdot h(g) \cdot h(g^{-1}) = e_H \cdot h(g^{-1}) = h(g^{-1}) \implies h(g)^{-1} = h(g^{-1})$. ■

Definição 1.2. Sejam G e H grupos e $h : G \to H$ um homomorfismo. O *núcleo* de h é o conjunto $\ker h$ definido por $\ker h := \{g \in G; h(g) = e_H\}$. A *imagem* de h é o conjunto $\operatorname{Im} h := \{b \in H; \exists g \in G$ com h(g) = b$\}$ ♣

Proposição 1.2. *Sejam G e H grupos e $h : G \to H$ um homomorfismo. Vale que* $\ker h$ *é subgrupo de G e Im h é subgrupo de H.* $\square$

Demonstração. Vejamos que $\ker h$ é subgrupo de G. Basta mostrarmos que, se $g, f \in \ker h$, arbitrários, então $gf^{-1} \in \ker h$. Pela proposição 1.1, $h(gf^{-1}) = h(g)h(f)^{-1} = e_H e_H^{-1} = e_H$

Vejamos agora que Im h é subgrupo de H. Seja $g, f \in$ Im h. Então existe $a \in G$ tal que $g = h(a)$. e $b \in G$ tal que $h(b) = f$. Então, $gf^{-1} = h(a)h(b)^{-1} = h(ab^{-1})$. Assim, existe um elemento de G (a saber, ab^{-1}) tal que $h(ab^{-1}) = gf^{-1}$, e portanto $gf^{-1} \in$ Im h. Desta forma, concluímos a nossa demonstração. ■

Proposição 1.3. *Sejam G, H grupos e $\theta : G \to H$ um homomorfismo. θ será um monomorfismo se, e somente se* $\ker \theta = \{e_G\}$

Demonstração. Sabemos que $\theta(e_G) = 0$. Assim, se θ for um monomorfismo, então $g \neq e_G \implies \theta(g) \neq e_H$, e portanto $g \notin \ker \theta$, $\forall g \in G, g \neq e_G$. Reciprocamente, suponha que $\ker \theta = \{e_G\}$ e tome $g, h \in G$. Se valesse $gh^{-1} = e_G$, teríamos $g = h$. Assim, supondo $g \neq h$, temos $gh^{-1} \neq e_G$, e portanto $gh^{-1} \notin \ker \theta$, isto é, $\theta(gh^{-1}) = \theta(g)\theta(h)^{-1} \neq e_H$, e portanto $\theta(g) \neq \theta(h)$, donde θ é injetiva. ■

Proposição 1.4. *Sejam G e H grupos e $\theta : G \to H$ um homomorfismo. Dado $M, N \in G$, vale que $\theta(M) = \theta(N) \iff MN^{-1} \in \ker \theta$.* $\square$

Demonstração. $\theta(M) = \theta(N) \iff \theta(M)\theta(N)^{-1} = e_H \iff \theta(MN^{-1}) = e_H \iff MN^{-1} \in \ker \theta$. ■

1.2 Classes de Conjugação e Subgrupos Normais

Definição 1.3. Seja G um grupo. Dois elementos $a, b \in G$ são ditos *conjugados* se existe $g \in G$ tal que $gag^{-1} = b$ ♣

Proposição 1.5. *Seja G um grupo. A relação $\sim$ tal que $a \sim b$ se, e somente se a e b forem conjugados é uma relação de equivalência. As classes de equivalência por tal $\sim$ são chamadas de* classes de conjugação. $\square$

Demonstração. (i) Reflexividade. Dado $a \in G$, para ver que $a \sim a$, basta tomarmos $g = e$. Daí $eae^{-1} = eae = a$. (ii) Simetria. Sejam $a, b \in G$ tais que $a \sim b$. Então existe $g \in G$ tal que $gag^{-1} = b$. Mas daí, $ga = bg$ e $a = g^{-1}bg$. Tomando $h = g^{-1}$ teremos $hbh^{-1} = a$, donde $b \sim a$. (iii) Transitividade. Suponhamos que $a \sim b$ e $b \sim c$. Então existem $g, h \in G$ tais que $gag^{-1} = b$ e $hbh^{-1} = c$. Então teremos $(hg)a(hg)^{-1} = hgag^{-1}h^{-1} = hbh^{-1} = c$, donde $a \sim c$. ■

Elementos conjugados de um grupo costumam possuir propriedades semelhantes. No caso dos grupos estudados por nós, seus elementos serão funções, e elementos conjugados se relacionarão, por exemplo, quanto aos seus pontos fixos. Destacamos, como exemplo, que os elementos conjugados de um grupo de matrizes são chamados de matrizes semelhantes. Entre algumas propriedades elementares citamos: i) {e} é sempre uma classe de conjugação e ii) em um grupo abeliano, as classes de conjugação são conjuntos unitários, uma vez que, $\forall g \in G$, $gag^{-1} = gg^{-1}a = a$.

Definição 1.4. Sejam G um grupo e $N \subset G$ um subgrupo. Dizemos que N é um *subgrupo normal* se, $\forall g \in G$, $\forall n \in N$, $gng^{-1} \in N$. ♣

Da definição, segue diretamente que qualquer subgrupo de grupos abelianos é normal. Este resultado também pode ser enunciado como corolário da proposição 1.6 a seguir. Além disso, é trivial ver que, em qualquer grupo G, os subgrupos $\{e_G\}$ e o próprio G são subgrupos normais. Estes subgrupos normais são chamados de subgrupos triviais.

Como pode se esperar, o conceito de subgrupo normal está intimamente relacionado ao conceito de elementos conjugados. Se $n \in N$ (N normal), então todos os seus elementos conjugados também pertencem a N. Em outras palavras, $[n] \subset N$, onde $[n]$ é a classe de conjugação de N. Com esta observação, fica evidente provar a seguinte proposição:

Proposição 1.6. *Sejam G um grupo e $N \subset G$ um subgrupo. Então N é normal se, e somente se, é a união de classes de conjugação de G.* □

Definição 1.5. Um grupo é dito simples quando seus únicos subgrupos normais são os triviais. ♣

As proposições seguintes demonstram que alguns subgrupos específicos são normais, o que nos será útil adiante.

Proposição 1.7. *Sejam G, H grupos e $h : G \to H$ um homomorfismo. O núcleo de h é um subgrupo normal de G.* □

Demonstração. Pela proposição 1.2, já sabemos que $\ker h$ é um subgrupo. Falta provarmos que é normal. Queremos provar que, dado $n \in \ker h$ e $g \in G$, $h(gng^{-1}) = e_H$. Ora, $h(gng^{-1}) = h(g)h(n)h(g)^{-1}$. Usando que $n \in \ker h$: $h(gng^{-1}) = h(g)e_H h(g)^{-1} = e_H$. ■

Para a próxima proposição, devemos notar que a intersecção arbitrária de subgrupos de um grupo também é subgrupo. A demonstração deste fato não é difícil.

Proposição 1.8. *Seja G um grupo e seja N um subgrupo normal de G. Então, se S é um subgrupo arbitrário de G, $N \cap S$ é um subgrupo normal de S.* □

Demonstração. Tome $g \in S$ e $n \in N \cap S$ arbitrários. Queremos provar que $gng^{-1} \in N \cap S$. Primeiro, note que, em especial, $g \in G$ e $n \in N$. Pela definição de subgrupo normal, vem então que $gng^{-1} \in N$. Por outro lado, $g, n \in S$, donde $gng^{-1} \in S$. Unindo as duas relações de pertinência completamos a demonstração. ■

Além do mais, a intersecção arbitrária de subgrupos normais também é normal, o que nos permite falar em menor subgrupo normal contendo um certo subconjunto do grupo. Este subgrupo é chamado de fecho normal.

1.3 Cosets e Grupo Quociente

Muitas vezes, um grupo possui mais informação do que precisamos. Neste caso, convém dividi-lo em conjuntos que compartilhem a propriedade desejada de forma a formar um novo grupo. Este novo grupo é chamado de grupo quociente. Os conjuntos são classes de equivalências, e para entendê-las, precisamos entender o que é um coset. Um coset é basicamente um subgrupo afim, isto é, a translação de um subgrupo. Como a operação de grupo não necessariamente é comutativa, podemos ter cosets à direita ou cosets à esquerda. Precisamente, temos a seguinte

Definição 1.6. Sejam G um grupo e $H \subset G$ um subgrupo. Dado $g \in G$, o conjunto $gH := \{gh; h \in H\}$ é chamado de *coset à esquerda de H com relação a g* e o conjunto $Hg := \{hg; h \in H\}$ é chamado de *coset à direita de H com relação a g.* ♣

Proposição 1.9. *Sejam G um grupo e $H \subset G$ um subgrupo. Sejam também $\sim_D$ e $\sim_E$ as relações em G tais que $x \sim_D y$ se, e somente se existe $h \in H$ tal que $x = hy$ e $x \sim_E y$ se, e somente se existe $h \in H$ tal que $x = yh$. Então vale que ambas são relações de equivalência e além disso, $D \subset G$ é coset à direita de H se e somente se existe $g \in G$ tal que $D = [g]_D$ e $E \subset G$ é coset à esquerda de H se, e somente se existe $g \in G$ tal que $E = [g]_E$.* □

Demonstração. Provemos que $\sim_D$ é uma relação de equivalência: (i) (reflexiva) Tome $g \in G$. Como H é subgrupo, $e_G \in H$. Assim, existe $h \in H$ tal que $g = gh$ (a saber, $h = e_G$), donde $g \sim_D g$. (ii) (simétrica) Sejam $x, y \in G$ tais que $x \sim_D y$. Então existe $h \in H$ tal que $x = yh$. Como H é subgrupo, $h^{-1} \in H$. Assim, $x = yh \implies xh^{-1} = y$, donde $y \sim_D x$. (iii) (transitiva) Sejam $x, y, z \in G$ tais que $x \sim_D y$ e $y \sim_D z$. Então existem $g, h \in H$ tais que $x = yg$ e $y = zh$. Estas duas igualdades implicam $x = (zh)g = z(hg)$. Ora, $hg \in H$, por motivos óbvios, donde $x \sim_D z$.

Se D é coset à direita de H, por definição, existe $g \in G$ tal que $D = gH$. Assim, $x \in D$ se, e somente se, existe $h \in H$ tal que $x = gh$, que é equivalente a dizer que $x \sim_D g$, e portanto $x \in D$ se, e somente se $x \in [g]_D$. Portanto, $D = [g]_D$.

Os casos para cosets à esquerda e $\sim_E$ se faz de modo análogo. ■

Como estamos interessados no caso em que os cosets formam um grupo, passaremos estudar algumas condições para tal. Uma forma bem natural de garantir que seja um grupo é exportando-as do grupo, isto é: $(gN)(hN) = (gh)N$ e $(gN)^{-1} = (g^{-1})N$. Estas operações nem sempre estão bem definidas...

os cosets à direita e à esquerda coincidem. Nestes casos, o conjunto dos cosets forma um grupo. Este grupo, derivado das classes de equivalências, é chamado de grupo quociente. As hipóteses suficientes e necessárias para que isto ocorra é que H seja um subgrupo normal de G. (vemos aqui, finalmente, a utilidade da definição de subgrupo normal). Estes resultados estão formalmente enunciados a seguir:

Proposição 1.10. *Sejam G um grupo e N um subgrupo de G. Uma condição necessária e suficiente para que $gN = Ng$, $\forall g \in G$ é que N seja um subgrupo normal de G.* □

Demonstração. Por definição, $x \in gN$ se, e somente se existe $n \in N$ tal que $x = gn$. Isto, porém, é verdade se, e somente se $xg^{-1} = gng^{-1}$. Mas $x \in Ng$ se, e somente se existe $m \in N$ tal que

$xg^{-1} = m$, ou seja, se, e somente se $xg^{-1} = gng^{-1} \in N$. Concluímos que $gN = Ng$ se, e somente se $gng^{-1} \in N$, $\forall g \in G$, $\forall n \in N$, que é precisamente a definição de N ser normal. ■

Proposição 1.11. *Sejam G um grupo e N um subgrupo normal de G. Então o conjunto $G/N = \{gN = Ng; g \in G\}$, munido da operação $(gN)(hN) = ghN$ forma um grupo, chamado de* grupo quociente de G por N*, onde $e_G N = N$ é o elemento neutro e $g^{-1}N$ é o elemento inverso de gN.* □

Aqueles que estão mais habituados com álgebra linear hão de perceber a semelhança entre as definições de grupo quociente e espaço quociente. Perceberão também o quão mais fácil é definir o segundo (sem necessidade de recorrer a objetos análogos aos cosets à esquerda e à direita). Isto se deve ao fato de, num espaço vetorial, as operações serem comutativas.

Podemos pensar num grupo como uma figura cujos pixels são os elementos e no grupo quociente como uma versão com menor resolução da figura. Esta visão reforça o que queríamos dizer quando dissemos que o grupo quociente é útil quando o grupo tem mais informações do que precisamos. Um caso mais concreto é dado pela segiunte situação: Seja $\theta : G \to H$ um homomorfismo. Esta função pode não ser injetiva. Pelo fato do núcleo de um homomorfismo ser normal, podemos fazer $G \backslash \ker \theta$. Não é difícil ver que o homomorfismo será constante em cada coset e a nova versão do homomorfismo será injetora. A próxima seção tratará de resultados relacionados a essas ideias.

Além de útil na definição de grupo quociente, os cosets desempenham papel fundamental na demonstração do Teorema de Lagrange, que será enunciado aqui apenas à título de informação.

Definição 1.7. Seja G um grupo. A ordem de G é definida como a cardinalidade do conjunto G. Um grupo G é dito finito se tem ordem finita ♣

Teorema 1.12. *(Teorema de Lagrange) Sejam G um grupo finito e $H \subset G$ um subgrupo. Então a ordem de G é múltipla da ordem de H.* □

Enfatizo que "ordem de um elemento"é um conceito diferente de "ordem de um grupo"e será tratada na seção 1.8.

1.4 Homomorfismo Revisitado

Existem alguns tipos de grupos quocientes que são de particular interesse. Um deles é aquele quocientado pelo núcleo de um homomorfismo. Este caso possui um resultado importante que será contemplado pelo Primeiro Teorema de Isomorfismos. Antes dele, vejamos um teorema mais geral, que é um dos resultados mais importantes deste capítulo.

Teorema 1.13. *(Teorema Fundamental de Homomorfismos). Sejam G e H grupos e $\phi : G \to H$ um homomorfismo. Seja N um subgrupo normal de G tal que $N \subset \ker \phi$. Então a função $\psi : G/N \to H$ definida por $\psi(gN) = \phi(g)$ é um homomorfismo.* □

Corolário 1.13.1. *(Primeiro Teorema de Isomorfismos). Existe um isomorfismo ψ entre $G/\ker \phi$ e Im ϕ dado por $\psi(g \ker \phi) = \phi(g)$* □

Corolário 1.13.2. *Se ϕ é um epimorfismo, então existe um isomorfismo ψ entre $G/\ker \phi$ e H dado por $\psi(g \ker \phi) = \phi(g)$* □

1.5 Grupos de Simetria

Definição 1.8. Seja M um conjunto. O conjunto de todas as bijeções de M para M forma um grupo sob a operação de composição e é chamado de grupo de simetria de M, denotado por $\mathrm{Sym}(M)$. Um subgrupo de um grupo de simetria é chamado de grupo de permutação. ♣

Na verdade, o grupo de simetria de um objeto matemático é o grupo de funções que preservam a estrutura deste objeto. Por exemplo, se for um espaço métrico, o grupo de simetria é o grupo de isometria. Se for um espaço topológico, é o grupo de homeomorfismos, se for uma variedade diferenciável, é o grupo de difeomorfismos... E assim por diante.

1.6 Ações de Grupos

O cerne do estudo dos grupos é "simetria". Este fato é bem conhecido mas, à primeira vista, a definição de um grupo não parece estar nitidamente relacionada com simetria. Isto porque os elementos de um grupo, a grosso modo, são as transformações de simetria em si, desprovidas do objeto a qual se referem. Esta é a motivação da definição de ação de um grupo. A ação de um grupo é a razão de ser de um grupo, aquilo para o qual ele foi construído. Como o nome indica, a ação de um grupo torna explícito a atuação de um grupo num objeto que possui simetrias. Novamente, devido ao fato de um grupo não ser necessariamente comutativo, podemos ter dois tipos de ação, à esquerda e à direita. Sem mais delongas, passemos à definição formal.

Definição 1.9. Sejam M um conjunto não-vazio e G um grupo. Seja $\alpha : G \times M \to M$ uma função. Considere as seguintes propriedades:

1) Dado $g \in G$, a função $\alpha(g, \cdot) : M \to M$ é bijetiva

2) Se e é o elemento neutro de G, então a função $\alpha(e, \cdot) : M \to M$ é a identidade. Ou seja, $\alpha(e, x) = x$, $\forall x \in M$.

3e) Vale $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$, $\forall g, h \in G$, $\forall x \in M$

3d) Vale $\alpha(g, \alpha(h, x)) = \alpha(hg, x)$, $\forall g, h \in G$, $\forall x \in M$

Se α satisfaz 1, 2 e 3e, então é α é dita ser uma ação à esquerda de G. Se for 1, 2, 3d, α é dita ser uma ação à direita de G. ♣

Note que se G é comutativo, ações à direita e à esquerda empatam. Quando dissermos apenas "ação", estará subentendido que pode se tratar de uma ação à direta ou à esquerda. Além disso, muitas das demonstrações serão feitas com apenas um tipo mas a demonstração para o outro caso é completamente simétrica.

Vamos olhar com mais atenção a esta definição e extrair o significado intuitivo de cada propriedade. Dado um elemento g do grupo, a função $\alpha(g, \cdot) : M \to M$ se torna uma transformação de simetria de M. Nada mais justo do que ser bijetiva. A condição 2 apenas impõe que a transformação de simetria que não faz nada (a identidade), corresponda ao elemento neutro. A terceira condição é a mais interessante. Ela diz que realizar uma transformação de simetria por g em M já transformado por h deve corresponder à transformação de simetria composta gh ou hg. Em outras palavras, aplicar h e depois aplicar g deve empatar com aplicar gh (ou hg) de uma vez só. Tal observação, juntamente com a proposição a seguir, tornam uma outra notação funcional. De agora em diante, podemos escrever gx no lugar de $\alpha(g, x)$. A observação que acabou de ser feita torna-se um tipo de associatividade: $(gh)x = g(hx)$. Também:

Proposição 1.14. *Se $\alpha(g, x) = y$, então $\alpha(g^{-1}, y) = x$.* □

Demonstração. Basta substituir a primeira na segunda:

$$\alpha(g^{-1}, y) = \alpha(g^{-1}, \alpha(g, x)) = \alpha(g^{-1}g, x) = \alpha(e, x) = x.$$ ■

Na notação alternativa, a proposição acima só está dizendo que $gx = y \implies x = g^{-1}y$.

Note que uma ação dá origem a um homomorfismo Π entre G e $\mathrm{Sym}(G)$, onde $\Pi(g)(x) = \alpha(g, x)$

Existem algumas características de ações que são importantes de serem vistas. Uma delas é a transitividade:

Definição 1.10. Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. α é dita *transitiva* se, $\forall x, y \in M$, existir $g \in G$ tal que $y = \alpha(g, x)$. Será dita *simplesmente transitiva* quando tal g for único. Mais ainda, α é dita ser *k-transitiva* se, para todo par de k-uplas $(x_1, ..., x_k)$ e $(y_1, ..., y_k)$ com elementos distintos dois a dois, existir um $g \in G$ tal que $y_i = \alpha(g, x_i)$, $1 \leqslant i \leqslant k$. ♣

A estrutura de grupo (juntamente com o fato da definição de ação ser compatível com essa estrutura) permite que a ação de verificar se um grupo é transitivo seja ligeiramente facilitada. Em vez de considerar pares quaisquer de pontos do espaço, podemos fixar um deles e provar que qualquer outro ponto pode ser mandado para ele. Mais geralmente:

Proposição 1.15. *Seja G um grupo agindo num espaço X. A ação é k-transitiva se, e somente se, existe uma k-upla de pontos distintos de X, $(p_1, ..., p_k)$ tal que, para toda k-upla $(x_1, ..., x_k)$ de elementos distintos existe um elemento $g \in G$ tal que $p_i = \alpha(g, x_i), 1 \leqslant i \leqslant k$.* □

Demonstração. Que k-transitividade implica na propriedade acima é imediato da definição. Reciprocamente, suponha que a propriedade valha e tome um par de k-uplas com elementos distintos $(x_1, ..., x_k)$ e $(y_1, ..., y_k)$. Então existem $g, h \in G$ tais que $p_i = \alpha(g, x_i)$ e $p_i = \alpha(h, y_i), 1 \leqslant i \leqslant k$. Então existe um elemento de G, a saber, $h^{-1}g$ tal que $y_i = \alpha(h^{-1}g, x_i), 1 \leqslant i \leqslant k$. De fato, $\alpha(h^{-1}g, x_i) = \alpha(h^{-1}, \alpha(g, x_i)) = \alpha(h^{-1}, p_i) = y_i, 1 \leqslant i \leqslant k$, por (1.14). ■

A partir de agora, quando provarmos que uma ação é transitiva, provaremos simplesmente a propriedade acima.

Definição 1.11. Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Dado $m \in M$, a *órbita de m pela ação de α* é o conjunto $Orb_\alpha(m) := \{\alpha(g, m); g \in G\}$. ♣

Em palavras, a órbita de m por α é o conjunto dos pontos de M que m pode assumir, considerando todas as transformações de simetria que G pode oferecer através de α. A proposição seguinte abrirá o caminho para uma definição alternativa de ação transitiva.

Proposição 1.16. *Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Dado $m \in M$, vale que $n \in Orb_\alpha(m) \implies Orb_\alpha(n) = Orb_\alpha(m)$.* □

Demonstração. Suponhamos que α seja uma ação à esquerda. Mostraremos inicialmente que $Orb_\alpha(n) \subset Orb_\alpha(m)$. Seja $x \in Orb_\alpha(n)$. Então existe $g \in G$ tal que $\alpha(g, n) = x$. Por outro lado, como $n \in Orb_\alpha(m)$, então existe $h \in G$ tal que $\alpha(h, m) = n$. Ora, substituindo uma coisa na outra, $x = \alpha(g, \alpha(h, m)) = \alpha(gh, m)$, em que usamos a definição de ação à esquerda. A última igualdade mostra que $x \in Orb_\alpha(m)$.

Veja que mostramos a seguinte implicação: $n \in Orb_\alpha(m) \implies Orb_\alpha(n) \subset Orb_\alpha(m)$. Trocando o papel das letras, teremos: $m \in Orb_\alpha(n) \implies Orb_\alpha(m) \subset Orb_\alpha(n)$. Assim, se mostramos que $n \in Orb_\alpha(m) \implies m \in Orb_\alpha(n)$, obteremos automaticamente a igualdade desejada. Novamente, a hipótese nos leva a crer que existe $h \in G$ com $\alpha(h, m) = n$. Pela definição de ação, $\alpha(e, m) = m$. Donde $\alpha(h^{-1}h, m) = m \implies \alpha(h^{-1}, \alpha(h, m)) = m \implies \alpha(h^{-1}, n) = m$ e portanto $m \in Orb_\alpha(n)$.

Para ações à direita, a demonstração é análoga. ■

Corolário 1.16.1. *A relação definida por $m \sim n$ se, e somente se $m \in Orb_\alpha(n)$ é uma relação de equivalência. (e portanto as órbitas formam classes de equivalência em M)* □

Corolário 1.16.2. *Se existe $m \in M$ tal que $Orb_\alpha(m) = M$, então $Orb_\alpha(n) = M$, $\forall n \in M$.* □

Demonstração. Dado $n \in M$, como $Orb_\alpha(m) = M$, então $n \in Orb_\alpha(m)$. Pela proposição, $Orb_\alpha(n) = Orb_\alpha(m) = M$. ■

Corolário 1.16.3. *Uma ação α é transitiva se, e somente se existir $m \in M$ tal que $Orb_\alpha(m) = M$* □

Demonstração. Pelo corolário acima, existir $m \in M$ tal que $Orb_\alpha(m) = M$, é equivalente a dizer que $Orb_\alpha(x) = M$, $\forall x \in M$. Dados $x, y \in M$ arbitrários, pelo que foi dito, $Orb_\alpha(x) = M$, donde $y \in Orb_\alpha(x)$, donde existe $g \in G$ tal que $y = \alpha(g, x)$. Pela arbitrariedade de x e y, concluímos que α é transitiva. Reciprocamente, sendo α transitiva e, fixado um $x \in M$, para todo $y \in M$, existe $g \in G$ tal que $y = \alpha(g, x)$, e portanto $y \in Orb_\alpha(x)$, $\forall y \in M$ e, portanto, $Orb_\alpha(x) = M$. ■

Definição 1.12. Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. α é dita *efetiva* se dados $g, h \in G$, $g \neq h \implies \exists x \in M; \alpha(g, x) \neq \alpha(h, x)$. Tal α também costuma ser dita *fiel.* α é dita *livre* se, dados $g, h \in G$, $g \neq h \implies \alpha(g, x) \neq \alpha(h, x), \forall x \in M$. ♣

Apesar de parecidas, não se engane: a definição de ação efetiva e livre são diferentes, sendo a última mais forte, ou seja, toda ação livre é efetiva mas não vale a recíproca. Em palavras, uma ação é efetiva se transformações de simetria diferentes fazem algum ponto ficar diferente. Ou, se duas transformações de simetria deixam o espaço igual, então os elementos correspondentes do grupo são iguais. Assim como fizemos com a transitividade, forneceremos uma equivalência para a definição de livre e efetiva:

Proposição 1.17. *Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Então α é efetiva se, e somente se $\alpha(g,x) = x$, $\forall x \in M \implies g = e$. Além disso, α é livre se, e somente se $\exists x \in M$; $\alpha(g,x) = x \implies g = e$.* □

Demonstração. Suponha, primeiramente, que valha a implicação: $\alpha(a,y) = y, \forall y \in M \implies a = e$. Provarei a contrapositiva da tese. Tome $g, h \in G$ arbitrários e ponha $a = gh^{-1}$ e $y = \alpha(h,x)$. Daí $\alpha(gh^{-1}, \alpha(h,x)) = \alpha(h,x)$, $\forall x \in M \implies gh^{-1} = e \implies g = h$. Note que não há perda de generalidade ao tomarmos $\alpha(h,x)$ no lugar de y pois $\alpha(h,\cdot)$ é uma bijeção por definição (e portanto alcança todos $y \in M$). Mas, por definição de ação, $\alpha(gh^{-1}, \alpha(h,x)) = \alpha(g,x)$. E portanto temos a seguinte implicação: $\alpha(g,x) = \alpha(h,x), \forall x \in M \implies g = h$. A tese segue da arbitrariedade na escolha de g e h. Reciprocamente, suponhamos que α seja efetiva. Suponhamos também que, para algum g, $\alpha(g,x) = x$, $\forall x \in M$. Quero provar que $g = e$. Para tanto, notemos que $\alpha(e,x) = x$, $\forall \ inM$, pela definição de ação. Assim, a hipótese acima enunciada pode ser reescrita como: $\alpha(g,x) = \alpha(e,x)$, $\forall x \in M$. Ora, da efetividade de α segue que $g = e$.

Em segundo lugar, suponhamos que α seja livre. Isto é, que se existe $x \in M$ tal que $\alpha(g,x) = \alpha(h,x)$, então $g = h$. Suponhamos também que existe $x \in M$ tal que $\alpha(g,x) = x$. Provemos que $g = e$. Note que, por definição de ação, $x = \alpha(e,x)$. Daí, existe $x \in M$ tal que $\alpha(g,x) = \alpha(e,x)$. Mas como a ação é livre, chegamos que $g = e$. Reciprocamente, suponhamos que valha a implicação: $\exists x \in M; \alpha(g,x) = x \implies g = e$. Suponhamos também que exista x tal que $\alpha(g,x) = \alpha(h,x)$. Quero provar que $g = h$. Note que $\alpha(g,x) = \alpha(h,x) \implies \alpha(gh^{-1}, \alpha(h,x)) = \alpha(h,x)$. Por hipótese, isso nos dá $gh^{-1} = e$, donde $g = h$, concluindo nossa demonstração. ■

Essa proposição nos dá, em palavras, o seguinte: o uma ação é efetiva se a única transformação que deixar *tudo* igual for a identidade; uma ação é livre se a única transformação que deixar *alguma coisa* igual for a identidade.

Note que $\ker \Pi$ consiste de todos os elementos do grupo que agem em X como a identidade. Assim, uma ação ser efetiva é equivalente a $\ker \Pi = \{e\}$. Caso contrário, há uma redundância na ação: vários elementos agindo da mesma maneira. Essa redundância sempre pode ser eliminada através do quociente $G/\ker(f)$, cuja ação será sempre efetiva. Assim, toda ação dá origem a uma ação efetiva. Lembrando que o kernel de um homomorfismo ser a identidade corresponde ao homomorfismo ser um monomorfismo. Isso esclarece no nome "fiel", já que na matemática este adjetivo costuma corresponde a algum tipo de injetividade.

Proposição 1.18. *Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. α é simplesmente transitiva se, e somente se α é transitiva e livre.* □

Demonstração. Suponha que α é transitiva e livre. Então, dados $x, y \in M$, existe $g \in G$ tal que $\alpha(g,x) = y$. Daí, se $\alpha(h,x) = y$, teremos $\alpha(g,x) = \alpha(h,x)$, para algum x. Como α é livre, isso nos dá $g = h$, donde g é o único elemento tal que $\alpha(g,x) = y$, o que demonstra ser α transitiva.

Reciprocamente, suponha que α é simplesmente transitiva. Isto é, dados $x, y \in M$, se $\alpha(g,x) = y$ e $\alpha(h,x) = y$, então $g = h$ (além do mais, tal g sempre existe). Ora, que α é transitiva é óbvio. Dado $x \in M$ e $g \in G$, tome $y = \alpha(g,x)$. Daí valerá que se $\alpha(h,x) = y$, isto é, que se $\alpha(g,x) = \alpha(h,x)$ para algum x, então $g = h$, donde α é livre. ■

Definição 1.13. Seja G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Um subconjunto $S \subset M$ é dito invariante se $\alpha(g, s) \in S$, para todo $s \in S$ e $g \in G$.

Proposição 1.19. *Dado $x \in G$, $Orb(\alpha, x)$ é um conjunto invariante. A união e intersecção arbitrária de conjuntos invariantes é invariante. Um conjunto é invariante se, e somente se, é a união de órbitas.*

Demonstração. Tome $y \in \mathrm{Orb}(\alpha, x)$. Então $\mathrm{Orb}(\alpha, x) = \mathrm{Orb}(\alpha, y)$. Assim, dado qualquer $g \in G$, $\alpha(g, y)$ está na órbita de y por definição de órbita, e portanto na órbita de x. A conclusão segue pela arbitrariedade de y e g.

Seja $(X_i)_{i\in I}$ uma família arbitrária de conjuntos invariantes e sejam X^*, X_* a união e a intersecção desta família, respectivamente. Tome $x \in X^*$ e $g \in G$. $x \in X_i$ para algum i, e portanto $y = \alpha(g, x) \in X_i$ pois este é invariante, donde $y \in X^*$. Se, por outro lado, $x \in X_*$, então $x \in X_i$ para todo i. Como cada um deles é invariante, $y = \alpha(g, x) \in X_i$ também para todo i, donde $y \in X_*$.

Já vimos que a órbita de um ponto é invariante, donde qualquer união de órbita também o é, pelo parágrafo anterior. Agora, tome X um conjunto invariante e escrevamos-no como uma tal união. É óbvio que $X \subset \bigcup_{x\in X} \mathrm{Orb}(\alpha, x)$, mas provaremos também a inclusão recíproca. Tome um elemento y da união. Por construção, $y \in \mathrm{Orb}(\alpha, x)$, para algum $x \in X$, isto é, $y = \alpha(g, x)$ para algum $g \in G$. Então fica claro que $y \in X$ pela invariância deste conjunto. ■

Corolário 1.19.1. *Um conjunto ser invariante é equivalente a valer a seguinte implicação: $x \in X \implies Orb(\alpha(g, x) \subset X$.*

Corolário 1.19.2. *Seja G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Então a ação é transitiva se, e somente se, os únicos subconjuntos invariantes forem $\emptyset$ e M.*

Demonstração. Suponha que a ação seja transitiva e tome um conjunto invariante não vazio. Mostrarei que esse conjunto é M. Tome y em M qualquer. Como o conjunto é não vazio, existe x em M. Como a ação é transitiva, existe g com gx = y. Como o conjunto é invariante, deve-se ter y em X.

Suponha agora que a ação não é transitiva. Então devem existir, necessariamente, $x, y \in X$ tais que não haja nenhum $g \in G$ com $gx = y$. Neste caso, a órbita de x é um conjunto invariante que não é vazio, pois tem x, não é o todo, pois não tem y, e é invariante pelo simples fato de ser uma órbita. ■

Definição 1.14. (grupo de isotropia). Seja G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. Se $g \in G$ e $m \in M$ tais que $\alpha(g, m) = m$, então m é dito ser um ponto fixo de g. Seja $X \subset M$. O conjunto $\{g \in G; \alpha(g, x) \in X, \forall x \in X\}$ é chamado de *grupo estabilizador de G com relação a X* ou *grupo de isotropia* e denotado por $G_{\alpha,X}$ ou G_X quando a ação estiver subentendida. Assim, X é invariante se, e somente se $G_X = G$. ♣

Proposição 1.20. *Sejam G um grupo, $\alpha : G \times M \to M$ uma ação de G em M e $X \subset M$. Então G_X é um subgrupo de G.* □

Demonstração. Suponha que $g, h \in G_X$. Então $\alpha(g, x) = x$, $\forall x \in X$ e também $\alpha(h, x) = x$, $\forall x \in X$. Daí, dado $x \in X$ arbitrário, $\alpha(gh, x) = \alpha(g, \alpha(h, x)) = \alpha(g, x) = x$, donde $gh \in G_X$. Ainda, $\alpha(e, x) = x \implies \alpha(g^{-1}g, x) = x \implies \alpha(g^{-1}, \alpha(g, x)) = \alpha(g^{-1}, x) = x$. ■

Proposição 1.21. *Sejam G um grupo e $\alpha : G \times M \to M$ uma ação de G em M. A ação α é livre se, e somente se $G_x = \{e_G\}$, $\forall x \in X$.* □

Demonstração. Suponha que α é livre. Dado $g \in G_X$, temos que $\alpha(g, x) = x$. Porém, pela proposição 1.17, como α é livre, isto implica que $g = e$, donde $G_X = \{e\}$.

Reciprocamente, suponha que temos $G_X = \{e\}$. Então $\alpha(g, x) = x$ implica que $g \in G_X$ e portanto $g = e$. Temos então a implicação $\alpha(g, x) = x \implies g = e$, que pela proposição 1.17 é o mesmo que dizer que α é livre. ■

Por fim, citaremos um resultado evidente que usa a noção de classe de conjugação com o estudo de pontos fixos:

Proposição 1.22. *Seja G um grupo agindo em um conjunto X. Se $g \in G$ tem n (resp. infinitos) pontos fixos em X, então todo elemento conjugado a g também tem n (resp. infinitos) pontos fixos em X.* □

A partir de agora, usaremos a notação mais simplificada $Gx = Orb_\alpha(x)$, consistente com a notação $\alpha(g, x) = gx$, introduzida anteriormente.

Proposição 1.23. *A função $\phi : G/G_x \to Gx$ dada por $\phi(gG_x) = gx$ está bem definida e é uma bijeção*

Demonstração. Para mostrar que a função está bem-definida, mostraremos que $gG_x = hG_x \implies gx = hx$. Como $gG_x = hG_x$ e $e \in G_x$, temos que $ge = g = ha$, para algum $a \in G_x$. Daí, $gx = (ha)x = h(ax) = hx$, pois $a \in G_x$.

A sobrejetividade é evidente. Para mostrar a injetividade, mostraremos que $gx = hx \implies gG_x = hG_x$. Com efeito, se $gx = hx$, então $h^{-1}gx = x$, donde $h^{-1}g \in G_x$. Daí, $h(h^{-1}g) \in hG_x$. Por outro lado, $h(h^{-1}g) = g = ge \in gG_x$, donde gG_x e hG_x se intersectam. Como são classes de equivalência, temos que $gG_x = hG_x$. ■

A seguir, veremos alguns exemplos importantes.

Exemplo 1. *Seja M um conjunto e Sym(M) o grupo de simetrias de M. Então a função $\alpha(f, x) = f(x)$ é uma ação de grupos. Note que esta ação é transitiva, fiel mas não livre (se M tiver mais de um elemento). A órbita de todo elemento é M.*

Exemplo 2. *Subgrupos de Sym(G) induzem uma ação naturalmente através da restrição da ação supracitada. No caso de G ser um grupo, há alguns subgrupos relevantes. Enfatizo, por exemplo, Aut(G) $\subset$ Sym(G), de automorfismos de G. Há também o dos automorfismos internos, Inn(G), que serão vistos na próxima seção.*

Exemplo 3. *Seja G um grupo. Há uma ação canônica $\alpha : G \times G \to G$ dada por $\alpha(g, h) = gh$. Note que esta ação é simplesmente transitiva e a órbita de cada elemento é G*

Exemplo 4. *À semelhança do que foi feito no último exemplo, podemos definir uma ação análoga em $G\backslash H$. Não iremos exigir que H seja normal, e portanto $G\backslash H$ deve ser visto apenas como conjunto de cosets (esquerda ou direita, não faz diferença), e não como um grupo (quociente). Esta ação é transitiva.*

1.7 Comutatividade

O objetivo desta seção é usar as ferramentas que já temos para medir o quão abeliano um grupo é. Começaremos definindo o centro de um grupo (ou centralizador).

Definição 1.15. Seja G um grupo. O *centro de G* é o subconjunto de G definido por: $Z(G) := \{z \in G; zg = gz, \forall g \in G\}$. ♣

Em palavras, o centro de um grupo é o conjunto dos elementos que comutam com todos os outros. Enfatizo que este conjunto nunca é vazio, pois contém sempre o elemento neutro de G. Mais ainda:

Proposição 1.24. *Sejam G um grupo e $Z(G)$ seu centro. Vale que $Z(G)$ é um subgrupo normal de G.* □

Demonstração. Suponha que $x, y \in Z(G)$. Quero provar que $xy \in Z(G)$. Para tanto, tome $g \in G$ arbitrário. Pela associatividade, e como $y \in Z(G)$, $xyg = xgy$. Como $x \in Z(G)$, $xgy = gxy$. Assim, $(xy)g = g(xy)$. Pela arbitrariedade de g, resta-nos que $xy \in Z(G)$. Provaremos também que $x^{-1} \in Z(G)$. Ora, $xg = gx \implies g = x^{-1}gx \implies gx^{-1} = x^{-1}g$. Concluímos que $Z(G)$ é de fato subgrupo de G.

Falta provar que $Z(G)$ é normal. Ora, dados $z \in Z(G)$ e $g \in G$, $gzg^{-1} = gg^{-1}z = z \in Z(G)$. Isto completa nossa demonstração. ■

Como o leitor pode imaginar, enfatizar que o centro de um grupo é um subgrupo normal revela que iremos tomar o quociente do grupo pelo centro em algum momento. O leitor está correto.

Definição 1.16. Seja G um grupo. Dado $g \in G$, a função $\phi_g : G \to G$ definida por $\phi_g(h) = ghg^{-1}$ é chamada de *automorfismo interno*. O conjunto de todos automorfismo internos de um grupo G é denotado por $\mathrm{Inn}(G)$. Um automorfismo que não é interno é dito externo. ♣

Proposição 1.25. *Os automorfismos internos de um grupo G são de fato automorfismos. Além disso, Inn(G) é um subgrupo normal sob a operação de composição.*

□

Demonstração. Precisamos mostrar que, dado $g \in G$, ϕ_g é um isomorfismo. Primeiramente, vamos mostrar que $\phi_g(xy) = \phi_g(x)\phi_g(y)$. De fato, por definição, $\phi_g(xy) = gxyg^{-1} = gxeyg^{-1}$, onde e é o elemento neutro. Como $g^{-1}g = e$, $\phi_g(xy) = gxg^{-1}gyg^{-1} = (gxg^{-1})(gyg^{-1}) = \phi_g(x)\phi_g(y)$. Agora já sabemos que ϕ_g se trata de um endomorfismo. Só falta ver que ϕ_g é uma bijeção. Seja $y \in G$. Tome $x = g^{-1}yg \in G$. Então $\phi_g(x) = y$. De fato, $\phi_g(x) = gg^{-1}ygg{-}1 = eye = y$, e portanto ϕ_g é sobrejetiva. Suponha agora que $x \neq y$. Então $gx \neq gy \implies gxg^{-1} \neq gyg^{-1} \implies \phi_g(x) \neq \phi_g(y)$, donde ϕ_g é injetiva e, portanto, um automorfismo.

Falta mostrar que $\mathrm{Inn}(G)$ é um grupo. Inicialmente, mostraremos que a operação de composição está bem definida neste conjunto. Seja ϕ_g e ϕ_h elementos de $\mathrm{Inn}(G)$. Note que $\phi_g \circ \phi_h : G \to G$ é dada por $x \mapsto ghxh^{-1}g^{-1} = (gh)x(gh)^{-1} = \phi_{gh}(x)$, donde $\phi_g \circ \phi_h = \phi_{gh} \in \mathrm{Inn}(G)$. Automaticamente temos que esta operação é associativa, pois a composição de funções o é. Afirmamos que ϕ_e (onde e é o elemento neutro de G) é o elemento neutro de $\mathrm{Inn}(G)$. De fato, como vimos, $\phi_g \circ \phi_e = \phi_{ge} = \phi_g = \phi_{eg} = \phi_e \circ \phi_g$. Finalmente, dado $\phi_g \in \mathrm{Inn}(G)$, mostraremos que seu elemento oposto é ϕ_h, onde $h = g^{-1}$. Ora, $\phi_g \circ \phi_h = \phi_{gh} = \phi_e$, o elemento neutro.

Por fim, seja $N \in \mathrm{Inn}(G)$ e $T \in \mathrm{Aut}(G)$, mostrarei que $T^{-1}NT \in \mathrm{Inn}(G)$. Ora, por hipótese existe $g \in G$ tal que $N(x) = g^{-1}xg$, donde $(T^{-1}NT)(x) = T^{-1}(N(T(x))) = T^{-1}(g^{-1}T(x)g) = T^{-1}(g^{-1})T^{-1}(T(x))T^{-1}(g)$, pois T é um homomorfismo. Assim, $(T^{-1}NT)(x) = (T^{-1}(g))^{-1}xT^{-1}(g)$, donde $T^{-1}NT$ é interno. ■

Definição 1.17. O quociente $\mathrm{Aut}(G)\backslash\mathrm{Inn}(G)$ é chamado de grupo de automorfismos externos e denotado por $\mathrm{Out}(G)$

Proposição 1.26. *Sejam G um grupo e $Inn(G)$ o grupo de automorfismos internos de G. Então a função $\theta : G \to Inn(G)$ que associa a cada $g \in G$ o automorfismo interno ϕ_g, isto é, $\theta(g) = \phi_g$, é um epimorfismo.* □

Demonstração. Para provar que é um homomorfismo, basta ver que $\theta(gh) = \phi_{gh} = \phi_g \circ \phi_h$ (a última igualdade está provada na demonstração da proposição 1.25). Ademais, por definição de $\mathrm{Inn}(G)$, temos que se $f \in \mathrm{Inn}(G)$, então existe $g \in G$ tal que $f = \phi_g = \theta(g)$, por definição de θ. E portanto, θ é sobrejetiva. ■

Proposição 1.27. *Seja G um grupo. $g \in Z(G) \iff \phi_g = \phi_e$. Em outras palavras,* $\ker\theta = Z(G)$. □

Demonstração. $g \in Z(G) \implies \phi_g(x) = gxg^{-1} = gg^{-1}x = x = exe^{-1} = \phi_e(x)$. Reciprocamente, $\phi_g = \phi_e \implies \phi_g(x) = x, \forall x \in G \implies gxg^{-1} = x \implies gx = xg, \forall x \in G$.

A segunda afirmação se prova com a seguinte sequências de desigualdades: $g \in \ker\theta \iff \theta(g) = \phi_e \iff \phi_g = \phi_e \iff g \in Z(G)$. ■

A proposição acima mostra que um elemento de G faz parte do centro se, e somente se o homomorfismo θ definido anteriormente o mapeia para a identidade (i.e. para o elemento neutro de $\mathrm{Inn}(G)$). Vemos então que, quanto mais abeliano um grupo é, maior é seu centro e mais elementos são mapeados por θ para a identidade. Uma outra forma de ver isto é a seguinte: quanto menos abeliano é um grupo, mais automorfismos internos existem (distintos da identidade). Desta forma, o tamanho do centro de um grupo é "inversamente proporcional"ao tamanho do grupo de automorfismos internos deste. A intuição nos diz, então, que deve existir uma certa relação entre $G/Z(G)$ e $\text{Inn}(G)$ e é precisamente isto que o corolário abaixo diz. Esta relação é, na verdade, um isomorfismo!

Corolário 1.27.1. *$G/Z(G) \cong Inn(G)$.* □

Demonstração. Basta aplicar as Proposições 1.26 e 1.27 no Corolário 1.13.1. Existirá um isomorfismo ψ entre $G/\ker\theta = G/Z(G)$ e $\text{Inn}(G)$ dado por $\psi(gZ(G)) = \theta(g) = \phi_g$. ■

1.8 Geradores

Antes de passarmos para a próxima definição, convém lembrarmos que a intersecção arbitrária de subgrupos de um grupo ainda é um subgrupo. Notamos também que o mesmo não vale para uniões de subgrupos.

Definição 1.18. Seja G um grupo e $X \subset G$ um subconjunto. O *subgrupo de G gerado por X* e denotado por $\langle X \rangle$ é definido como:

$$\langle X \rangle := \bigcap_{S<G;X\subset S} S$$

O subgrupo gerado por $X = \emptyset$ é $\langle X \rangle = \{e\}$.

♣

Como a intersecção arbitrária de subgrupos é subgrupo, temos automaticamente que $\langle X \rangle$ é subgrupo de G.

Agora, trabalharemos para apresentar uma definição equivalente de subgrupo gerado, começando com a definição de *palavra*:

Definição 1.19. Sejam G um grupo e $X \subset G$ um subconjunto não-vazio. Um elemento de G é dito uma palavra de X quando pode ser escrito como $x_1...x_n$, onde x_i ou x_i^{-1} pertence a X, para todo i natural entre 1 e n. ♣

Isso nos dá a seguinte:

Proposição 1.28. *Sejam G um grupo e X um subconjunto não vazio de G. Vale que $\langle X \rangle$ é o conjunto de todas as palavras de X.* □

Demonstração. Seja W o conjunto de todas as palavras de X. É simples ver que W forma um subgrupo. Mais trivial ainda é ver que este subgrupo contém X. Portanto, $\langle X \rangle \subset W$, por definição. Reciprocamente, seja $w \in W$ arbitrário. Dado um subgrupo S que contém X, por ser em particular um grupo, é fechado com relação a operação de grupo e de elemento oposto. Como w é justamente formada por elementos de X (e portanto de S) e seus elementos inversos, então teremos $w \in S$. Pela arbitrariedade de w e S temos que W está contido na interseção de todos os subgrupos que contenham X. Isto é, $W \subset \langle X \rangle$. ■

Definição 1.20. Seja G um grupo e $X \subset G$. Se $\langle X \rangle = G$, dizemos que X *gera* G e que os elementos de X são *geradores* de G. Se existe algum X finito que gera G, dizemos que G é finitamente gerado. ♣

Definição 1.21. Seja G um grupo. Se existir $x \in G$ tal que $G = \langle x \rangle$, dizemos que G é um grupo cíclico. ♣

Definição 1.22. Seja G um grupo e $x \in G$ um elemento. Então a ordem de x é definida como a ordem do grupo $\langle x \rangle$. ♣

Proposição 1.29. *Seja G um grupo. $x \in G$ tem ordem infinita se, e somente se todas as potências de x são distintas (pela proposição anterior* (1.28)*, o conjunto das potências de x empata com $\langle x \rangle$).* □

Demonstração. Se todas as potências de x são distintas, $\langle x \rangle$ é obviamente infinito. Reciprocamente, iremos provar a contra-positiva. Isto é, supondo que existam potências iguais, digamos $x^m = x^l$, provaremos que $\langle x \rangle$ é finito. Sem perda de generalidade, consideremos $m > l$. Então teremos que $x^{m-l} = e_G$. Daí, o conjunto dos naturais tais que $x^k = e_G$ é não vazio. Pelo princípio da boa ordenação, podemos achar o menor natural que valha a igualdade. Denotaremos este natural por n. Pelo algoritmo da divisão, tomando um inteiro m qualquer, podemos reescrevê-lo como $m = nq + r$, com $0 \leqslant r \leqslant n-1$. Então $x^m = (x^n)^q + x^r = (e_G)^q + x^r = x^r$, o que mostra que $\langle x \rangle = \{e_G, x, ..., x^{n-1}\}$, donde x tem ordem finita. ■

Corolário 1.29.1. *Se G é um grupo finito, então, dado $x \in G$, existe um inteiro positivo n tal que $x^n = x^{-1}$.* □

Demonstração. Basta ver que, como G é finito e como $\langle x \rangle \leqslant G$, devem existir m, l tais que $x^m = x^l$. Pelo que foi falado na demonstração da proposição, concluiremos que $\langle x \rangle = \{e_G, x, ..., x^{n-1}\}$. Como $x^{-1} \in \langle x \rangle$, segue a asserção. ■

Corolário 1.29.2. *Se G é um grupo finito, vale que $\langle x \rangle$ é o conjunto cujos elementos são da forma $x_1^{\epsilon_1} ... x_n^{\epsilon_n}$, com $\epsilon_i \geqslant 0$, para todo i.* □

1.9 Grupos Livres

Definição 1.23. Sejam F um grupo, X um conjunto não-vazio e $\sigma : X \to F$ uma função. Então (F, σ) é dito livre em X se, para todo grupo G e toda função $\alpha : X \to G$, existir e for único um homomorfismo $\beta : F \to G$ tal que $\alpha = \beta \circ \sigma$.

♣

Esta definição pode parecer muito estranho a uma primeira vista. Gastaremos um tempo buscando uma intuição para ela. Para tanto, faremos uma analogia com espaços vetoriais.

A base de espaço vetorial possui uma característica singular. Qualquer transformação linear fica unicamente determinada uma vez que se sabe o valor que esta assume nos elementos de uma base. Isto significa que há a total *liberdade* de escolhermos os valores da transformação nos vetores da base, sem correr o risco de chegar numa inconsistência. Isto se deve ao fato de que os vetores de uma base não satisfazem relações desnecessárias, só satisfazem as relações impostas pela definição de espaço vetorial. Este fato é sinônimo do fato de que os vetores são linearmente *independentes*. Por exemplo, se os vetores u, v, w satisfizerem a relação de dependência $2u + v + 3w = 0$, então não podemos escolher arbitrariamente os valores de uma transformação linear $T(u), T(v), T(w)$ sem corrermos o risco de sermos contraditórios. $T(u), T(v), T(w)$ devem necessariamente satisfazer a mesma relação. Agora, basta trocarmos "espaço vetorial"por grupo e "transformação linear"por homomorfismo. Assim, um grupo F é livre sobre um conjunto de geradores X se qualquer homomorfismo de domínio F fica unicamente determinado pelos valores que assume nos elementos de X (que podem ser escolhidos "livremente"). Isso é análogo a X ser uma base de F. Ser "livre"significa que os elementos de X são "livres"de relações desnecessárias, isto é, que não aquelas impostas pelo fato de F ser grupo.

A principal diferença reside no fato de que todo espaço vetorial possui uma base, enquanto que nem todo grupo é livre sobre algum conjunto de geradores. Quanto um grupo é livre, então é possível traçar um análogo de "dimensão"no caso de espaços vetoriais, e que é chamado de posto.

Esta questão possui algumas facilidades de ser tratadas no caso de grupos abelianos. Isto decorre do fato de que um grupos abelianos podem ser vistos como módulos sobre $\mathbb{Z}$, isto é, são "quase"espaços vetoriais. Daí, um grupo abeliano livre de posto n é isomorfo a $\mathbb{Z}^n$.

Note que, implicitamente, esta discussão supôs que $X \subset F$. Na verdade, esta suposição não é exatamente injustificada. Sempre que um grupo for livre em um conjunto X, há, no mínimo, uma cópia de X dentro de F, que difere no máximo pela natureza dos elementos. Formalmente, a função $\sigma : X \to F$ é injetiva:

Proposição 1.30. *Se G é um grupo livre em X, então a função $\sigma : X \to F$ da definição é injetiva.* □

Demonstração. Será por absurdo. Suponha que σ não é injetiva. Então existem $x_1, x_2, \in X$ distintos tais que $y = \sigma(x_1) = \sigma(x_2)$. Tome $G = \{g_1, g_2\}$ composto de dois elementos distintos e uma função $\alpha : X \to G$ tal que $\alpha(x_1) = g_1$ e $\alpha(x_2) = g_2$. Tomemos β tal que $\beta \circ \sigma = \alpha$, que existe pela definição de grupo livre. Entretanto, $(\beta \circ \sigma)(x_1) = \beta(\sigma(x_1)) = \beta(\sigma(x_2)) = (\beta \circ \sigma)(x_2)$, o que é um absurdo, pela definição de α. ■

Não é difícil ver, a partir daí, que F seria livre também em $\mathrm{Im}\sigma$, substituindo σ pela função inclusão $i : \mathrm{Im}\sigma \to F$, donde todo grupo livre é livre sobre algum subconjunto. É possível mostrar que $\mathrm{Im}\sigma$ gera F. Mais que isso, um outro aspecto importante sobre um grupo ser livre é que F é, em certo sentido, o "maior"grupo que $\mathrm{Im}\sigma$ pode gerar. Enunciaremos o teorema sem demonstrá-lo, pois a demonstração possui diversos detalhes técnicos chatos, apesar de que a ideia da demonstração é simples, e sera dada.

Proposição 1.31. *Seja X um conjunto não-vazio. Então existe um grupo F e uma função $\sigma : X \to F$ tal que (F, σ) é livre em X e $F = \langle Im\sigma \rangle$* □

A demonstração é completamente por construção. Primeiro, devemos tomar um conjunto X^{-1}, disjunto a X e com a mesma cardinalidade, de forma que a cada elemento de X, corresponderemos um de X^{-1} para cumprir o papel de seu elemento inverso. Daí, tomaremos o conjunto S de todas as palavras (isto é, sequências finitas de elementos) de $X \cup X^{-1}$, sendo que a palavra "vazia" cumprirá o papel de identidade. O produto de duas palavras é definida como a concatenação delas e o inverso se faz trocando cada elemento da palavra pelo seu inverso e invertendo também a ordem. S é quase F. Temos que fazer apenas mais um ajuste. Do jeito que construímos, palavras como $x^{-1}xa$, $xx^{-1}a$ e a são diferentes uma da outra, enquanto que num grupo, elas seriam iguais. O resto do trabalho agora se reduz a unir palavras assim formando classes de equivalência, definindo operações da forma natural, provando que tudo está bem definido e de fato forma um grupo. F será o conjunto das classes de equivalência.

Por fim, deve-se provar que a função $\sigma : X \to F$ por $\sigma(x) = [x]$ faz de (F, σ) livre. Dada uma função $\alpha : X \to G$, construamos primeiro uma função $\overline{\beta} : S \to G$, do conjunto de todas as palavras, mapeando uma palavra $x_1...x_n$ para $g_1...g_n$, onde $g_i = \alpha(x_i)$. É fácil ver que elementos na mesma classe de equivalência serão mapeados para o mesmo elemento em G, donde definimos $\beta : F \to G$ por $\beta([w]) = \overline{\beta(w)}$. É fácil mostrar que tal β se torna um homomorfismo.

Capítulo 2

Topologia

2.1 Definições Básicas

Definição 2.1. Sejam (X, τ) um espaço topológico e $S \subset X$. Um ponto $p \in S$ é dito ponto de acumulação de S se toda vizinhança de p possui um elemento de S que não seja o próprio p. ♣

Em alguns lugares aparece uma outra definição de ponto de acumulação, geralmente usada com sequências: um ponto p geralmente é dito de acumulação para a sequência (x_n) se toda vizinhança de p contém infinitos elementos da sequência.

Proposição 2.1. *Seja $S \subset X$ um conjunto. Então S é fechado e discreto se, e somente se, não possui pontos de acumulação.*

Demonstração. Suponha que S tenha ponto de acumulação. Como S é fechado, o ponto de acumulação pertence a S, mas então ele não pode ser discreto. Reciprocamente, se S não fosse discreto, conteria um elemento que seria ponto de acumulação de si mesmo. Além disso, como um conjunto é fechado se, e somente se possui todos os seus pontos de acumulação, um conjunto que não possui pontos de acumulação é automaticamente fechado ■

2.2 Axiomas de Separação

Muitas propriedades não valem para todos espaços topológicos, mas valem para todos os espaços topológicos de interesse. Isso dá origem a sistemas de classificação de espaços topológicos. Um sistema de classificação bastante útil refere-se a possibilidade de separar pontos ou conjuntos do espaço com abertos, e cada tipo de espaço é geralmente denotado por T seguido de um número.

Definição 2.2. Um espaço topológico T_0 (ou Kolmogorov) é um espaço tal que, dados dois pontos quaisquer, pelo menos um deles tem uma vizinhança aberta que não contém o outro.

Um espaço topológico T_1 é um espaço tal que, dados dois pontos quaisquer, ambos possuem uma vizinhança aberta contendo ele e não o outro. ♣

O último é também chamado de Frèchet, mas este uso não é recomendado pois "espaço de Frèchet"é algo completamente diferente em análise funcional, e é o significado mais usual do nome. Veremos uma importante propriedade de espaços $T1$.

Proposição 2.2. *Dado um espaço topológico T_1, um ponto p é de acumulação de S se, e somente se, toda vizinhança aberta de p possui infinitos pontos de S.*

□

Demonstração. A volta é direta. Reciprocamente, suponha, por absurdo, que exista $S \subset X$, p ponto de acumulação de S e V vizinhança aberta de p tal que $V \cap S$ é finito. Para cada ponto desta intersecção, seja U_i a vizinhança aberta de p que não contém x. Tal vizinhança existe pelo fato de estarmos num espaço T_1. Tome agora $U = \bigcap U_i$, que é uma intersecção finita de abertos, donde U é uma vizinhança aberta. Ora, $U \cap V$ ainda é uma vizinhança aberta de p que, entretanto, não possui nenhum ponto de S, o que é um absurdo pois p é de acumulação ■

Um espaço ser T_2 é o mesmo de ser Hausdorff. Enfatizo que ser Hausdorff é geralmente o mínimo que geralmente se pede de um espaço topológico, principalmente em geometria. A proposição a seguir sintetiza algumas propriedades importantes dos espaços Hausdorff:

Proposição 2.3. *Seja X um espaço topológico Haudorff. Então, se $K \subset X$ é compacto, K é fechado. Além disso, se X é compacto, todo subconjunto fechado de X também é compacto. Também, o limite de uma rede é único se, e somente se, o espaço for Hausdorff.* □

Um espaço T_3, ou regular, é um espaço onde, para qualquer ponto e qualquer subconjunto fechado, há abertos disjuntos contendo cada um deles. Um espaço T_4, ou normal, é análogo ao T_3, com a diferença que a definição usa dois conjuntos fechados no lugar de um ponto e um conjunto fechado. Todo espaço métrico é normal (portanto todo espaço topológico metrizável também).

2.3 Bases e Subbases

Definição 2.3. Seja (X, τ) um espaço topológico. Um conjunto $B \subset \tau$ é dito uma base da topologia τ se todo aberto de τ pode ser escrito como uma união de elementos de B. Isto é:

$$\forall U \in \tau, \exists A_\lambda \subset B; U = \bigcup_{A \in A_\lambda} A$$

♣

A proposição seguinte será bem útil nesta seção.

Proposição 2.4. *Seja (A_i) uma família de conjuntos. Então, existir uma subfamília (A_λ) tal que $U = \bigcup A_\lambda$ é equivalente a dizer que, para todo $x \in U$, existe λ tal que $x \in A_\lambda \subset U$.* $\square$

Demonstração. A ida é bem direta. Para a volta, para cada $x \in U$, tome A_x tal que $x \in A_x \subset U$, que existe por hipótese. Então, $U \subset \bigcup A_x \subset U$, pois $A_x \subset U$ para todo x, donde temos a igualdade entre os conjuntos. $\blacksquare$

E ela nos fornece de forma direta uma equivalência para a definição de base.

Proposição 2.5. *Seja (X, τ) um espaço topológico. Um conjunto $B \subset \tau$ é uma base de τ se, e somente se, para todo $U \in \tau$ e todo $x \in U$, existe $V \in B$ tal que $x \in V \subset U$.* $\square$

Veremos agora que uma base não pode formar topologias diferentes

Proposição 2.6. *Seja τ uma topologia sobre X e B uma base de τ. Se B é uma base de σ, então $\tau = \sigma$.* $\square$

Demonstração. Provaremos que $\tau \subset \sigma$ e o outro lado seguirá por simetria. Tome $U \in \tau$. Como B é base de τ, $U = \bigcup A_\lambda$ com $A_\lambda \in B$, para todo λ. Mas B também é base de σ, e aí, por definição, todo elemento de B é elemento de σ. Em particular, $A_\lambda \in \sigma$ para todo λ. Como σ é uma topologia, é fechada por uniões arbitrárias. Em particular, $U = \bigcup A_\lambda \in \sigma$, como queríamos demonstrar. $\blacksquare$

Apesar de uma base determinar univocamente uma topologia, uma topologia não determina univocamente uma base, isto é, uma mesma topologia pode ter mais de uma base diferentes, o que pode ser facilmente verificado com a construção de exemplos.

Note que, até agora, sempre estamos supondo que um conjunto é base de *alguma* topologia, ou seja, estamos tomando bases a partir de topologias "prontas. Entretanto, podemos considerar o problema análogo: construir topologias a partir de bases. Entretanto, dado um conjunto X, nem todos os subconjuntos de $\mathcal{P}(X)$ podem ser base de alguma topologia. Com efeito, tome o conjunto $X = \{1, 2, 3\}$ e $B = \{\{1, 2\}, \{1, 3\}\}$. Suponha que B seja a base para alguma topologia τ de X. Então $\{1, 2\}$ e $\{1, 3\}$ são abertos desta topologia. Como a intersecção de abertos é aberto, $\{1\}$ é um aberto desta topologia. Sendo B uma base, $\{1\}$ pode ser escrito como uma união de elementos de B, o que é um absurdo.

A seguir, veremos um critério necessário e suficiente para que um subconjunto do conjunto das partes de um conjunto seja base para *alguma* topologia.

Proposição 2.7. *Seja X um conjunto e $B \subset \mathcal{P}(X)$. Então, existe alguma topologia τ tal que B é base de τ se, e somente se:*

$$\bigcup_{A \in B} A = X$$

E também, dados $U, V \in B$ e $x \in U \cap V$, existe $W \in B$ tal que $x \in W \subset (U \cap V)$. $\square$

Demonstração. Suponha que uma tal topologia exista. Como $B \subset \mathcal{P}(X)$, é óbvio que $\bigcup A \subset X$, basta provar a outra continência. Como τ é topologia, $X \in \tau$, e aí existe alguma família de elementos de B, (A_λ), tal que $\bigcup A_\lambda = X$, em particular $X \subset \bigcup A_\lambda$, que por sua vez está contido na união de todos os elementos de B. Além do mais, como B é base de τ, dados $U, V \subset B$, U, V são abertos de τ. Como topologias são fechadas por intersecções finitas, $U \cap V$ é um aberto de τ, e podemos aplicar a proposição 2.5 para mostrar que, dado $x \in U \cap V$, existe $W \in B$ tal que $x \in W \subset (U \cap V)$.

Reciprocamente, seja $B \subset \mathcal{P}(X)$ satisfazendo as propriedades citadas. Seja τ a coleção de todas as uniões arbitrárias de conjuntos de B. Mostrarei que τ é uma topologia e aí seguirá a existência que queremos mostrar, por construção.

Em primeiro lugar, $\emptyset \in \tau$ pois é a união de uma família vazia de elementos de B. Além disso, $X \in \tau$ pois a união de todos os elementos de B é X, por hipótese. Em segundo lugar, se temos uma união arbitrária de elementos de τ, cada elemento é, por definição, uma união arbitrária de elementos de B, o que nos dá uma coleção arbitrária de famílias, $(\mathcal{F}_i)$. Tomando a família formada pela união, $\mathcal{F} = \mathcal{F}_i$, é fácil ver que a união de elementos de B indexada por $\mathcal{F}$ nos dá a união arbitrária de elementos de τ, e seguirá que a união está em τ.

Por fim, provaremos que a intersecção de dois elementos de τ está em τ e uma indução nos permitirá estender o resultado para interseções finitas. Sejam $U, V \in \tau$. Então existem famílias I, J tais que $U = \bigcup U_i$ e $V = \bigcup V_j$. É fácil verificar que:

$$\left(\bigcup_i U_i\right) \cap \left(\bigcup_j V_j\right) = \bigcup_{i,j} (U_i \cap V_j)$$

Como já provamos que τ é fechada por uniões arbitrárias, basta mostrar que $C \cap D \in \tau$ para $C, D \in B$. Sabemos que para $x \in C \cap D$, existe $W \in B$ tal que $x \in W \subset (C \cap D)$. A proposição 2.4 mostra então que existe uma família de elementos de B cuja união é $C \cap D$, donde $C \cap D \in \tau$. ■

Enfatizo o fato evidente que um conjunto ser fechado por intersecções (duas a duas) implica na segunda condição para que seja base de alguma topologia, e temos que uma condição suficiente para tal é um conjunto cobrir o espaço e ser fechado por intersecção.

É comum dizermos que a topologia formada pela união arbitrária de elementos de uma base, essa que construímos na demonstração anterior, é a *topologia gerada pela base*. Vale fazer algumas observações pertinentes a essa topologia. A primeira é que, pela proposição 2.4, isso equivale à topologia cujos abertos são os conjuntos U tais que, para todo elemento x pertencente U existe um elemento B da base tal que $x \in B \subset U$. A demonstração usando esta definição fica mais legível, tente refazê-la.

Em segundo lugar, demonstramos o seguinte:

Proposição 2.8. *Se B é base de alguma topologia, o conjunto formado pelas uniões arbitrárias de elementos de B é uma topologia.* □

Pela unicidade estabelecida na proposição 2.6, concluímos que se um certo conjunto é base de uma topologia, esta topologia é necessariamente o conjunto formado pelas uniões arbitrárias de elementos da base, e não há outra topologia diferente que este conjunto continue sendo base.

Em terceiro lugar, o modo como se escreve os abertos da topologia, como uniões arbitrárias de elementos da base, não é único. Isto faz o conceito de base na topologia ser diferente daquele da álgebra linear, onde todo vetor é escrito de maneira única como combinação linear dos vetores da base.

Por último, veremos agora se esta terminologia, "topologia gerada pela base", é consistente com o significado mais familiar de "gerado".

Para tanto, dado um conjunto X e um subconjunto $B \subset \mathcal{P}(X)$, diremos que a topologia gerada por B, denotando-a por $\langle B \rangle$, é a intersecção de todas as topologias de X que contêm B. Neste caso, dizemos que B é uma *subbase* para a topologia $\langle B \rangle$. O exercício simples de que esta intersecção está sempre bem definida e é uma topologia é deixada ao leitor, juntamente com as propriedades fáceis: (i) $A \subset B \implies \langle A \rangle \subset \langle B \rangle$ e (ii) $\langle \langle B \rangle \rangle = \langle B \rangle$. Uma maneira equivalente de definir seria dizer que $\langle B \rangle$ é a menor topologia que contém B. Esta topologia é única pois coincide com a intersecção de modo que um conjunto define unicamente uma topologia de qual é subbase, mas uma topologia pode ter mais de uma subbase.

Observação útil: todas as coisas do parágrafo anterior continuam valendo trocando a palavra "topologia"por "σ=álgebra"(com exceção da nomenclatura "subbase") ou alguma outra estrutura semelhante.

Notamos que, diferentemente do conceito de base, qualquer conjunto é uma subbase de alguma topologia.

Veremos agora em quais casos uma subbase de uma topologia também é base dela. Como já foi observado, se B é arbitrário, o conjunto de todas as uniões de elementos de B não necessariamente é uma topologia, enquanto que $\langle B \rangle$ sempre é, donde podemos ver de antemão que não vale a igualdade entre os conceitos neste caso. Entretanto, a proposição seguinte garante a igualdade dada uma hipótese fraca.

Proposição 2.9. *Sejam X um conjunto, $B \subset \mathcal{P}(X)$ e τ o conjunto de todas as uniões arbitrárias de elementos de B. Se τ é uma topologia, então $\langle B \rangle = \tau$, isto é, se B é base de τ, então B também é subbase de τ.* $\square$

Demonstração. Evidentemente, $B \subset \tau$. Como τ é uma topologia, pela definição de topologia gerada por um conjunto, $\langle B \rangle \subset \tau$. Reciprocamente, se $U \in \tau$, U é a união de alguma família de elementos de B. Como $B \subset \langle B \rangle$ e topologias são fechadas por uniões, tem-se que $U \in \langle B \rangle$. $\blacksquare$

Sintetizando vários dos resultados provados neste seção, concluímos que se B é base de alguma topologia, essa topologia é a formada por uniões arbitrárias e B também é subbase.

Analisemos agora a recíproca, isto é, B é base de $\langle B \rangle$? Já vimos que há conjuntos que não são base de nenhuma topologia, então isso naturalmente não pode ser verdade. E se soubermos que B é base de *alguma* topologia?

Proposição 2.10. *Sejam X um conjunto e $B \subset \mathcal{P}(X)$ base de alguma topologia. Então B é base de $\langle B \rangle$, isto é, se ele gera, ele é base.* $\square$

Demonstração. Seja τ a topologia de qual B é base, que existe por hipótese. Por definição, $\langle B \rangle \subset \tau$. Obviamente, $B \subset \langle B \rangle$. Além disso, seja $U \in \langle B \rangle$. Então $U \in \tau$. Como B é base de τ, U pode ser escrito como uma união de elementos de B. Dada a arbitrariedade de U, temos que B é base de $\langle B \rangle$. $\blacksquare$

Os dois resultados acima indicam que, se B é base de alguma topologia, então B gerar uma topologia e ser base dela são equivalentes.

Temos ainda:

Proposição 2.11. *Seja $B \subset \mathcal{P}(X)$ e T o conjunto de todas as uniões de elementos de B. Então vale que $\langle B \rangle = \langle T \rangle$* $\square$

Vale que $\langle B \rangle \subset \langle T \rangle$ pois evidentemente $B \subset T$. Além disso, como topologia é fechada por uniões arbitrárias, $T \subset \langle B \rangle \implies \langle T \rangle \subset \langle \langle B \rangle \rangle = \langle B \rangle$, onde usamos propriedades citadas anteriormente.

Por fim, discutiremos um pouco mais sobre a definição de subbase, que pode variar ligeiramente entre autores.

Já discutimos que B é subbase de τ se, e somente se, (1) τ é a menor topologia que contém B, que equivale a dizer que (2) é a intersecção de todas as topologias que contém B. Outra maneira de definir é de que $\langle B \rangle$ (3) são todas uniões arbitrárias de todos os conjuntos que podem ser formados tomando intersecções de um número finito de elementos de B, que por sua vez é equivalente a (4) o conjunto que consiste de todas as intersecções de um número finito de elementos de B é base de $\langle B \rangle$. Para definir destas duas últimas maneiras, devemos usar a convenção de que uma intersecção indexada por um conjunto vazio dá o espaço todo, X, a fim de que $X \in \langle B \rangle$. Caso esta convenção não for adotada, (3) e (4) deixam de ser equivalentes às anteriores e passa a ser necessário fazer adaptações para que sejam. Uma maneira de consertar seria, por exemplo, dizer que o conjunto que consiste de todas as intersecções finitas de elementos de B juntamente com X forma uma base para $\langle B \rangle$. Por fim, muitos autores costumam ainda supor que B cobre X, a fim de evitar que haja confusão com relação à convenção citada. Esta última definição não é equivalente às anteriores de modo geral. Entretanto, supondo que o espaço é $T1$, as definições voltam a ser equivalentes.

Proposição 2.12. *Seja (X, τ) um espaço topológico $T1$ que contenha dois ou mais pontos. Então, se B é uma subbase de τ, B cobre X* $\square$

Demonstração. Provaremos que se B é subbase de τ mas existe $x \notin \bigcup B$, então τ não é T_1.

Seja I' o conjunto de todas as intersecções finitas de elementos de B, sem contar a intersecção "vazia". Então, dado $A \in I'$, existe $U \in B$ tal que $A \subset U \subset \{x\}^c$, donde $x \notin A$, para todo $A \in I'$. Seja agora $I = I' \cup \{X\}$. O único elemento de I que contém x é X. Sabemos que I é base de τ, donde todo elemento de τ pode ser escrito como união de elementos de I. Suponha que X não esteja nesta união, então é óbvio que a união resultante não terá x. Por outro lado, se X tiver, a união resultante será o próprio X. Assim, o único elemento de τ que possui x é X. Como X tem no mínimo dois elementos por hipótese, seja $y \in X$ com $y \neq x$. Então, não há nenhum aberto contendo x e não contendo y, donde τ não é T_1, como queríamos demonstrar. ■

2.4 Sistema de Vizinhanças e Base Local

Seja X é um espaço topológico. O sistema de vizinhanças para $x \in X$, denotado por $\mathcal{V}(x)$ é o conjunto de todas as vizinhanças de x. [1]. Para os mais familiarizados, um sistema de vizinhanças forma um filtro (isto é, não tem o vazio, tem o todo, é fechado por intersecções finitas e se um conjunto contém algum conjunto do filtro, ele também está no filtro).

Uma base para vizinhanças de x ou base local de x é um subconjunto de $\mathcal{V}(x)$ tal que toda vizinhança de x contém um elemento da base. Em espaços métricos, o conjunto $\{B(x, 1/n), n \in \mathbb{N}\}$ é uma base local para x.

2.5 Construções Usuais

2.5.1 Topologia Final e Inicial

Nesta seção abordaremos, primeiramente, formas típicas de construir topologias, como inicial, final, quociente e produto. As duas primeiras podem ser consideradas mais importantes pois é possível definir as outras em termos delas.

Estas construções, do tipo “topologia inicial”e “topologia final”são muito comuns na matemática, principalmente no que se refere a uma álgebra de conjuntos. Prova disso é o conceito de $\sigma-$álgebra inicial e final. Devido ao fato deste tipo de construção ser tão comum e tão semelhante nos diversos casos, antes de partirmos para as definições formais, generalizaremos informalmente esta construção para estruturas quaisquer onde esta faz sentido.

Sejam A e B conjuntos e $f : A \to B$. Se pudermos construir uma “estrutura”a partir de B (por “estrutura”, pode-ser ler topologia, $\sigma-$álgebra ou outros), digamos $\mathcal{B}$, então f induzirá naturalmente uma “estrutura”em A, chamada de “estrutura”inicial e definida como $\mathcal{A} := f^{-1}(\mathcal{B})$. A “estrutura”inicial é a menor estrutura que torna a função f um “morfismo”entre $(A, \mathcal{A})$ e $(B, \mathcal{B})$. (Por morfismo, pode-se ler função contínua no caso de topologia ou função mensurável no caso de $\sigma-$álgebra). Isto pode ser generalizado se tivermos uma família de espaços $(B_i, \mathcal{B}_i)$ e funções $f_i : A \to B_i$. A estrutura inicial vai ser a menor estrutura que torna a função f um morfismo entre $(A, \mathcal{A})$ e $(B_i, \mathcal{B}_\rangle)$, $\forall i$. Entretanto, em geral não basta tomar pré-imagens.

Se, por outro lado, tivermos uma “estrutura”a partir de “A”e não em B, então f pode induzir naturalmente uma “estrutura”em B chamada de “estrutura”final e definida por $\mathcal{B} := \{Y \subset B; f^{-1}(Y) \in \mathcal{A}\}$. Esta é a maior “estrutura”em B que torna f um “morfismo”entre $(A, \mathcal{A})$ e $(B, \mathcal{B})$.

[1]Neste texto, vizinhança é um conjunto qualquer que contém um aberto

Novamente, o conceito pode ser estendido para uma família de espaços mas também se torna mais adequado definir como maior estrutura que torna todas as funções morfismos e não a definição com pré-imagens.

Passemos às definições formais para o caso específico de topologia.

Definição 2.4. Sejam X um conjunto, (Y_i, τ_{Y_i}) uma família arbitrária de espaços topológicos indexadas por I e (f_i) uma família de funções indexadas pelo mesmo conjunto tal que $f_i : X \to Y_i$, para todo $i \in I$. Então, a *topologia inicial induzida por* $(f_i)_{i\in I}$ é definida como a menor topologia que torna todas (f_i) contínuas. ♣

No caso de apenas um contradomínio, isto coincidirá com a topologia $\tau_X := \{f_i^{-1}(U), U \in \tau_{Y_i}, i \in I\}$.

Definição 2.5. Sejam (X_i, τ_{X_i}) uma família arbitrária de espaços topológicos indexadas por I e (f_i) uma família de funções indexadas pelo mesmo conjunto tal que $f_i : X_i \to Y$, para todo $i \in I$. Então, a *topologia final induzida por* $(f_i)_{i\in I}$ é definida como a maior topologia que torna todas as funções contínuas

♣

No caso de apenas um domínio, isto coincidirá com a topologia $\tau_Y := \{U \subset Y; f_i^{-1}(U) \in \tau_{X_i}, i \in I\}$.

Como foi dito, estas definições nos permitirão manejar melhor outras, como o conceito de topologia induzida, quociente e produto, apresentados ambos a seguir.

2.5.2 Topologia Induzida e Quociente

Definição 2.6. Seja (X, τ) um espaço topológico e $S \subset X$. A topologia induzida por τ em X é definida por $\tau_S := \{S \cap U, U \in \tau\}$. ♣

Proposição 2.13. *Seja (X, τ) um espaço topológico e $S \subset X$. Então a topologia induzida em S corresponde à topologia inicial induzida pela função inclusão $i : S \to X$, $i(x) = x$.* ☐

Relembremos alguns conceitos que serão importantes para a definição de grupo fuchsiano.

Definição 2.7. Seja (X, τ) um espaço topológico. Dado um subconjunto $S \subset X$, $x \in S$ é dito *ponto isolado* se existe um aberto $U \in \tau$ tal que $S \cap U = \{x\}$. Se todo ponto de S é isolado, S é chamado de conjunto discreto. ♣

Proposição 2.14. *Seja (X, τ) um espaço topológico. Um subconjunto $S \subset X$ é discreto se, e somente se a topologia induzida por τ em S for a topologia discreta de S (isto é, todo subconjunto de S estiver na topologia).* ☐

Demonstração. Suponha que S é discreto. Dado $x \in S$, tomemos U_x tal que $U_x \cap S = \{x\}$. Daí, $\{x\}$ são abertos da topologia induzida, por definição. Como topologias são fechadas por uniões arbitrárias, segue que todo subconjunto de S é aberto e portanto a topologia induzida empata com a topologia discreta de S. Reciprocamente, suponha que a topologia induzida em S seja a discreta. Então, dado $x \in S$, $\{x\} \in \tau_S$. Ora, pela definição de topologia induzida, há de haver um aberto U de τ tal que $S \cap U = \{x\}$, donde x é isolado. Pela arbitrariedade de x, segue que S é discreto. ■

Definição 2.8. Sejam (X, τ_X) um espaço topológico e $\sim$ uma relação de equivalência em X. Então o espaço quociente de (X, τ_X) por $\sim$ é definido como (Y, τ_Y) onde $Y = X/\sim$, isto é, $Y = \{[x] := \{v \in X; v \sim\}; x \in X\}$ e onde $\tau_Y = \{U \subset Y; \bigcup_{[x] \in U}[x] \in \tau_X\}$. ♣

Em um certo sentido, a topologia do espaço quociente é mais "grossa". Primeiro, note que, se um certo conjunto é um aberto do espaço quociente, então também é um aberto do espaço original, ("abrindo"as classes de equivalência). Por outro lado, a topologia quociente exclui os abertos que possuem elementos sem possuir toda a classe de equivalência deste elemento. (lembre-se que, no espaço quocientado, um elemento e sua classe de equivalência se confundem). Um exemplo fácil é o seguinte. Considere o plano com a topologia usual. Considere a relação de equivalência que quocienta o plano em linhas verticais, isto é, dois pontos estão na mesma classe de equivalência se, e somente se suas abscissas forem iguais. Uma faixa do plano sem a fronteira é um aberto tanto do plano quanto do espaço quociente. Entretanto, bolas abertas não são abertos do espaço quociente pois este conjunto não contém nenhuma classe de equivalência, isto é, nenhuma linha vertical, deixando de fazer sentido falar até mesmo que esse conjunto é subconjunto do espaço quociente. De certa maneira, podemos dizer que a topologia quociente é a maior topologia "contida"na topologia usual e que os abertos façam sentido para o espaço quocientado. Esta afirmação pode ser rigorosamente expressa na seguinte proposição:

Proposição 2.15. *Sejam (X, τ_X) um espaço topológico e (Y, τ_Y) o espaço quociente de (X, τ_X) por uma relação de equivalência $\sim$. Seja $q : X \to Y$ a projeção canônica de $\sim$, isto é, a função que mapeia um elemento $x \in X$ na sua respectiva classe de equivalência $[x] \in Y$. Então τ_Y empata com a topologia final induzida por q.* □

Demonstração. Seja $\mathcal{F}$ a topologia final induzida por q. Tome $U \subset Y$. Temos que $\bigcup_{[x] \in U}[x] := \{a \in X; a \in [x], \text{ para algum } [x] \in U\} = \{a \in X; [a] \in U\}$. De fato, se a está no segundo conjunto, então $[a] \in U$. Em particular, existe $[x] \in U$ tal que $a \in [x]$, a saber, $[x] = [a]$. Se, por outro lado, a está no primeiro conjunto, então existe $[x] \in U$ tal que $a \in [x]$. Por outro lado, pelo fato de serem classes de equivalência, temos que $[a] = [x]$. Assim, $[a] \in U$ e portanto a está no segundo conjunto. Pois bem. Com isso em mãos, para completarmos a prova basta considerarmos as seguintes equivalências: $U \in \mathcal{F} \iff q^{-1}(U) \in \tau_X \iff \{x \in X; q(x) \in U\} \in \tau_X \iff \{x \in X; [x] \in U\} \in \tau_X \iff \bigcup_{[x] \in U}[x] \in \tau_X \iff U \in \tau_Y$. ■

2.5.3 Topologia Produto e das Caixas

Nosso objetivo agora é de definir uma topologia para o produto cartesiano de espaços topológicos. A forma mais natural seria a de considerar a topologia gerada pelo conjunto de todos os produtos cartesianos de abertos de cada topologia. Esta construção de fato existe, é chamada de topologia das caixas e é bastante adequada para o produto cartesiano de um número finito de espaços. Entretanto, ela passa a não se tornar mais tão adequada no caso de um produto infinito, como veremos adiante. Uma construção melhor é chamada de "topologia produto". Esta topologia empata com a topologia das caixas no caso finito, e é mais adequada no caso infinito. A topologia produto é sempre mais "grossa"(isto é, possui um número menor de abertos) do que a topologia das caixas. Passemos às definições formais.

Definição 2.9. Seja $(X_i, \tau_i)_{i \in I}$ uma coleção arbitrária de espaços topológicos. O espaço produto desta coleção é definido como (X, τ) tal que

$$X = \bigtimes_{i \in I} X_i$$

e τ é a topologia inicial induzida pela família de funções $\pi_i : X \to X_i$, onde π_i são as projeções canônicas. ♣

Para que π_j seja contínua, é necessário e suficiente que $\pi_j^{-1}(U)$ seja aberto, para todo $U \in \tau_j$. É fácil ver que $\pi_j^{-1}(U) = \bigtimes U_i$, onde $U_i = X_i$ se $i \neq j$ e $U_j = U$. Conjuntos desse tipo são chamados de cilindros abertos. Assim, para eu todas as projeções sejam contínuas, é necessário que $\pi_i^{-1}(U_i)$ esteja na topologia, para todo $U_i \in \tau_i$ e todo i. Dizer então que a topologia produto é a menor topologia que torna as projeções contínuas significa dizer que é a menor topologia que contém todos os cilindros abertos. Em outras palavras, a topologia produto é a topologia gerada pelos cilindros. Ou ainda, topologia produto é a topologia cuja sub-base é o conjunto $\mathcal{S} = \bigcup \mathcal{S}_j$, $\mathcal{S}_j = \{\bigtimes U_i, i = j \implies U_i \in \tau_j, i \neq j \implies U_i = X_i\}$. Já vimos que o conjunto das intersecções finitas de elementos de uma sub-base forma uma base. É fácil ver que a intersecção de finitos elementos de $\mathcal{S}$ é um produto cartesiano $\bigtimes U_i$ onde $U_i \neq X_i$ apenas para um número finito de i, nos quais $U_i \in \tau_i$. Tais elementos são chamados de caixas aberto e o conjunto de todos eles serão denotados por $\mathcal{B}$, que é base da topologia produto. Assim, poderíamos definir topologia produto equivalentemente como a topologia gerada pela base $\mathcal{B}$.

Vale mencionar que a construção da topologia produto é análogo a muitas outras construções "produto"na matemática.

Vamos, agora, comparar um pouco a topologia produto com a das caixas. Em primeiro lugar, note que $\mathcal{B}' = \{\bigtimes U_i, U_i \in \tau_i\}$ é base de alguma topologia, pois $X \in \mathcal{B}'$ e $\mathcal{B}'$ é fechado por intersecções, já que:

$$\left(\bigtimes U_i\right) \cap \left(\bigtimes V_i\right) = \bigtimes U_i \cap V_i$$

É fácil ver que para produtos cartesianos finitos elas coincidem. Vejamos agora que para infinitos não coincidem. Seja $\mathbb{R}^{\mathbb{N}}$ o conjunto das sequências infinitas enumeráveis de números reais. O conjunto $(0,1)^{\mathbb{N}}$ é obviamente aberto na topologia das caixas, mas será que o é na topologia produto? suponha que sim. Então este conjunto é uma união arbitrária de elementos de $\mathcal{B}$. Não pode ser uma união vazia pois o conjunto não é vazio. Assim, existe $B \in \mathcal{B}$ tal que $B \subset (0,1)^{\mathbb{N}}$, pois os conjuntos que compõem uma união estão contidos nela. Por definição, B é um produto cartesiano com infinitos termos iguais a $\mathbb{R}$. Mas é um absurdo que um conjunto tal esteja contido em $(0,1)^{\mathbb{N}}$.

Depois de termos visto que as topológicas de fato são diferentes, daremos o principal motivo pelo qual a topologia produto é mais interessante:

Proposição 2.16. *Seja $f : X \to \bigtimes Y_i$ uma função tal que $f(x) = (f_i(x))$, com $f_i : X \to Y_i$. Se $\bigtimes Y_i$ estiver munido da topologia produto, vale que f é contínua se, e somente se todas f_i são contínuas.* □

Tal resultado não vale se trocarmos a topologia produto pela topologia das caixas.

2.6 Conexidade

Proposição 2.17. *Um conjunto conexo por caminhos é conexo. Um conjunto aberto e conexo é conexo por caminhos.* □

Proposição 2.18. *Uma união arbitrária de conjuntos conexos com pelo menos um ponto em comum é conexa* □

Proposição 2.19. *Se $A \subset B \subset cl(A)$ e A é conexo, então B é conexo. Em particular, o fecho de um conjunto conexo é conexo.* □

2.7 Recobrimentos

Definição 2.10. Seja X um espaço topológico. Um recobrimento de X é um par (Y, p) tal que Y é um espaço topológico e $p : Y \to X$ é uma função contínua sobrejetiva tal que, para cada $x \in X$, existe uma vizinhança aberta U de x tal que $p^{-1}(U)$ é uma união disjunta de abertos em Y, cada um dos quais é homeomorfo a U por p. Um recobrimento é dito universal se é simplesmente conexo. ♣

Vale que o recobrimento universal é recobrimento de qualquer recobrimento conexo de X.

Capítulo 3

Grupos Topológicos

Um grupo topológico é só um conjunto que possuem tanto estrutura de grupo quanto de espaço topológico e, além disso, estas estruturas conversam bem entre si. (especificamente, as operações de grupo são contínuas).

Definição 3.1. Uma tripla $(G, +, \tau)$ é dita um grupo topológico se $(G, +)$ é um grupo, (G, τ) é um espaço topológico e se as funções $+ : G \times G \to G$ e $- : G \to G$ tais que $+(g, h) = g + h$ e $-(g) = -g$ são contínuas, onde entende-se implicitamente que $G \times G$ está munido com a topologia produto. ♣

Proposição 3.1. *G é um grupo topológico se, e somente se, a função $f : G \times G \to G$ definida por $f(g, h) = gh^{-1}$ é contínua.* ☐

A demonstração segue de fatos básicos sobre continuidade de funções e sobre a topologia produto.

Proposição 3.2. *Seja G um grupo topológico e seja V uma vizinhança de e. Então existe uma vizinhança W de e tal que $p^{-1}q \in V$, para todos $p, q \in W$.* ☐

Demonstração. Para ver isso, considere a função $f : G \times G \to G$ dada por $f(p, q) = p^{-1}q$. Pela proposição 3.1, f é contínua. Assim, $f^{-1}(V)$ é uma vizinhança de (e, e). Por isso, e pelo fato do produto cartesiano de abertos ser base da topologia produto, temos $(e, e) \in V_1 x V_2 \subset f^{-1}(V)$. Assim, basta tomar $W = V_1 \cap V_2$ ■

Dado $g \in G$, G grupo topológico, não é difícil ver que a função $m_g(x) : G \to G$ definida por $m_g(x) = gx$ é uma bijeção contínua, sendo $m_{g^{-1}}$ sua função inversa (também contínua). Assim, m_g é um homeomorfismo, para todo $g \in G$. Dados $x, y \in G$ arbitrários, então $m_{x^{-1}y}(x) = y$, o que indica que o grupo de homeomorfismos de um grupo topológico é transitivo Isto é, qualquer ponto de G é topologicamente semelhante a qualquer outro ponto. Espaços topológicos que gozem desta propriedade são comumente denominados espaços homogêneos. Mas esta denominação também é usada para um espaço métrico em que o grupo de isometrias aja transitivamente sobre ele. Como isometrias são homeomorfismos, é fácil ver que todo espaço métrico homogêneo é um espaço topológico homogêneo, se a topologia é induzida pela métrica.

Uma consequência de um espaço topológico ser homogênea é de que toda vizinhança é homeomorfa a alguma vizinhança do elemento neutro:

Proposição 3.3. *Seja G um espaço topológico, $g \in G$ e U uma vizinhança aberta de g. Então existe V vizinhança aberta de e (elemento neutro) tal que U e V são homeomorfos.* □

Demonstração. Considere a função m_g como definida acima. Seja $V := m_g^{-1}(U)$. É fácil ver que V de fato é uma vizinhança aberta de e. Pela bijetividade de m_g, vale que $m_g(V) = U$. Defina então $f : V \to m_g(V) = U$ por $f(x) = m_g(x)$. f será um homeomorfismo. É fácil ver que f será bijetiva e contínua, por ser restrição de função contínua, e também é fácil ver que f^{-1} é restrição de $m_g^{-1} = m_{g^{-1}}$, contínua, donde f^{-1} é contínua. ■

Esta semelhança entre as várias partes de um grupo topológico permite-nos encontrar um critério mais simples para decidir se um grupo é discreto, basta ver se apenas um elemento é isolado.

Proposição 3.4. *Sejam G um grupo topológico, H um subgrupo de G e $g \in H$. Então H é um subgrupo discreto se, e somente se, g for um ponto isolado de H.*

□

Demonstração. Se H é discreto, segue trivialmente que existe g isolado (tome por exemplo a identidade) Por outro lado, seja U um aberto tal que $U \cap H = \{g\}$. Tome $h \in H$ arbitrário e tome o homeomorfismo $m_{gh^{-1}}$. Note que $m_{gh^{-1}}(h) = g$, donde $V := m_{gh^{-1}}^{-1}(U)$ é uma vizinhança aberta de h. Mostraremos, agora, que $V \cap H = \{h\}$. Seja $x \in V \cap H$. Pela definição de V, temos que $m_{gh^{-1}}(x) = gh^{-1}x \in U$. Por construção, $gh^{-1}x \in H$, donde devemos ter $gh^{-1}x = m_{gh^{-1}}(x) = g$ necessariamente, por hipótese. Entretanto, como $m_{gh^{-1}}$ é bijetiva e $m_{gh^{-1}}(h) = g$, devemos ter necessariamente $h = x$, portanto h é um ponto isolado. Pela arbitrariedade de h, concluímos que H é um subgrupo discreto. ■

Corolário 3.4.1. *Seja G um grupo topológico. Então G é discreto se, e somente se, existir um aberto da topologia contendo apenas um elemento do grupo.* □

Demonstração. Se G é discreto, obviamente haverá abertos contendo apenas um elemento. Por outro lado, suponha que $\{g\}$ é um aberto, para algum $g \in G$. então g é um ponto isolado de G, donde G é discreto, pela proposição. ■

Lema 3.5. *Seja G um grupo topológico $T1$ e H um subgrupo discreto de G. Então H é fechado.* □

Demonstração. Iremos provar que H contém todos os seus pontos de aderência, por absurdo. De fato, suponha que exista um x no fecho de H e que não está em H. Então x é um ponto de acumulação de H. Como o espaço é $T1$, toda vizinhança de x possui infinitos elementos de H. Acharemos, entretanto uma vizinhança de x cuja intersecção com H é unitária.

Tome V uma vizinhança de e cuja intersecção com H é unitária, e que existe pelo fato de H ser discreto. Além do mais, tome também uma outra vizinha W de e de tal modo que $p^{-1}q \in V$,

$\forall p, q \in W$. Tal vizinhança existe pela proposição 3.2. Provaremos que xW é a vizinhança de x desejada. Primeiro, é claro que tal conjunto é uma vizinhança de x pelo fato de W ser uma vizinhança de e e como escólio da proposição 3.3. Mostrarei que $xW \cap H$ é unitário. Tome dois pontos de $xW \cap H$, digamos xw_1 e xw_2. Então $(xw_1)^{-1}(xw_2) = w_1^{-1}w_2$. Como H é fechado pela operação, devemos ter $w_1^{-1}w_2 \in H$. Por construção, temos também $w_1^{-1}w_2 \in V$. Pela escolha de V, devemos ter necessariamente $w_1^{-1}w_2 = e$, e portanto $w_1 = w_2$, donde $xw_1 = xw_2$, como queríamos demonstrar. ■

Capítulo 4

Alguns Grupos de Matrizes

4.1 O grupo SL(2, $\mathbb{R}$)

O leitor já deve saber que SL(2, $\mathbb{R}$) é o subgrupo de $GL(2, \mathbb{R})$ das matrizes com determinante 1. Entretanto, há um jeito elegante de definir não apenas este conjunto, mas também tratar do caso geral $SL(n, \mathbb{K})$, $\mathbb{K}$ um corpo arbitrário[1], e que ainda nos dá de graça o fato de $SL(n, \mathbb{K})$ ser um grupo. Para tanto, note que podemos pensar no determinante como um homomorfismo $\det : GL(n, \mathbb{K}) \to \mathbb{K}^*$, onde $\mathbb{K}^* := \mathbb{K} \backslash \{0\}$ é tratado como grupo multiplicativo, uma vez que $\det(MN) = \det(M) \det(N)$. Com efeito, o determinante será um epimorfismo. Definimos então $SL(n, \mathbb{K}) := \ker(\det)$ e o fato de ser um grupo vem diretamente da proposição 1.2. Uma aplicação do corolário 1.13.2 nos informa que há um isomorfismo entre $\mathbb{K}^*$ e $GL(n, \mathbb{K}) \backslash SL(n, \mathbb{K})$.

Geometricamente, podemos pensar em $SL(n, \mathbb{R})$ como o conjunto das transformações lineares de $\mathbb{R}^n$ que preservam área orientada.

Pois bem, nesta seção estudaremos algumas propriedades algébricas e topológicas do grupo SL(2, $\mathbb{R}$). Por exemplo, a decomposição de Iwasawa, além de nos fornecer um conjunto gerador para SL(2, $\mathbb{R}$), nos permitirá ver que este grupo é homeomorfo ao produto de um círculo pelo semiplano, que implicará que o mesmo grupo é homeomorfo ao produto de um círculo pelo plano, ou de um círculo por um disco, que corresponde a um toro sólido sem a casca.

O primeiro passo, então, é definirmos uma topologia para $SL(2, \mathbb{R})$. O jeito usual de definir topologia num grupo de matriz qualquer $M_n(X)$ sobre um espaço topológico X é de identificar um elemento de $M_n(X)$ com um elemento de X^{n^2}. Se X for métrico, podemos induzir uma métrica em $M_n(X)$ de maneira análoga. Assim, a topologia de $SL(2, \mathbb{R}$ é a mesma do conjunto $\{(a, b, c, d) \in \mathbb{R}, ad - bc = 1\}$.

[1]Na verdade, para a maioria dos resultados deste capítulo, é possível generalizar ainda mais para o caso de um anel comutativo.

4.2 Grupos Projetivos

Grupos projetivos são uma classe de grupos muito importantes para nossos propósitos. Por isso, e por serem pouco familiares para muitas pessoas, decidimos escrever uma seção especialmente para introduzir os conceitos por trás deles.

Grupos projetivos são, em geral, grupos quocientes de grupos de matrizes, sendo os mais importantes os grupos $PGL(n, \mathbb{K})$ e $PSL(n, \mathbb{K})$, onde $\mathbb{K}$ é um corpo.

A motivação por detrás dos grupos projetivos vem, como o próprio nome sugere, da geometria projetiva. Ingenuamente, podemos pensar que na geometria projetiva, uma transformação multiplicada por um escalar é idêntica à transformação original. Isto é, transformações que só diferem por multiplicação de escalar são iguais, o que nos leva a querer criar classes de equivalência e tratar uma transformação do espaço projetivo como todo um conjunto de transformações que agem da mesma forma.

O mesmo fenômeno ocorre com um grupo de transformações conhecido como transformações de Möbius. Estas transformações são muito importantes para a análise complexa (por ser o grupo de automorfismos da esfera de Riemann) e para a geométrica hiperbólica (pois alguns de seus subgrupos estão ligados a grupo de isometrias de modelos do espaço hiperbólico).

Convém conhecermos um pouco sobre as transformações de Möbius para motivarmos o estudo dos grupos projetivos, uma vez que acreditamos ser de mais fácil compreensão que geometria projetiva (e mais úteis para os nosso propósitos).

Veremos com mais detalhes na próxima seção que, em geral, uma transformação de Möbius é uma função da forma:

$$z \mapsto \frac{az + b}{cz + d}$$

Com $a, b, c, d \in \mathbb{C}$ e $ac - bd \neq 0$. Veremos também que estas funções formam um grupo, denotado por Möb($\mathbb{C}$). Este grupo é bastante parecido, no formato de suas operações, com alguns grupos familiares de matrizes. De fato, um cálculo direto nos mostra que:

Proposição 4.1. *A função $\theta : GL(2, \mathbb{C}) \to Möb(\mathbb{C})$ que mapeia:*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az + b}{cz + d}$$

É um epimorfismo

Em particular, esta proposição nos diz que as composição de transformações e suas as inversas possuem coeficientes que coincidem com os coeficientes da multiplicação e inversa das matrizes associadas. Entretanto, θ não se configura como um isomorfismo, pois $\theta(\lambda M) = \theta(M)$, $\forall \lambda \in \mathbb{C}\backslash\{0\}$ e $\forall M \in GL(2, \mathbb{C})$. De fato:

$$\lambda \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix} \mapsto \frac{(\lambda a)z + (\lambda b)}{(\lambda c)z + (\lambda d)} = \frac{\lambda(az+b)}{\lambda(cz+d)} = \frac{az+b}{cz+d}$$

Eis uma situação em que convém tratar um conjunto de matrizes que diferem apenas por multiplicação de escalar como uma coisa só, justamente pela estrutura de uma transformação de Möbius. Como já comentamos (e como era de se esperar), faremos isso quocientando o grupo de matriz em questão. Mas quocientar com qual subgrupo (normal)?

A escolha mais natural para os nosso propósitos seria usar o próprio núcleo de θ, que já sabemos ser um subgrupo normal (proposição 1.7), e é o que de fato faremos, a princípio:

Proposição 4.2. *O núcleo* $\ker\theta$ *é o conjunto* $\{\lambda\mathbb{I}, \lambda \in \mathbb{K}^*\}$. □

Demonstração. É óbvio ver que o núcleo contém esse conjunto, pois $\theta(\lambda\mathbb{I})(z) = (\lambda z)/\lambda = z$, que é a identidade. Reciprocamente, se $\theta(M)$ for a identidade, então $\theta(M) = z \implies az+b = z(cz+d) = cz^2 + dz \implies cz^2 + (d-a)z - b = 0$. Como esta função deve ser zero para todo z, então deve-se ter $b = c = 0$ e $a = d$, isto é, a matriz é $M = a\mathbb{I}$, como gostaríamos. Note que a não pode ser zero pois $a\mathbb{I} \in GL(2,\mathbb{C}) \iff a \neq 0$ ■

Esta escolha, entretanto, depende da existência de um tal θ, e, em contextos mais gerais, pode-se querer construir os grupos projetivos sem depender de homomorfismos específicos. Tratarei sobre o que é feito nesses casos: sabemos que o centralizador de um grupo é um subgrupo normal (proposição 1.24). Mas será que, neste caso, quocientar com o centralizador resultaria no resultado desejado? Em outras palavras, o centralizador de $GL(n,\mathbb{K})$ é o conjunto de "transformações escalares" $\lambda\mathbb{I}_n$, $\lambda \in \mathbb{K}^*$? A resposta é sim!

Note que $\mathbb{K}^* := \mathbb{K}\backslash\{0\}$ e multiplicar uma matriz por um escalar é o mesmo que multiplicar por essas "transformações escalares".

Proposição 4.3. *O centralizador* $Z(n,\mathbb{K})$ *do grupo* $GL(n,\mathbb{K})$ *é o conjunto* $\{\lambda\mathbb{I}_n, \lambda \in \mathbb{K}^*\}$. □

Assim, para n arbitrário e $\mathbb{K}$ um corpo qualquer, o resultado anterior nos motiva a seguinte:

Definição 4.1. O grupo linear projetivo $PGL(n,\mathbb{K})$ é definido como o grupo quociente $GL(n,\mathbb{K})\backslash Z(n,\mathbb{K})$. ♣

Para o caso específico de $n = 2$ e $\mathbb{K} = \mathbb{C}$, sabendo que $\ker\theta = Z(2,\mathbb{C})$, pelo Primeiro Teorema de Isomorfismos (1.13.1), ganhamos de graça que $PGL(2,\mathbb{C}) \cong \text{Möb}(\mathbb{C})$.

Enquanto que a definição de PGL é relativamente unânime entre os autores (alguns quocientam com o núcleo de um homomorfismo específico, ou não chamam a atenção para o fato do conjunto ser centralizador... mas o conjunto sempre é o mesmo), a definição de PSL possui duas definições (realmente) diferentes, embora os grupos resultantes sejam isomorfos.

Na definição que seguiremos, somos motivados a quocientar dessa vez o grupo $SL(n,\mathbb{K})$, de uma maneira parecida com a que fizemos para o grupo $GL(n,\mathbb{K})$, de modo que o grupo quociente resultante herde de SL a propriedade de que seus elementos *tenham* determinante 1. Mais especificamente, queremos que cada elemento de PSL seja uma classe de equivalência com *todos* os elementos tendo determinante 1 e que difiram apenas pela multiplicação por um escalar.

Na outra definição, simplificadamente, PSL seria formado por classes de equivalência de elementos que *podem* ter determinante 1 *ao serem multiplicados por um escalar adequado*. E, portanto, não seria um quociente de SL. Veremos mais para frente o que seria.

Mais uma vez: na nossa definição, PSL será um grupo cujos elementos são classes de equivalência contendo *apenas* matrizes de determinante 1. Na definição alternativa, PSL seria o grupo cujos elementos são classes de equivalência contendo matrizes que *podem* ter determinante 1 ao serem multiplicadas por um escalar, ou seja, classes de equivalência que contém *algum* elemento de determinante 1. A possibilidade de multiplicar a matriz por um escalar de modo que seu novo determinante seja 1 depende do corpo considerado. Veremos que, quando tratamos do corpo dos complexos, isso sempre é possível. Assim, *na definição alternativa*, valeria que $PGL(n,\mathbb{C}) = PSL(n,\mathbb{C})$. Esta igualdade não se verificará entretanto na nossa definição, sendo substituída por um isomorfismo.

Dada a motivação do que queremos que PSL seja, acabará que nossa definição será parecida com a do PGL. A única diferença é que devemos quocientar por um subgrupo de SL, não podendo ser, portanto, $Z(n,\mathbb{K})$ completamente. Quocientaremos, portanto, pelo grupo $SZ(n,\mathbb{K}) := \lambda\mathbb{I}_n \cap SL(n,\mathbb{K})$ que é um subgrupo normal por (proposição 1.8).

Definição 4.2. O grupo linear especial projetivo $PSL(n,\mathbb{K})$ é definido como o grupo quociente $SL(n,\mathbb{K})\backslash SZ(n,\mathbb{K})$. ♣

Vejamos como seria a cara do grupo $PSL(2,\mathbb{C})$, a título de exemplo. O grupo $SZ(2,\mathbb{C})$ é o grupo das matrizes $\lambda\mathbb{I}$ com determinante 1. Ora:

$$\det\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix} = \lambda^2$$

E então queremos que $\lambda^2 = 1$, que só é possível caso $\lambda = 1$ ou $\lambda = -1$. Explicitamente isso nos dá $SZ(2,\mathbb{C}) = \{\mathbb{I}, -\mathbb{I}\}$ e os elementos de $PSL(2,\mathbb{C})$ são parzinhos da forma $\{M, -M\}$, com $M \in SL(2,\mathbb{C})$. Analogamente, temos também que os elementos de $PSL(2,\mathbb{R})$ são parzinhos da forma $\{M, -M\}$, com $M \in SL(2,\mathbb{R})$.

Em especial, a identidade de $PSL(2,\mathbb{C})$ seria então o parzinho $\{\mathbb{I}, -\mathbb{I}\}$, que claramente não possui $2\mathbb{I}$ como elemento, apesar de que $2\mathbb{I}$ é elemento da identidade de $PGL(2,\mathbb{C})$, que é exatamente $Z(2,\mathbb{C})$. Isso mostra que $PGL(2,\mathbb{C}) \neq PSL(2,\mathbb{C})$, definitivamente.

Agora, para atingir o objetivo da definição alternativa, teríamos que definir $PSL(n,\mathbb{K})$ como a imagem de $SL(n,\mathbb{K})$ pela aplicação quociente $q : GL(n,\mathbb{K}) \to PGL(n,\mathbb{K})$. Esta aplicação q manda cada matriz de GL no respectivo conjuntinho de PGL que a contém.

Definição 4.3. O grupo linear especial alternativo $PSL*(n,\mathbb{K})$ é definido como $\{q(M), M \in SL(n,\mathbb{K})\}$, onde $q : GL(n,\mathbb{K}) \to PGL(n,\mathbb{K})$ é a aplicação quociente. ♣

A identidade de $PSL*(2,\mathbb{C})$ é então $q(\mathbb{I}) = Z(2,\mathbb{C})$, que, em particular, contém o elemento $2\mathbb{I}$. Não apenas isso mas, mais geralmente, temos $PGL(n,\mathbb{C}) = PSL*(n,\mathbb{C})$. A prova deste fato usa violentamente o fato que $\mathbb{C}$ é algebricamente fechado, o que motiva o seguinte resultado:

Proposição 4.4. *$PGL(n,\mathbb{K}) = PSL*(n,\mathbb{K})$ se, e somente se $\mathbb{K}$ possui a raíz $n-$ésima de todo elemento.* □

Lembrando que Möb($\mathbb{C}$) é isomorfo a $PGL(2,\mathbb{C})$, o que a proposição anterior está nos dizendo é que é possível escrever qualquer transformação de Möbius de modo que $ad-bc=1$. De fato, basta dividir em cima e embaixo por $\sqrt{ad-bc}$.

Resumo da ópera:

Proposição 4.5. *Möb*$(\mathbb{C}) \cong PGL(2,\mathbb{C}) = PSL*(2,\mathbb{C}) \cong PSL(2,\mathbb{C})$ □

O que acontece porém, com $PSL*(2,\mathbb{R})$, uma vez que $\mathbb{R}$ não possui raiz quadrada de números negativos? Qual é o conjunto $q(N)$ com $N \in SL(2,\mathbb{R})$? Em outras palavras, quais são as matrizes M tais que $\det(\lambda M) = 1$, para algum $\lambda \in \mathbb{R}$?

Afirmo que M não pode ter determinante negativo. De fato, $\det(\lambda M) = \lambda^2 \det(M)$. Como λ é real, λ^2 só pode ser positivo, e se fosse $\det(M) < 0$, teríamos $\det(\lambda M) < 0$, para todo λ, nunca podendo ser igual a 1. Reciprocamente, todas as matrizes de determinante positivo são equivalentes a alguma matriz com determinante 1, bastando multiplicar por $1/\sqrt{\det(M)}$, de forma que $PSL*(2,\mathbb{R})$ é o conjunto de classes de equivalência que contém apenas matrizes de determinante positivo. Podemos então definir Möb$_+(\mathbb{R})$ como o conjunto das transformações de Möbius a coeficientes reais tais que $ad-bc>0$ e valerá que Möb$_+(\mathbb{R}) \cong PSL*(2,\mathbb{R})$, donde qualquer transformação de Möbius assim pode ser escrita de modo que $ad-bc=1$.

Há uma vantagem na nossa definição com relação à alternativa: com ela, é mais natural estender o conceito de traço para grupos projetivos e transformações de Möbius. Seja $g \in PSL(2,\mathbb{K})$ com $\mathbb{K} = \mathbb{R}$ ou $\mathbb{R}$. Então $g = \{M, -M\}$, com $M \in SL(2,\mathbb{K}$. Assim, se definimos o traço de g como o traço das matrizes que o compõe, a função não estará bem definida, pois $tr(-M) = -tr(M)$. Entretanto, as funções $tr^2(g)$ e $Tr(g) := |tr(g)|$ estão, o que não seria o caso se usássemos a definição alternativa. O traço de uma transformação de Möbius T corresponde então ao traço do elemento de PSL(2, $\mathbb{K}$) correspondente.

Como os grupos projetivos que estamos lidando são quocientes de espaços topológicos, muni-los-emos com a topologia mais adequada: a topologia quociente. Esta topologia é, em geral, metrizável e a próxima proposição nos fornecerá uma métrica possível para o caso de $PSL(2,\mathbb{R})$, e que nos será útil adiante.

Proposição 4.6. *Sejam $[M] = \{M, -M\}$ e $[N] = \{N, -N\}$ elementos de $PSL(2,\mathbb{R})$. A métrica definida por $d([M],[N]) = \min\{d(M,N), d(M,-N)\}$ induz a topologia de $PSL(2,\mathbb{R})$.* □

Parte II

Grupos Fuchsianos

Capítulo 5

Transformações de Möbius

5.1 Noções de Análise Complexa

5.1.1 A Esfera de Riemann

Há algumas vantagens em usar o corpo dos números complexos. Uma delas é de que os complexos são algebricamente fechados. Outra, é de que uma função que é diferenciável uma vez em um domínio[1] é diferenciável infinitas vezes neste domínio e, para todo ponto dele, podemos escrever a função como uma série de potências convergente em uma determinada vizinhança do ponto (são as chamadas funções *holomorfas* ou *analíticas*). Duas desvantagens dos complexos são: não poder dividir por zero e $\mathbb{C}$ não ser compacto (e portanto poder haver sequências sem nenhuma subsequência convergente). Estas duas desvantagens podem ser ambas dribladas acrescentando ∞ aos complexos. Este novo conjunto será chamado de *plano complexo estendido*, e denotado por $\overline{\mathbb{C}} := \mathbb{C} \cup \infty$. Para os leitores preocupados, a natureza do elemento ∞ não importa - a princípio, pode tomar qualquer objeto que não pertença a $\mathbb{C}$.

Proposição 5.1. *Seja $S^2 \subset \mathbb{R}^3$ a esfera unitária de centro em $(0,0,0)$. A projeção estereográfica $\overline{\pi} : S^2/(0,0,1) \to \mathbb{C}$ é um homeomorfismo.* $\square$

Demonstração. A projeção estereográfica é dada por:

$$\overline{\pi}(x,y,z) = \frac{x}{1-z} + i\frac{y}{1-z} \tag{5.1}$$

Para ver que a projeção é contínua, basta ver que cada uma das funções coordenadas são contínuas (nunca se tem $z = 1$).

Falta ver que a função inversa é contínua. Para tanto, achemos agora uma expressão para esta. Definindo $u := \mathrm{Re}(\overline{\pi})$ e $v := \mathrm{Im}(\overline{\pi})$, temos:

[1]Neste capítulo, domínio será sinônimo de aberto conexo, quando subentendido pelo contexto.

$$u^2 + v^2 + 1 = \frac{x^2 + y^2 + (1-z)^2}{(1-z)^2} = \frac{x^2 + y^2 + z^2 + 1 - 2z}{(1-z)^2}$$

Como $(x, y, z) \in S^2$, $x^2 + y^2 + z^2 = 1$ e daí:

$$u^2 + v^2 + 1 = \frac{2 - 2z}{(1-z)^2} = \frac{2}{1-z}$$

Por (1), temos que:

$$u = \frac{x}{1-z} \implies x = u(1-z) = \frac{2u}{2/(1-z)} = \frac{2u}{u^2 + v^2 + 1}$$

Analogamente:

$$y = \frac{2v}{u^2 + v^2 + 1}$$

Para achar z em função de u e v, veja que:

$$u^2 + v^2 + 1 = \frac{2}{1-z} \implies 1 - z = \frac{2}{u^2 + v^2 + 1} \implies z = 1 - \frac{2}{u^2 + v^2 + 1}$$

$$\implies z = \frac{u^2 + v^2 + 1 - 2}{u^2 + v^2 + 1} \implies z = \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1}$$

Com isto, chegamos finalmente que:

$$\overline{\pi}^{-1}(u, v) = \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1} \right) \tag{5.2}$$

Que é evidentemente contínua pois cada função coordenada é contínua. ■

Podemos construir uma bijeção $\pi : S^2 \to \overline{\mathbb{C}}$ estendendo $\overline{\pi}$ de forma que $\pi(0, 0, 1) = \infty$. A semelhança entre $\overline{\mathbb{C}}$ e S^2 se estende também para a topologia, por isso os dois objetos costumam ser identificados e costumamos chamar $\overline{\mathbb{C}}$ de *esfera de Riemann*.

Definiremos a topologia de $\overline{\mathbb{C}}$ como as imagens por π de abertos de S^2, e teremos:

Proposição 5.2. *π é um homeomorfismo.* □

Demonstração. Seja A um aberto de $\overline{\mathbb{C}}$. Quero mostrar que $\pi^{-1}(A)$ é aberto de S^2. Ora, por definição de aberto em $\overline{\mathbb{C}}$, existe B aberto de S^2 tal que $A = \pi(B)$. Como π é uma bijeção, $\pi^{-1}(A) = \pi^{-1}(\pi(B)) = B$, que é aberto de S^2. Está provado que π é contínua.

Agora, dado B aberto de S^2, quero provar que $\tau^{-1}(B)$ é aberto de $\overline{\mathbb{C}}$, onde τ é a função inversa de π. Ora, a imagem inversa da função inversa de uma função, coincide com sua imagem direta, donde $\tau^{-1}(B) = \pi(B)$, que é aberto por definição. ■

Corolário 5.2.1. $\overline{\mathbb{C}}$ *é compacto* □

Demonstração. S^2 é compacto e $\overline{\mathbb{C}}$ é a imagem de S^2 sob π. Como π é homeomorfismo, em particular é contínua. Como a imagem de compactos por funções contínuas é compacto, a proposição segue. ■

Proposição 5.3. *Todo aberto de* $\overline{\mathbb{C}}$ *ou é um aberto de* $\mathbb{C}$*, ou é o complementar de um compacto de* $\mathbb{C}$ *unido com* ∞ □

O processo que acabamos de fazer é o caso particular de uma técnica mais geral conhecida como compactificação de espaços topológicos (mais especificamente, *compactificação de Alexandrov*).

Sobre a aritmética, definimos: $\infty + z = \infty$ para todo $z \in \mathbb{C}$, bem como para $z = \infty$. Se $z \neq 0$, então $z\infty = \infty$ (mesmo se z for infinito). As operações $\infty - \infty$ e 0∞ ficam indeterminadas e portanto a esfera de Riemann não se configura como um corpo.

5.1.2 Funções Racionais

Ao transformar o plano complexo na esfera de Riemann, o conceito de holomorfismo pode ser generalizado. Agora, funções com certas singularidades podem se tornar bem definidas, se definirmos o valor da função na singularidade como sendo infinito. Por questões de bom comportamento, estas singularidades devem ser pólos e devem ser isoladas. Uma função complexa holomorfa em todo plano exceto em um conjunto discreto de pontos que são pólos é chamada de função meromorfa. Assim, a generalização do conceito de "holomorfismo"para a esfera de Riemann passa a ser uma adaptação do conceito de meromorfismo no plano. Alguns autores chamam esta generalização de "holomorfismo"para a esfera de Riemann de holomorfismo de fato, enquanto outros ainda usam o termo meromorfismo. Usaremos o segundo.

Assim como holomorfismo no plano implicava em continuidade, a proposição seguinte nos garante que:

Proposição 5.4. *Se* f *é meromorfa em* $\overline{\mathbb{C}}$*, então* f *é contínua em* $\overline{\mathbb{C}}$ □

Diferentemente das funções holomorfas no plano, é possível descrever a forma de todas as funções meromorfas:

Teorema 5.5. *Uma função de* $\overline{\mathbb{C}}$ *para* $\overline{\mathbb{C}}$ *é meromorfa se, e somente se é uma função racional* □

E além disso:

Proposição 5.6. *Uma função racional $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ é injetiva se, e somente se for composta por polinômios de grau igual a 1.* $\square$

O conjunto das funções meromorfas em $\overline{\mathbb{C}}$ forma um corpo $\mathbb{C}(z)$. Este corpo possui um subcorpo isomorfo a $\mathbb{C}$ (o das funções constantes) e portanto, $\mathbb{C}(z)$ pode ser considerado uma extensão de $\mathbb{C}$

5.2 Círculos e Retas

Nesta seção, iremos generalizar os conceitos de círculos e retas em dois sentidos. O primeiro se refere ao tratamento para dimensões arbitrárias. O segundo tem a motivação de tratar círculos e retas de maneira unificada, de maneira que seremos capazes de pensar numa reta como um círculo de raio infinito.

Note que uma reta, no plano, é um subespaço afim de dimensão 1. No espaço tridimensional, o análogo a uma reta seria um plano, que é um subespaço afim de dimensão 2. Feitas estas observações, parece razoável dizer que a generalização da reta num espaço $n-$dimensional é um subespaço afim de dimensão $n-1$. A diferença entre a dimensão de um subespaço e o espaço que o contém é chamada de codimensão. Assim, um hiperplano é um subespaço de codimensão 1.

Antes de tratarmos do caso afim, tratemos do caso linear. É um resultado bem conhecido de álgebra linear que o complemento ortogonal de um subespaço de dimensão $n-1$ num espaço $n-$dimensional será um subespaço de uma dimensão, isto é, gerado por um vetor. Assim, qualquer subespaço de dimensão $n-1$ pode ser visto como complemento ortogonal de um vetor não-nulo. Assim, um subespaço de codimensão 1 equivale ao conjunto $\{v \in V; \langle p, v\rangle = 0\}$, para algum p não-nulo. Levando em consideração o produto interno usual, um subespaço é o conjunto de vetores $(x_1, ..., x_n)$ tais que:

$$a_1x_1 + ... + a_nx_n = 0$$

Onde $(a_1, ..., a_n) = p$. Para duas dimensões, recuperamos a equação da reta que passa na origem: $ax + by = 0 \iff y = -(a/b)x$.

Temos ainda que um subespaço afim é um conjunto de pontos que seria um subespaço se todos os pontos fossem subtraídos por um mesmo vetor, a fim de passarem pela origem. Isso nos diz que um hiperplano é o conjunto de vetores: $\{v \in V; \langle (v-w), p\rangle = 0\} = \{v \in V; \langle v, p\rangle - \langle w, p\rangle = 0\} = \{v \in V; \langle v, p\rangle = b\}$, com $b = \langle w, p\rangle$. Isso nos leva a seguinte definição:

Definição 5.1. Um hiperplano em $\mathbb{R}^n$ é o conjunto de pontos x que satisfaz $p \cdot x = b$, com $p \in \mathbb{R}^n$ e b escalar. ♣

Em outros termos, um hiperplano é o conjunto de pontos que satisfaz:

$$a_1x_1 + ... + a_nx_n = b$$

Que recupera a equação geral da reta: $ax + by = c \iff y = -(a/b)x + c/b$ e a equação do plano: $ax + by + cz = d$.

Vamos agora unificar os conceitos de reta e círculo em $\mathbb{C}$. Ora, já vimos que uma reta no plano é o conjunto de pontos que satisfaz $ax + by + c = 0$ para a, b, c reais. Escrevendo x, y em termos de $z = x + iy$ e seu complexo conjugado, temos que uma reta no plano complexo é o conjunto de pontos que satisfazem:

$$a\left(\frac{1}{2}(z + z^*)\right) + b\left(\frac{1}{2i}(z - z^*)\right) + c = 0$$

Rearranjando, temos:

$$\beta z + \beta^* z^* + c = 0$$

Com $\beta = (a - ib)/2$. (Verifique!)

Por outro lado, um círculo no plano complexo tem a forma $|z - z_0| = r^2$. Podemos reescrever esta relação no formato: $(z - z_0)(z - z_0)^* = r^2 = (z - z_0)(z^* - z_0^*) = r^2$.

$$\implies zz^* - z_0^* z - z_0 z^* + z_0 z_0^* - r^2 = 0$$

Tomando $\beta = -z_0^*$ e $c = z_0 z_0^* - r^2$, a equação acima se escreve:

$$zz^* + \beta z + \beta^* z^* + c = 0$$

Sabendo que multiplicar a equação da reta e do círculo por um escalar não-nulo não altera o conjunto de pontos que são solução, temos que ambas se apresentam na forma $\alpha zz^* + \beta z + \beta^* z^* + \gamma = 0$, onde uma reta ocorre quando $\alpha = 0$ e um círculo ocorre caso contrário. Isso nos motiva a seguinte definição:

Definição 5.2. Um círculo no plano complexo é o conjunto de pontos que satisfaz a seguinte equação:

$$\alpha zz^* + \beta z + \beta^* z^* + \gamma = 0$$

Com α, γ reais.

Esta definição entra em conflito com a definição usual de círculo, mas a distinção será clara pelo contexto.

5.3 Conformalidade

Intuitivamente, uma transformação conforme é uma transformação que preserva ângulos. Um jeito de tornar esta ideia um pouco mais rigorosa é dizer que, para qualquer ponto e quaisquer duas curvas que cruzam este ponto, o ângulo entre os vetores tangentes a estas curvas é igual ao ângulo dos vetores tangentes às curvas que são imagem das curvas originais.

Um critério equivalente mas um pouco mais simples é de que uma transformação é conforme quando sua jacobiana, a cada ponto, é uma constante vezes uma matriz ortogonal. A razão para isto ficará mais clara na subseção 6.2.3.

Um conceito relacionado é o de transformação conforme que preserva orientação. Uma tal função preservaria ângulos orientados e a sua jacobiana, a cada ponto, deve ser um múltiplo positivo de uma matriz ortogonal de determinante 1.

Para o caso específico de uma função holomorfa, as condições de Cauchy-Riemann garantem que qualquer função holomorfa seja conforme, desde que sua derivada não seja nula.

5.4 Definição e propriedades básicas

Definição 5.3. Uma função $T : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ é dita ser uma *Transformação de Möbius* se existem $a, b, c, d \in \mathbb{C}$ com $ad - bc \neq 0$ tais que:

$$T(z) = \frac{az + b}{cz + d}$$

$\forall z \in \overline{\mathbb{C}}$. ♣

Note que é necessário que a função tome valores nos complexos estendidos, para estar bem definida (pelo menos se $c \neq 0$), caso contrário teríamos um pólo em $-d/c$.

É fácil ver que as transformações de Möbius formam um grupo sob a operação de composição. Entretanto, para que isto seja verdade, torna-se relevante o fato de $ad - bc \neq 0$, como segue da proposição 5.7. Este grupo será denotado por Möb($\mathbb{C}$). Já vimos brevemente este grupo na seção 4.2. Recomendo fortemente a leitura desta seção antes da presente, mas relembrarei que os principais fatos são: 1) Möb($\mathbb{C}$) $\cong PSL(2, \mathbb{C})$. Em particular, os coeficientes da composição são dados pelos coeficientes da multiplicação das matrizes associadas[2]; e 2) qualquer transformação de Möbius pode ser escrita de modo que $ad - bc = 1$. Assim, a partir de agora, ao considerarmos uma transformação de Möbius, podemos assumir implicitamente que seus coeficientes satisfazem $ad - bc = 1$.

Proposição 5.7. *Seja* $T : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ *dada por:*

[2]Esse fato poderia fazer o leitor atencioso suspeitar que transformações com $ad - bc = 0$ correspondem a matrizes com determinante nulo (não-invertíveis) e portanto as próprias transformações não teriam inversa

$$T(z) = \frac{az+b}{cz+d}$$

Com $a, b, c, d \in \mathbb{C}$ quaisquer. Então, T é constante se, e somente se (estiver bem-definida e) $ad - bc = 0$

Demonstração. O requerimento de que a função esteja bem definida é para evitar patologias como todos os coeficientes serem nulos, o que levaria a indeterminações. Excluindo estes casos, sabemos por um resultado elementar de álgebra linear que se o determinante de uma matriz $(ab|cd)$ for nulo, então (a, b) e (c, d) são proporcionais, isto é, existe λ tal que $(az+b)/(cz+d) = \lambda(cz+d)/(cz+d)$, donde a transformação é a constante λ. Reciprocamente, teríamos $az + b = \lambda cz + \lambda d \implies (a - \lambda c)z + (b - \lambda d)z = 0$ para todo z, e portanto $ad - bc = \lambda ab - \lambda ba = 0$. ∎

Proposição 5.8. *As transformações de Möbius são bijeções.*

Demonstração. Se $c = 0$, a transformação é afim, e portanto bijetiva. Para $c \neq 0$, provaremos primeiro a injetividade. Tome $z, w \in \mathbb{C}\backslash\{-d/c\}$. Se $T(z) = T(w)$, então $(az + b)(cw + d) = (aw + b)(cz + d)$. Abrindo as contas chegamos em $(ad - bc)z = (ad - bc)w$. Como $ad - bc \neq 0$, $z = w$. Para a sobrejetividade, $T(z) = y$ significa $az + b = czy + dy \implies (a - cy)z = dy - b$ e portanto $z = (dy - b)/(a - cy)$. Se $y = a/c$, então $z = \infty$. ∎

Na presente seção, iremos nos dedicar a estudar as propriedades de Möb($\mathbb{C}$) relacionadas com sua ação em $\overline{\mathbb{C}}$. Na verdade, podemos pensar nas transformações de Möbius como a expressão da ação do grupo $PSL(2, \mathbb{C})$ em $\overline{C}$. Neste sentido, esta seção irá se preocupar em estudar o grupo $PSL(2, \mathbb{C})$ enquanto agente deste espaço. Assim, estudaremos questões de pontos fixos, estabilizadores, transitividade, etc.

Convém estabelecer uma notação que será seguida no decorrer de todo o texto, daqui em diante. Dada um matriz M, seja de um grupo linear quanto de um grupo projetivo, denotaremos por f_M a respectiva transformação de Möbius.

Voltando um pouco em seu papel na análise complexa, as transformações de Möbius são precisamente os automorfismos da esfera de Riemann. Isto significa que estas transformações preservam a estrutura de variedade complexa da esfera. Sem entrar em muitos detalhes agora, mas preservar a estrutura de variedade complexa significa que a transformação é uma bijeção meromorfa com inversa também meromorfa. A proposição seguinte demonstra isto, juntamente com o fato já conhecido que Möb($\mathbb{C}$) é um grupo.

Proposição 5.9. *Uma função $T : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ pertence a Möb($\mathbb{C}$) se, e somente se T é uma bijeção meromorfa* □

Demonstração. Basta usar em conjunto o teorema 5.5 e a proposição 5.6 ∎

Corolário 5.9.1. *Se $T \in$ Möb($\mathbb{C}$), então T é homeomorfismo de $\overline{\mathbb{C}}$.* □

Demonstração. Pela proposição, T é meromorfa. Pela proposição 5.4, T é contínua em $\overline{\mathbb{C}}$. Como Möb($\mathbb{C}$) é um grupo, $T^{-1} \in$ Möb($\mathbb{C}$), donde T^{-1} é meromorfa e portanto contínua. A afirmação segue. ■

Chamo a atenção para o fato de que a proposição acima diz que preservar a estrutura de variedade complexa nos dá automaticamente que a topologia também é preservada.

Busquemos agora um gerador para Möb. A proposição seguinte nos fornece isto.

Proposição 5.10. *Sejam R_θ definida por $R_\theta(z) = e^{i\theta}z$, J definida por $J(z) = 1/z$, S_r definida por $S_r(z) = rz$ e T_t definida por $T_t(z) = z + t$ funções da esfera de Riemann na esfera de Riemann. O conjunto $X = \{R_\theta, S_r, T_t, J | \theta \in \mathbb{R}, r \in \mathbb{R}^+, t \in \mathbb{C}\}$ gera Möb.* □

Devido a esta proposição, há uma interpretação geométrica bacana para transformações de Möbius. Imagine que temos uma esfera em cima de um plano e uma lâmpada em cima da esfera, que gera uma projeção estereográfica sobre o plano. Então as transformações de Möbius correspondem a uma composição das seguintes operações com a esfera: i) rotação sobre o eixo vertical (R_θ), ii) rotação sobre o eixo horizontal que passa por 1 e -1 (J), iii) contração ou expansão da esfera (S_r), iv) translações da esfera (T_t).

Uma outra característica importante das transformações de Möbius é o que se segue:

Teorema 5.11. *Uma transformação da esfera de Riemann em si mesma é um homeomorfismo conforme que preserva orientação se, e somente se, é uma transformação de Möbius.* □

Transformações de Möbius diferentes da identidade possuem no máximo dois pontos fixos. Isto é, se uma transformação de Möbius deixa três pontos fixos, então é a identidade. Podemos usar a interpretação acima e então este fato pode ser expresso nos seguinte termos: é impossível fazer quaisquer movimentos numa esfera fixando três dedos nela (em pontos diferentes).

5.5 Pontos fixos, transitividade e classes de conjugação

Proposição 5.12. *Uma transformação de Möbius T fixa ∞ se, e somente se $c = 0$.* □

Demonstração. Evidente, basta ver que $T(\infty) = a/c$. ■

Teorema 5.13. *Sejam $M \in PSL(2, \mathbb{C})$ e $T = f_M$ (tal que $ad - bc = 1$). Se $tr^2 T \neq 4$, então T fixa dois pontos. Se $tr^2 T = 4$ e não for a identidade, então fixa um ponto.* □

Demonstração. Separaremos em casos. Primeiro, suponha que $c \neq 0$. então, se z é um ponto fixo:

$$\frac{az + b}{cz + d} = z \implies cz^2 + (d - a)z - b = 0$$

A equação acima possui duas soluções se $\Delta \neq 0$ e uma se $\Delta = 0$. Ora:

$$\Delta = (d-a)^2 + 4bc = (a+d)^2 - 4 = tr^2T - 4$$

Assim, T fixa dois pontos se $tr^2T \neq 4$ e um se $tr^2 = 4$.

Vejamos agora o caso em que $c = 0$. Neste caso, $ad = 1$ e a transformação pode ser escrita como $T(z) = a^2z + ab$, e, como já foi dito, ∞ é um ponto fixo. Buscando por outros pontos fixos, temos:

$$a^2z + ab = z$$

Agora separaremos no caso que $a^2 = 1$ e $a^2 \neq 1$. Ora, no primeiro deles temos que $T(z) = z+b$, donde ∞ é o único ponto fixo se $b \neq 0$, isto é, se T não for a identidade. Caso contrário, teremos outro ponto fixo dado por:

$$z = \frac{ab}{1-a^2}$$

Até agora temos o seguinte: Se $c \neq 0$, T tem dois pontos fixos se $tr^2T \neq 4$ e um se $tr^2T = 4$. Se $c = 0$ e T não é a identidade, então T tem dois pontos fixos se $a^2 \neq 1$ e apenas um se $a^2 = 1$.

Para concluir a demonstração, basta agora mostrarmos que $a^2 = 1 \iff tr^2 = (a+d)^2 = 4$

Suponha que $a^2 = 1$, então $d^2 = 1$, pois $ad = 1 \implies a^2d^2 = 1 \implies d^2 = 1$. Daí, $tr^2T = (a+d)^2 = a^2 + 2ad + d^2 = 4$.

Por outro lado, temos $a^2 + 2add^2 = 4 \implies a^2 + 1/a^2 = 2$. Chamando a^2 de x (que deve ser diferente de zero), temos $x^2 - 2x + 1 = 0$, cuja única solução é $x = a^2 = 1$. ∎

Corolário 5.13.1. *Se T é uma transformação de Möbius que fixa três (ou mais) pontos, então T é a identidade.* □

Corolário 5.13.2. *O grupo de Möbius não é 4-transitivo.* □

Demonstração. Não existe uma transformação que leva $(0, 1, \infty, 2)$ em $(0, 1, \infty, 3)$, por exemplo.q ∎

Aproveitando o corolário acima, daremos uma pausa para estudarmos a transitividade do grupo.

Proposição 5.14. *O grupo de Möbius é simplesmente 3-transitivo.* □

Demonstração. Seja (z_1, z_2, z_3) uma tripla qualquer de números complexos distintos. Iremos achar uma transformação de Möbius que leva essa tripla em $(0, 1, \infty)$, e isto bastará.

Caso os três forem finitos, é fácil verificar que a seguinte transformação leva cada ponto em $(0, 1, \infty)$ respectivamente:

$$\frac{(z - z_1)(z_2 - z_3)}{(z_1 - z_2)(z_3 - z)}$$

Para o caso em que algum for infinito, basta tomar o limite na expressão acima, o que resulta em, respectivamente:

$$\frac{z_2 - z_3}{z - z_3}, \frac{z - z_1}{z - z_3}, \frac{z - z_1}{z_2 - z_1}$$

A demonstração da existência fica concluída pelo fato que $ad - bc \neq 0$ em todas as transformações; o que é matéria de fácil verificação, dado que os três pontos são distintos.

A unicidade é feita da seguinte forma. Seja U uma transformação que leva (z_1, z_2, z_3) em $(0, 1, \infty)$ e provaremos que $U = T$. De fato, U^{-1} leva $(0, 1, \infty)$ em (z_1, z_2, z_3), donde TU^{-1} fixa os pontos $(0, 1, \infty)$. Pelo que foi visto, TU^{-1} deve ser a identidade, e segue que $T = U$. ∎

O estudo dos pontos fixos irá nos conduzir naturalmente ao estudo das classes de conjugação. Antes, vale fazer uma observação. Mediante cálculo direto, podemos ver que $tr(AB) = tr(BA)$, para quaisquer matrizes A, B. Assim, valerá que $tr^2(UT) = tr^2(TU)$. O resultado prático disto é que o traço é invariante por classes de conjugação, isto é, se TeU são transformações conjugadas, então $tr^2(T) = tr^2(U)$, basta ver que $T = G^{-1}UG$ para algum G, donde $tr^2(T) = tr^2(G^{-1}UG) = tr^2(G^{-1}GU)$, pela observação acima. Entretanto, poderia ser o caso de haver duas classes de conjugação com o mesmo traço. Isto não ocorre, e as classes de conjugação possuem uma relação bijetiva com o traço (à exceção da identidade que, como já vimos, forma uma classe de conjugação por si própria). Antes de provarmos isso, provaremos o seguinte lema de natureza mais técnica.

Lema 5.15. *Seja $U_\lambda = \lambda z$ se $\lambda \neq 1$ e $U_1 = z + 1$. Então, toda transformação de Möbius, que não é a identidade, é conjugada a U_λ para algum $\lambda \neq 0$. Além disso, U_λ é conjugado a U_κ se, e somente se $\lambda = 1/\kappa$.* □

Demonstração. Já vimos que, se uma transformação não é a identidade, então ela possui um ou dois pontos fixos.

Consideremos primeiro uma transformação T que fixa um ponto, digamos z_0. Pela transitividade do grupo de Möbius, existe uma transformação S que leva ∞ em z_0, de modo que a transformação $S^{-1}TS$ fixa ∞ e somente ∞. Por um escólio do teorema 5.13, temos que $S^{-1}TS(z) = z+b$. Tomando a homotetia $V(z) = bz$ vemos facilmente que $V^{-1}(S^{-1}TS)V(z) = z + 1$, donde T é conjugado a U_1 mediante SV.

Agora, seja T uma transformação que fixa dois pontos, digamos z_1, z_2. Novamente, pela transitividade, existe uma transformação W que manda $(0, \infty)$ em (z_1, z_2) de forma que $W^{-1}TW$ fixa 0 e ∞. Novamente pelo escólio do teorema 5.13, temos que a transformação é dada por $a^2z + ab$. Entretanto, por fixar 0, deve valer que $b = 0$, e aí a transformação é do tipo U_λ com $\lambda = a^2 \neq 1$

Em primeiro lugar, U_1 não pode estar conjugado a nenhum outro U_λ, pois o número de pontos fixos não é o mesmo (como já comentamos, elementos conjugados possuem o mesmo número de pontos fixos). Agora, suponha que U_λ é conjugado a U_κ, então $tr^2(U_\lambda) = tr^2(U_\kappa)$, isto é:

$$\lambda + \frac{1}{\lambda} + 2 = \kappa + \frac{1}{\kappa} + 2 \implies \lambda + \frac{1}{\lambda} = \kappa + \frac{1}{\kappa} + \implies (\kappa + \frac{1}{\kappa})\lambda = \lambda^2 + 1$$

$$\implies \lambda^2 - (\kappa + \frac{1}{\kappa})\lambda + 1$$

Por Báskara, isso nos dá $\lambda = \kappa$ ou $\lambda = 1/\kappa$.

Reciprocamente, se $J(z) = 1/z, J^{-1}(z) = 1/z$, daí $J^{-1}U_\lambda J = U_{1/\lambda}$ ■

Assim, o par $\{\lambda, 1/\lambda\}$ é chamado multiplicador da transformação de Möbius e duas transformações são conjugadas se, e somente se possuem o mesmo multiplicador. Além disso:

Teorema 5.16. *Duas transformações de Möbius são conjugadas se, e somente se o traço ao quadrado delas são iguais.* □

Demonstração. Basta mostrar que se possuem o mesmo traço são conjugados. Digamos que estas duas transformações sejam conjugadas a U_λ e U_κ, e isso está garantido pelo lema 5.15 acima. Então, vale também que $tr^2(U_\lambda) = tr^2(U_\kappa)$ e já vimos que isso implica que $\kappa = \lambda$ ou $\kappa = 1/\lambda$. De qualquer forma, U_λ e U_κ são conjugados, e portanto as transformações iniciais também. ■

O lema 5.15 serviu, basicamente, para reduzir o problema de provar que traços iguais implicam em classes de conjugações iguais para apenas uma classe específica de transformações de Möbius.

Vemos que, associada com uma certa transformação de Möbius há dois valores, o traço ao quadrado e o multiplicador. Estes valores, além de especificarem a classe de conjugação, também determinam o comportamento geométrico das transformações e o comportamento de $T^n(z)$ quando n tende a infinito, como veremos na próxima subseção, e estão relacionadas pela fórmula: $tr^2(T) = \lambda + 1/\lambda + 2$

5.6 Classificação geométrica e propriedades dinâmicas

Definição 5.4. Seja T uma transformação de Möbius. Dizemos que T é elíptica se $0 \leqslant tr^2(T) < 4$, que é parabólica se $tr^2(T) = 4$, que é hiperbólica se $tr^2(T) > 4$ e, por fim, que é loxodrômico caso contrário, isto é, se $tr^2(T) < 0$ ou não é real. ♣

Imediatamente vemos que uma transformação tem apenas um ponto fixo se, e somente se, for parabólica e não for a identidade. Como vimos, o protótipo de uma tal transformação é a transformação $z \mapsto z+1$.

Devo advertir que alguns autores incluem as transformações hiperbólicas nas loxodrômicas.

A proposição seguinte fornece uma tradução da definição acima em termos de multiplicadores, no lugar do traço.

Proposição 5.17. *T é hiperbólica se, e somente se, λ é real, positivo e $|\lambda| \neq 1$. T é loxodrômica se, e somente se, $|\lambda| \neq 1$ e real negativo ou complexo. T é elíptica se, e somente se $|\lambda| = 1$.* □

Demonstração. A prova é feita por cálculo direto. ■

Teorema 5.18.

1. *Se T é uma transformação parabólica de ponto fixo z_0, então $\lim T^n(z) = z_0$, entre outra palavras, o conjunto limite de uma transformação parabólica é seu ponto fixo.*

2. *Se T é hiperbólico ou loxodrômico, de pontos fixos z_1, z_2, $\lim T^n(z)$ é um dos dois pontos fixos, para z que não seja ponto fixo. Em outras palavras, o conjunto limite de uma transformação hiperbólica ou loxodrômica são seus dois pontos fixos. O que ocorre é que os pontos se afastam de um ponto fixo e vão em direção ao outro.*

3. *Se T é elíptico, $T^n(z)$ não tem limite.*

□

Demonstração.

1. Já vimos que toda transformação parabólica é conjugada à transformação $U_1(z) = z+1$. Ora, note que $\lim U_1^n(z) = \lim z+n = \infty$. Assim, $\lim T^n(z) = \lim V^{-1}U_1^nV(z) = V^{-1}(\infty) = z_0$.

2. Se T é hiperbólica ou loxodrômica, vimos que é conjugada à uma transformação U_λ com $|\lambda| \neq 1$. Se $|\lambda| < 1$, para $z \neq \infty$ temos $\lim U_\lambda^n(z) = \lim \lambda^n z = 0$, que é um ponto fixo. Caso $|\lambda| > 1$ e $z \neq 0$, $\lim U_\lambda^n(z) = \infty$, que também é ponto fixo. Em qualquer caso, os pontos tendem a um dos dois pontos fixos.

3. Neste caso, a transformação é conjugada a U_λ, com $|\lambda| = 1$, donde $\lambda = e^{i\theta}$, para θ real, donde $U_\lambda^n = e^{in\theta}z$, que não tem limite para $z \neq 0, \infty$

■

Corolário 5.18.1. *Se T é periódica (isto é, se existe um inteiro tal que T^n é a identidade), então T e elíptica.* □

Demonstração. T é conjugada a U_λ para algum λ. Não pode ser $\lambda = 1$ pois nesse caso $U_1^n = z+n$, que nunca é a identidade. Assim, necessariamente $U_\lambda^n = \lambda^n z = z$, donde deve ser $|\lambda| = 1$, e portanto elíptica. ■

Entretanto, deve-se notar que não são todas as transformações elípticas que são periódicas.

5.7 Generalização para Várias Dimensões

Vale ressaltar que há uma maneira de generalizar as transformações de Möbius. Primeiro, use a identificação de $\mathbb{C}$ com $\mathbb{R}^2$. Use também o fato que todas as transformações de Möbius podem ser vistas como inversões sobre círculos ou reflexões sobre planos. Generalize agora estes conceitos para esferas e hiperplanos. Com isso, as transformações de Möbius em $\mathbb{R}^n$ podem ser definidas como o conjunto de transformações geradas pelas inversões sobre esferas e reflexões sobre hiperplanos.

Capítulo 6

Introdução à Geometria Hiperbólica

6.1 O Espaço Hiperbólico

Lembremos que os cinco axiomas de Euclides permaneceram por muito tempo sendo os princípios norteadores da geometria. O quinto axioma, que é equivalente a dizer: "dada uma reta e um ponto fora da reta, há apenas uma outra reta paralela à primeira que passa pelo ponto"ou (o que é equivalente), "a soma dos ângulos internos de um triângulo é 180º", era o axioma menos natural de ser aceito sem demonstração, e muitas pessoas tentaram provar que este axioma podia ser provado a partir dos outros quatro. Na primeira metade do século XIX, Gauss provou que isto não era possível, construindo uma geometria em que não valia o quinto axioma mas mesmo assim era consistente. Especificamente, triângulos podiam ter a soma interna de seus ângulos menor do que 180º. Apesar de Gauss não ter publicado este trabalho, anos mais tarde, os matemáticos Lobachevsky e Bolyai publicaram versões semelhantes, dando origem ao embrião do que seria a geometria hiperbólica.

O ambiente natural para tratar de geometria hiperbólica é o espaço hiperbólico. O espaço hiperbólico aparece naturalmente ao se fazer a classificação das variedades riemannianas completas, simplesmente conexas e de curvatura seccional constante. Para cada dimensão n e cada valor de curvatura, há apenas um tipo de variedade assim (a menos de isometrias): $\mathbb{R}^n$, se a curvatura for nula, a esfera ($\mathbb{S}^n$), se a curvatura for positiva, e o espaço hiperbólico ($\mathbb{H}^n$), se a curvatura for negativa.

A exemplo deste teorema, uma compreensão mais aprofundada dos temas abordados aqui necessitam de uma ampla dose de geometria diferencial, mais especificamente, de variedades riemannianas. Não exigiremos, entretanto, tais requisitos e pessoas sem conhecimento em geometria diferencial poderão acompanhar a imensa maioria do texto sem problemas. A despeito disto, durante o texto poderão haver alguns comentários, tratando das nossas construções enquanto variedades riemannianas propriamente ditas, a fim de esclarecer construções obscuras, sanar a curiosidade dos leitores interessados ou que possuem os requisitos para tal. A próxima seção tem como objetivo introduzir esta abordagem, sem compromissos.

Um modo natural de se construir o espaço hiperbólico é como uma esfera num espaço de

Minkowski, mas há outras construções (isométricas, pelo teorema citado acima). Estas construções são chamadas nesta seção de “modelos do espaço hiperbólico”. Veremos principalmente três deles: o modelo do semiplano, o modelo do disco de Poincaré e o modelo do hiperboloide (que é a construção mencionada acima como a mais natural), nesta ordem. A princípio, veremos todos os modelos em apenas duas dimensões, mas é possível generalizá-los para uma dimensão n arbitrária.

Como iremos ver, a geometria hiperbólica é completamente diferente da euclidiana e causa disto é a diferença no modo de medir distância em cada geometria. Mais que isso, veremos que a forma de construir distância em cada modelo é peculiar e pode parecer arbitrária, à primeira vista. Por isso, optamos por introduzir brevemente o modo usual de definir distância numa variedade riemanniana, para motivar a definição de distância no nossos modelos. A situação é análoga para áreas.

Na verdade, a construção do grupo de isometrias, geodésicas e outros objetos será muito mais trabalhoso no primeiro modelo que iremos tratar. Como os demais modelos são isométricos ao primeiro, basta encontrar a isometria e ela poderá ser usada para transportar todos os resultados do primeiro modelo para os outros de forma relativamente fácil.

Grande parte dos resultados que serão expostos na seção 3 e seguintes estão detalhadamente provados em [Bon09], por isso não demonstraremos a maior parte deles em detalhes, dando apenas uma ideia ou roteiro da demonstração.

6.2 Um mínimo de Variedades Riemannianas

6.2.1 Distância

O modo de definir distância numa variedade riemanniana reflete muito sobre a estrutura deste objeto, por isto gastaremos um tempo falando disto. Para os leitores menos habituados, uma variedade é uma generalização do conceito de superfície (se for diferenciável, é uma superfície onde faz sentido falar de coisas como derivadas e diferenciabilidade). Uma variedade riemanniana é uma variedade que possui uma estrutura a mais, chamada de métrica riemanniana. Esta estrutura é semelhante a um produto interno, dando portanto noção de comprimentos e ângulos e não deve ser confundida com métrica no sentido de espaço métrico[1]. Para ser mais exato, a métrica riemanniana é um objeto g que a cada ponto P da variedade associa um produto interno $g_P(\cdot,\cdot)$ no espaço tangente a este ponto (generalização de plano tangente para uma superfície) e que também deve satisfazer a um certo critério de suavidade. Este objeto também costuma ser chamada de tensor métrico. O tensor métrico não nos permite calcular distâncias diretamente, mas ele induz uma norma em cada espaço tangente, que pode ser usada para calcular comprimento de vetores tangentes a uma curva, e portanto comprimentos de curvas na nossa variedade. Explicitamente, seja (M, g) uma variedade riemanniana e $\gamma : I \to M$ uma curva em M. Então, o comprimento de γ pode ser obtido assim:

[1] Para não confundir, passaremos a nos referir a métrica no sentido de espaço métrico como distância sempre que possível.

$$L(\gamma) = \int_I ||\gamma'(t)||dt = \int_I \sqrt{g_{\gamma(t)}(\gamma'(t), \gamma'(t))}dt$$

É possível mostrar que esta integral é invariante por diferentes parametrizações.

Mas isto ainda não é a distância entre pontos. Como podemos usar comprimento de curvas para obter informações sobre a distância entre pontos? Tomando emprestada a intuição já construída sobre o conceito de distância, gostaríamos que a distância entre dois pontos tome o menor valor, considerando todas as formas de partir de um e chegar no outro. Assim, podemos definir a distância entre dois pontos como o ínfimo dos comprimentos de curvas ligando ambos. Vale considerar um critério de suavidade, como curvas de classe $\mathcal{C}^1$ por partes ou de classe $\mathcal{C}^\infty$. Habitualmente tomaremos esta última[2]. Nos nossos modelos, construiremos distância justamente assim, sendo o ínfimo do comprimento de curvas ligando dois pontos, induzida por uma certa métrica e deixando clara a estrutura de variedade riemanniana destes espaços. É possível mostrar que uma função construída desta maneira é de fato uma métrica (no sentido de espaço métrico) em qualquer variedade riemanniana, mas a prova disto exige conhecimentos mais avançadas. Por isso, provaremos para os casos específicos que iremos tratar. Ainda, provaremos que existe e é única a curva que liga quaisquer dois pontos e cujo comprimento coincide com a distância (chamaremos de geodésica), sendo curvas análogas a retas, no caso euclidiano. A existência de tais curvas, apesar de ser uma propriedade dos espaços euclidiano e hiperbólico, não é verdade para qualquer espaço. Considere, por exemplo, $\mathbb{R}^2\backslash\{0\}$ com a métrica usual e tome dois pontos no eixo y, um positivo e outro negativo.

Vale observar que esta definição só faz sentido em variedades conexas, para que haja pelo menos uma curva ligando dois pontos.

Uma outra observação, agora um pouco mais técnica, é a seguinte. A definição formal de variedade demanda que esta seja um espaço topológico. É possível provar que a topologia induzida pela distância da forma como a construímos empata com a topologia que começamos, isto é, da variedade enquanto espaço topológico, e isto vale para toda variedade riemanniana (Hausdorff). Isso mostra que, em algum sentido, a distância que definimos é topologicamente consistente com nossa variedade.

Apesar de não provarmos que qualquer métrica riemanniana dá origem a uma distância bem definida, provaremos para uma classe bem geral de métricas. São métricas do tipo: $g_P(v, v) = f^2(P)\langle v, v\rangle$, onde $\langle \cdot, \cdot\rangle$ é o produto interno usual, definidas num subconjunto de $\mathbb{R}^2$. Esta classe de métricas receberá especial atenção no nosso texto. Na demonstração deste fato, usaremos que o produto interno usual induz de fato uma distância. A prova deste fato pode ser vista na próxima seção.

Lema 6.1. *Seja $U \subset \mathbb{R}^2$ um aberto conexo por caminhos e seja $f : U \to \mathbb{R}$ uma função contínua. Se g é uma métrica riemanniana em U definida por*

$$g_P(v, v) = f^2(P)\langle v, v\rangle$$

[2]Dá para mostrar que as duas escolhas citadas levam sempre aos mesmos resultados, mas a prova disto é difícil.

Então a distância d_f induzida pelo ínfimo dos comprimentos l_f de curvas $\mathcal{C}^\infty$ por partes :

$$l_f(\gamma) = \int |f|(P)||\gamma'(t)||dt$$

É uma métrica. □

Demonstração. Antes de tudo, notamos que a distância está bem definida pelo fato do domínio ser conexo. Assim, sempre haverá pelo menos uma curva ligando dois pontos do domínio e estaremos tomando o ínfimo de um conjunto sempre não-vazio.

É evidente que o integrando é sempre positivo, donde $l(\gamma) \geqslant 0$, para toda curva possível γ, e portanto $d(P,Q) \geqslant 0$, $\forall P, Q \in U$. Tomando $\gamma(t) = P$, a curva constante, temos que $\gamma'(t) = 0$, $\forall t$, e daí o integrando vai ser identicamente nulo, resultando em $l(\gamma) = 0$ e portanto $d(P,P) \leqslant 0$, $\forall P \in U$. Usando ambas as desigualdades, chegamos em $d(P,P) = 0$, $\forall P \in U$.

Seja agora $P \neq Q \in U$. Como U é aberto, existe $r_1 > 0$ tal que $B[P, r_1] \subset U$ (bola fechada com distância euclidiana). Além disso, seja $r_2 = (1/2).d_{euc}(P,Q) > 0$ e tome $r = \min\{r_1, r_2\} > 0$ de forma que $K := B[P,r] \subset U$. K é compacto, donde $|f|$ assume um mínimo em K. Seja $m > 0$ tal mínimo. Mostrarei que $m.r > 0$ é uma cota inferior para o conjunto dos comprimentos $l_f(\gamma)$, com γ adequada. Tome uma tal curva $\gamma : [a,b] \to \mathbb{R}^2$ arbitrária. Não é difícil ver que γ irá interceptar ∂K em pelo menos um ponto, seja $R \in \mathrm{Im}\gamma \cap \partial K$ o primeiro deles e seja $t^* \in [a,b]$ tal que $\gamma(t^*) = R$. Então:

$$l_f(\gamma) = \int_a^b f(\gamma(t))||\gamma'(t)||dt \geqslant \int_a^{t^*} f(\gamma(t))||\gamma'(t)||dt \geqslant m \int_a^{t^*} ||\gamma'(t)||dt$$

Pois $\gamma(t) \in K$, $\forall t < t^*$. Note que a última integral é o comprimento euclidiano da curva γ restrita a $[a, t^*]$ que é maior que a distância euclidiana de $\gamma(a) = P$ a $\gamma(t^*) = R$, por definição de distância euclidiana. Isto é, $l_f(\gamma) \geqslant |R - P|$. Agora, basta ver que, como $R \in \partial K$, $|R - P| = r > 0$, e aí $l_f(\gamma) > m.r$. A implicação $P \neq Q \implies d(P,Q) > 0$ segue da arbitrariedade de P, Q e γ.

Para ver que $d(P,Q) = d(Q,P)$, basta ver que reparametrizações que invertem o sentido da curva são difeomorfismos e portanto não alteram o valor do comprimento da curva, pelo teorema da mudança de variáveis na integral.

Por fim, provemos a desigualdade triangular. Sejam $P, Q, R \in U$ arbitrários. Pela definição de ínfimo, para todo n existem curvas γ_n e μ_n tais que $l(\gamma_n) \leqslant d(P,Q) + 1/2n$ e $l(\mu_n) \leqslant d(Q,R) + 1/2n$. Tomando α_n como as concatenações de γ_n e μ_n, sendo eventualmente necessário reparametrizações, teremos que:

$$d(P,R) \leqslant l(\alpha_n) = l(\gamma_n) + l(\mu_n) \leqslant d(P,Q) + d(Q,R) + \frac{1}{n}$$

Como esta desigualdade vale para todo n, obtemos a desigualdade triangular: $d(P,R) \leqslant d(P,Q) + d(Q,R)$ para quaisquer pontos. ■

6.2.2 Diferenciais

Lembre-se de que, se f é uma função diferenciável entre $\mathbb{R}^n$ e $\mathbb{R}^m$, então sua diferencial é uma aplicação que a cada ponto $p \in \mathbb{R}^n$ associa uma transformação linear $df_p : \mathbb{R}^n \to \mathbb{R}^m$. Em verdade, a transformação linear df_p é a transformação linear que melhor aproxima f nas proximidades do ponto p. A diferencial possui uma propriedade bastante interessante. Se $\gamma'(t)$ é o vetor tangente a uma curva γ no ponto p, então $df_p(\gamma'(t))$ será o vetor tangente à curva $f \circ \gamma$ no ponto $f(p)$. Esta propriedade parece nos indicar que, enquanto pontos se transformam através de funções do tipo f, vetores se transformam com os respectivos diferenciais. Esta propriedade motiva a generalização de diferencial para variedades.

Se M, N são variedades diferenciáveis, e $f : M \to N$ é uma função diferenciável, então a diferencial de f é a aplicação que a cada ponto $p \in M$ associa uma transformação linear entre os espaços vetoriais T_pM e $T_{f(p)}N$ (T_pM é o espaço tangente a p na variedade M) de maneira que se γ é uma curva em M e $\gamma'(t) \in T_pM$, então $df_P(\gamma'(t)) = (f \circ \gamma)'(t) \in T_{f(p)}N$.

A versão generalizada da regra da cadeia diz que, se $f : M \to N$ e $g : N \to O$, então:

$$d(g \circ f)_p = (dg_{f(p)}) \circ (df_p)$$

6.2.3 Conformalidade

À luz do que foi discutido, tendo também em mente a discussão da seção 5.3 sobre transformações conformes, vemos que uma transformação é conforme se sua diferencial preserva ângulos. Sabemos que uma transformação que preserva produto interno é uma transformação ortogonal. Entretanto, preservar ângulos é mais fraco que preservar produto interno. Geometricamente, mudar a norma dos vetores altera o produto interno mas não altera o ângulo. Além disso, o ângulo entre dois vetores é usualmente definido através da divisão do produto interno pelas normas dos vetores. Assim, o diferencial de uma transformação conforme não necessariamente precisa ser ortogonal, pode ser um múltiplo disto.

Como o múltiplo pode depender do ponto, temos uma transformação é conforme se seu diferencial é da forma $u(x)O(x)$, onde $O(x)$ é uma transformação ortogonal para todo x. Uma transformação conforme preserva orientação de u é uma função positiva e se $O(x)$ tem determinante positivo para todo x.

Dando prosseguimento ao raciocínio, é fácil ver que se duas métricas diferem apenas por uma constante, então os ângulos são preservados. Assim, fizemos que duas métricas riemannianas são conformemente equivalentes se $h = ug$, com u uma função em M.

6.2.4 Áreas

As funções mais famosas do produto interno são a de medir comprimento de vetores e medir ângulo entre dois vetores. Entretanto, uma função menos óbvia está relacionada com volume. De fato, se V é um espaço vetorial com produto interno, então o k-volume do k-paralelepípedo formado pelos vetores $v_1, ..., v_k$ pode ser obtido tomando a raíz quadrada do determinante da matriz cuja entrada a_{ij} é $\langle v_i, v_j \rangle$. Podemos usar esta propriedade para calcular o k-volume de uma variedade de dimensão k. Dado um sistema de coordenadas locais $\phi(P) = (x_1, ..., x_k)$, podemos escrever g_P como uma matriz $g_{ij}(P) = g_{ij}(x_1, ..., x_k)$. Definimos então o k-volume de um conjunto $E \subset M$ por:

$$\mu(E) = \int_{\phi(E)} \sqrt{\det(g_{ij}(x_1, ..., x_k))} dx_1...dx_k$$

No caso de uma métrica do tipo $g_P(v,v) = f^2(P)\langle v, v \rangle$, a matriz de g será do tipo $f^2.Id$ (na verdade, $(f \circ \phi^{-1})^2$), onde Id é a matriz identidade k x k. O determinante desta matriz será então f^{2k} e daí:

$$\mu(E) = \int_{\phi(E)} f^k dx_1...dx_k$$

No caso específico em que a dimensão da nossa variedade é 2 (que será o caso em que estaremos principalmente focados aqui), a área será simplesmente a integral da função que multiplica o produto interno.

6.3 Revisão de Geometria Euclidiana

Nesta seção, iremos apresentar o roteiro que seguiremos para os modelos do espaço hiperbólico e usaremos o plano euclidiano, que já nos é muito familiar, como primeiro exemplo. Antes de começarmos, é interessante ponderar um pouco sobre o significado do termo "geometria". Felix Klein, dentro do programa de Erlangen, define geometria da seguinte forma: "Dado um conjunto munido de uma estrutura e de um grupo de transformações que preservam tal estrutura, geometria é o estudo dos objetos que são invariantes por tais transformações. Nosso roteiro será então o seguinte: tomaremos um conjunto, definiremos uma distância nele (a partir da estrutura de variedade riemanniana, pelo processo já mencionado) e acharemos o grupo de transformações que preservam esta estrutura, isto é, o grupo de isometrias. Estudaremos também, em cada modelo, objetos invariantes por isometria, como geodésicas (informalmente, curvas que minimizam distância entre pontos) e áreas.

No caso da geometria euclidiana, o conjunto é $\mathbb{R}^2$ (ou n, mais geralmente). Neste caso específico, apesar de também ser uma variedade riemanniana, o plano possui uma estrutura natural de espaço métrico e portanto há um jeito fácil de escrever a métrica sem recorrer a ínfimo de comprimentos. Apesar desta estrutura métrica natural, iremos tratar o plano como uma variedade riemanniana dotada do produto interno usual e definir distância como mencionado acima, para sermos coerentes com nosso programa. Acharemos também o grupo de isometrias, usualmente denotado por $\text{Isom}(\mathbb{R}^2)$ e o grupo de isometrias que preservam orientação, denotado por $\text{Isom}^+(\mathbb{R}^2)$.

A métrica do plano euclidiano enquanto variedade riemanniana é dada por:

$$g_P(v,v) = \langle v, v \rangle$$

Onde $\langle \cdot, \cdot \rangle$ denota o produto interno usual. Note que a métrica não depende do ponto P, isto é, é uma métrica constante. Por algumas razões técnicas, este tipo de métrica é uma classe de métricas interessante de ser considerada. Veremos que as métricas dos modelos do espaço hiperbólico não são constantes, com exceção do modelo do hiperboloide. Isto é parte do motivo de considerarmos este modelo mais natural.

Com isso, o comprimento de uma curva γ é:

$$L_{euc}(\gamma) = \int \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt = \int ||\gamma'(t)|| dt$$

Para o caso euclidiano, faremos uma pequena alteração no roteiro. Em vez de provarmos primeiro que a distância definida usando ínfimo de comprimentos de curvas está bem definida, usaremos a expressão acima para acharmos o comprimento de uma reta ligando dois pontos e depois provaremos que qualquer outra curva que os conecte tem comprimento maior, donde o comprimento da reta empata com a distância entre os pontos. Desta forma, teremos uma expressão para a distância entre pontos e com ela concluiremos que de fato satisfaz as propriedades desejadas.

Em primeiro lugar, deve ser provado que o comprimento da reta ligando pontos P e Q coincide com $||Q - P||$. Isso mostra que a distância entre dois pontos P e Q é menor ou igual a $||Q - P||$. Para a outra desigualdade, basta mostrar que o comprimento de qualquer curva é maior ou igual a $||Q - P||$. Isto se faz usando a desigualdade de Cauchy-Schwartz para estimar $||\gamma'(t)||$ por baixo usando o $||Q - P||$. Depois disto, usa-se a relação do produto interno com a derivada de forma a poder se usar o teorema fundamental do cálculo. Uma vez que a distância entre dois pontos é dado por uma norma, segue que de fato é uma métrica, e ainda coincide com a usual.

Uma maneira alternativa de provar que retas realizam distância é provar inicialmente para o caso mais fácil em que os dois pontos estão na mesma abscissa e posteriormente usar rotações e translações para atingir quaisquer dois pontos.

Agora que já definimos a distância e justificamos que ela de fato o é (e de quebra achamos geodésicas), caracterizaremos o grupo de isometrias do espaço euclidiano. Você pode começar verificando que as translações e as transformações ortogonais são isometrias. De fato, toda isometria de $\mathbb{R}^n$ pode ser escrita como a composição de uma transformação corresponde a alguma matriz de $O(n, \mathbb{R})$ com uma translação. É matéria de fácil verificação que o conjunto destas funções forma um grupo. (e também, que coincide com o grupo gerado pelos subgrupos de transformações ortogonais e translações)

No caso específico do plano, pode ser mostrado que as transformações ortogonais podem ser apenas rotações em torno da origem ou reflexões em torno de uma linha. Para os leitores familiarizados com o estudo dos grupos de Lie, lembre-se que $O(2,\mathbb{R})$ é um grupo de Lie de uma dimensão, e portanto o grupos das isometrias do plano é um grupo de Lie de três dimensões.

Teorema 6.2. *Seja $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ uma isometria. então existem $A \in O(2,\mathbb{R})$ e $p \in \mathbb{R}^2$ tais que $\phi(v) = Av + p$, $\forall v \in \mathbb{R}^2$.* □

A ideia da demonstração é definir uma função $\psi(v) = \phi(v) - \phi(0)$, de forma que ψ fixe a origem. Dado que ϕ é isometria, é fácil ver que ψ também será. Daí, $||\psi(v)|| = d(\psi(v), 0)$. Como ψ preserva origem, $||\psi(v)|| = d(\psi(v), \psi(0)) = d(v, 0) = ||v||$, donde ψ preserva norma. Como a norma vem de um produto interno, temos que $2\langle v, w\rangle = ||v||^2 + ||w||^2 - ||v - w||^2$ e disto pode ser tirado que ψ preserva produto interno. Ainda, destes fatos podemos concluir que $||\psi(v+w) - \psi(v) - \psi(w)||^2 = 0$ e ——$\psi(tv) - t\psi(v)||^2 = 0$, e portanto ψ é linear.

Sendo ψ linear e que preserva produto interno, é fácil ver que é ortogonal. Dada uma base, vai haver uma matriz A que representa ψ. Daí, $\langle Av, Aw\rangle = (Av)^T(Aw) = v^T A^T Aw$ necessariamente deve ser igual a $\langle v, w\rangle = v^T w$, isto é, temos que $v^T A^T Aw = v^T w$. escolhendo v e w de maneira esperta podemos obter os componentes da matriz $A^T A$ e não é difícil ver que vai ser a identidade.

Chamando $\phi(0)$ de p, temos que $\phi(v) = Av + p$ com A ortogonal, como gostaríamos.

Seja $E(2)$ o subgrupo das bijeções de $\mathbb{R}^2$ gerado pela união dos conjuntos de transformações ortogonais e das translações. Pela proposição 1.28, este grupo empata com o grupo de todas as concatenações finitas de transformações ortogonais e translações. Não é difícil ver que esta concatenação sempre vai dar funções da forma do teorema anterior (6.2). Considere a composição de duas destas e depois use indução. Seja $\phi(v) = A(v) + p$ e $\psi(v) = B(v) + q \in E(2)$. Então $\phi \circ \psi = A(B(v) + q) + p = A(B(v)) + A(q) + p$. Como a composição de transformações ortogonais é ortogonal, seja $C \in O(2,\mathbb{R})$ tal que $C = A \circ B$ e seja $r = A(q) + p$. Então $\phi \circ \psi = C(v) + r$. Além disso, funções inversas também são da mesma forma. Tome $\xi(v) = D(v) + s$ onde $D = A^{-1}$ e $s = -D(p)$. Então $\xi \circ \phi = D(A(v) + p) + s = D(A(v)) + D(p) + s = A^{-1}(A(v)) + D(p) - D(p) = v$. Isto demonstra que $E(2) = Isom(\mathbb{R}^2)$.

Afirmamos, ainda, que $Isom^+(\mathbb{R}^2)$ é o subgrupo gerado pelas transformações lineares cujas matrizes estão em $SO(2,\mathbb{R})$ e pelas translações. Isto empata com as funções que podem ser escritas como $\phi(v) = Av + p$ com $A \in SO(2,\mathbb{R})$ e $p \in \mathbb{R}^2$. A demonstração é análoga àquela feita no parágrafo anterior.

Um assunto interessante, e que será de especial interesse neste estudo, é a chamada *tesselação*. Isto não é nada mais do que um ladrilhamento do plano, isto é, cobrir o plano com polígonos de forma que os polígonos não se superponham e de forma que vértices se encontrem com vértices. Caso os polígonos sejam regulares, as únicas tesselações possíveis do plano são com triângulos, quadrados ou hexágonos. Na geometria hiperbólica, entretanto, há um número infinito de tesselações possíveis, revelando uma estrutura mais rica do que a da geometria euclidiana. Os grupos fuchsianos, pelos quais estamos interessados, serão ferramentas importantes no estudo das tesselações do plano hiperbólico. Há uma associação entre uma tesselação e um grupo de isometria (o grupo de isometria que preserva a tesselação), conectando assim álgebra com geometria. E, no caso da geometria hiperbólica, estes grupos serão os fuchsianos.

6.4 O Semiplano

Partiremos agora, de fato, para a geometria hiperbólica. Começaremos considerando o modelo do semiplano complexo superior.

O conjunto que servirá de substrato para este modelo é $\mathbb{H} := \{z \in \mathbb{C}; \mathrm{Im}(z) > 0\} \approx \{(x, y) \in \mathbb{R}^2; y > 0\}$. Usaremos a extensão do plano complexo para abranger ∞ quando necessário, daí, rapidamente nota-se que $\mathbb{H}$ corresponde a um dos hemisférios (aberto) da esfera de Riemann.

Nossa métrica riemanniana será dada por:

$$g_P(v, v) = \frac{1}{y^2}\langle v, v\rangle$$

Onde $P = (x, y)$. A norma será então:

$$||v||_{\mathbb{H}} = \frac{1}{y}||v||$$

O comprimento de curvas de classe $\mathcal{C}^\infty$ por partes $\gamma : I \to \mathbb{H}$ será:

$$l_{\mathbb{H}}(\gamma) = \int_I \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)}dt$$

E a distância entre dois pontos como o ínfimo do valor do comprimento de todas as curvas $\mathcal{C}^\infty$ por partes que os ligam.

Chamaremos $l_{\mathbb{H}}$ e d_{hyp} apenas de l e d, mas ficará claro pelo contexto.

Para ver que esta métrica de fato induz uma distância, basta notar que ela é do tipo $g_P(v, v) = f^2(P)\langle v, v\rangle$ com $f(x, y) = 1/y$ e usar o lema 6.1.

Antes de passarmos a estudar o grupo de isometrias, analisaremos nesta seção algumas características particulares da distância hiperbólica definida. Isto servirá tanto para nos familiarizarmos e criarmos intuição com a geometria hiperbólica, tanto para deduzirmos resultados preliminares que serão necessários nas próximas etapas. Na verdade, há diversos resultados preliminares que precisam ser provados antes de conseguirmos achar o grupo de isometrias. Além de estudarmos como a distância hiperbólica se comporta, precisamos primeiros ver algumas isometrias e depois estudar geodésicas e algumas de suas propriedades antes de conseguirmos achar o grupo de isometrias.

Vamos começar abertamente comparando e confrontando a distância hiperbólica com a euclidiana. Note que o comprimento de curvas verticais ligando os pontos (x_0, a) e (x_0, b) é $\log(\frac{b}{a})$, o que significa que quanto menor for a, maior é o comprimento da curva, tendendo pra infinito. Essa é uma das razões pelas quais dizemos que o eixo real é chamado de círculo no infinito (o “círculo”vem da esfera de Riemann). O mesmo acontece quando b vai a infinito, mas é possível ver que pedaços de cima contribuem menos para o comprimento da curva. Por exemplo se temos uma reta de $(x_0, 1)$ para $(x_0, 3)$, esperaríamos, de acordo com a geometria euclidiana, que a reta de $(x_0, 1)$ a $(x_0, 2)$ tivesse o mesmo comprimento da reta de $(x_0, 2)$ a $(x_0, 3)$, mas na verdade, a primeira parte tem comprimento $\log 2 \approx 0,69 > 0,41 \approx \log 3/2$, que é o comprimento da segunda parte.

Por outro lado, retas horizontais ligando (a, y_0) a (b, y_0) possuem comprimento $(b-a)/y_0$. Isto é completamente novo com relação a nossa intuição de geometria euclidiana. O comprimento de retas com os mesmos extremos é tão maior quanto menor for a ordenada em que ela se encontra.

Vemos, de modo geral, que caminhar em pontos de ordenada pequena tende a contribuir bastante para o aumento do comprimento. Assim, nada mais razoável que imaginar que andar em linha “reta”entre dois pontos de diferentes abscissas talvez não seja o mais eficiente, mas sim subir e descer, tomando também cuidado para não subir e descer muito, o que também não seria eficiente. Assim, é possível imaginar que o caminho mais eficiente entre dois pontos de abscissas próximas, se existir (de fato, sempre vai existir) sobe e desce menos do que se forem mais afastados. Em verdade, veremos que os caminhos mais eficientes são dados por arcos de circunferência com centro no eixo x, no caso de pontos com abscissas diferentes (isto não é óbvio, poderiam ser pedaços de retângulo ou arcos de parábola).

Com estas propriedades, fica fácil ver, ainda que sem provar, que o centro de uma bola na distância hiperbólica é deslocada para baixo. Também não iremos provar, mas o formato da bola é o mesmo da bola euclidiana, com exceção do centro estar deslocado.

Agora, veremos uma maneira de estimar por baixo o valor da distância entre dois pontos quaisquer.

Proposição 6.3. *Sejam $P = (x_0, y_0)$ e $Q = (x_1, y_1)$ pontos de $\mathbb{H}$ quaisquer. Supondo sem perda de generalidade que $y_1 \geqslant y_0$ Então $d(P, Q) \geqslant \log(y_1/y_0)$.* □

Demonstração. Seja $\gamma : [a, b] \to \mathbb{H}$ uma curva de classe $\mathcal{C}^\infty$ por partes $\gamma = (x(t), y(t))$ tal que $\gamma(a) = P$ e $\gamma(b) = Q$. Naturalmente teremos $x'(t)^2 \geqslant 0$, donde

$$l(\gamma) = \int_a^b \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)} dt \geqslant \int_a^b \frac{\sqrt{y'(t)^2}}{y(t)} dt = \int_a^b \frac{|y'(t)|}{y(t)} dt = \int_a^b [log(y(t))]' dt$$

$$= \log\left(\frac{y(b)}{y(a)}\right) = \log\left(\frac{y_1}{y_0}\right)$$

■

Note que, apesar de valer para quaisquer dois pontos, esta estimativa deixa de ser útil para pontos de mesma ordenada, pois a estimativa dá zero e não acrescenta nada de novo. Entretanto, veremos uma utilidade fundamental desta estimativa no seguinte

Corolário 6.3.1. *Dados dois pontos P e Q na mesma abscissa, digamos, $P = (x_0, y_0)$ e $Q = (x_0, y_1)$, com $y_1 \geqslant y_0$ temos que $d(P,Q) = \log(y_1/y_0)$ e este valor é alcançado (unicamente) pelo segmento de reta vertical que os une, isto é, a reta realiza o menor comprimento.* □

Demonstração. Basta lembrar que a reta possui o mesmo comprimento da cota inferior estabelecida na proposição. Para a unicidade, note que, se a curva não for um segmento de reta, então teremos $x'(t) > 0$ para algum t (na verdade, para um certo intervalo, pela continuidade de $x'(t)$), donde a integral que fornece o comprimento será estritamente maior que a da reta vertical. ■

Neste caso, vemos que a reta realiza o menor comprimento tanto da geometria hiperbólica quanto euclidiana. Na verdade, isto só ocorre neste caso, como veremos adiante.

Apesar de já sabermos que a distância no semiplano está bem definida, apresentaremos uma demonstração alternativa que utiliza mais diretamente as características de $\mathbb{H}$. Em especial, a estimativa que acabamos de deduzir.

Proposição 6.4. $(\mathbb{H}, d)$ *é um espaço métrico.* □

Quero mostrar agora que pontos diferentes possuem distância maior de zero. Caso os pontos tenham ordenadas diferentes, isto pode ser visto usando o corolário 6.3.1. Vejamos agora o caso em que as ordenadas são iguais (e, portanto, as abscissas serão necessariamente diferentes). Seja $\gamma : [a,b] \to \mathbb{H}$ uma curva. Se a ordenada desta curva é limitada por $y^* = y_0 + 1$, podemos estimar o comprimento da curva por baixo supondo que ela está inteiramente nesta ordenada (isto vem do fato já discutido que quanto maior é a ordenada, menor é o comprimento da curva). Formalmente:

$$l_{\text{hyp}}(\alpha) = \int_a^b \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)} dt \geqslant \int_a^b \frac{|x'(t)|}{y(t)} dt \geqslant \int_a^b \frac{|x'(t)|}{y^*} dt \geqslant \frac{x_1 - x_0}{y^*} > 0$$

Se, entretanto, a curva não for limitado por y^*, pelo teorema do valor intermediário, deve haver um t^* tal que $\gamma(t^*) = y^*$. Então, o comprimento da curva será maior que o comprimento da restrição para $[a, t^*]$, que por sua vez terá um comprimento maior do que o comprimento se a curva tivesse apenas ordenada em y^*. Formalmente:

$$l_{\text{hyp}}(\alpha) \geqslant \int_a^{t^*} \frac{|y'(t)|}{y(t)} dt \geqslant \log\left(\frac{y^*}{y_0}\right) > 0$$

Uma vez que $y^* > y_0$ por definição. Em qualquer hipótese, temos:

$$l_{\text{hyp}}(\alpha) \geqslant \min\left\{\frac{x_1 - x_0}{y^*}, \log\left(\frac{y^*}{y_0}\right)\right\} > 0$$

que conclui nossa demonstração.

As outras propriedades se fazem do mesmo modo do lema 6.1.

6.4.1 Algumas isometrias

Relembramos que o semiplano superior pode ser tratado tanto como parte de $\mathbb{R}^2$ como $\mathbb{C}$. O intercâmbio entre os dois modos de ver ficará mais nítido nesta seção.

Daremos agora alguns exemplos de isometrias do plano hiperbólico. Para provar que uma certa função ϕ é uma isometria, note que é suficiente mostrar que, para toda curva γ, vale que $l(\phi \circ \gamma) = l(\gamma)$

Funções do tipo $T_\lambda : \mathbb{H} \to \mathbb{H}$, $T_\lambda(z) = \lambda z$ com $\lambda > 0$ são chamadas de homotetia. Estas funções criam o efeito de *zoom in* ou *zoom out* focadas na origem. Não é difícil ver que, para qualquer curva, $l(T_\lambda \circ \gamma) = l(\gamma)$, donde homotetias são isometrias. Este fato pode parecer surpreendente a princípio, dado o efeito de ampliação e redução que estas funções exibem, mas na verdade, estas funções também deslocam os objetos para cima e baixo. Por exemplo, uma reta que vai de $(-1, 1)$ a $(1, 1)$, se torna uma reta que vai de $(-2, 2)$ a $(2, 2)$. A reta aumenta de comprimento (no sentido euclidiano), mas também sobe no plano e um efeito anula o outro. Em coordenadas polares, é possível ver que uma homotetia leva um ponto de (r, θ) para $(\lambda r, \theta)$, aumentando a distância da origem mas mantendo o ângulo com o eixo real.

Translações no eixo real, do tipo $L_r : \mathbb{H} \to \mathbb{H}$ dadas por $L_r(z) = z + r$ com $r \in \mathbb{R}$ também são isometrias. Translações imaginárias não são. Isso mostra como o plano é simétrico na horizontal mas "não simétrico"na direção vertical, confirmando o que já sabíamos.

A reflexão $R(x, y) = (-x, y)$ também é uma isometria. Em notação complexa, $R(z) = -\bar{z}$.

Por fim, temos uma função que merece uma atenção um pouco maior. É a inversão. Dado um ponto P, a inversão leva este ponto para um outro ponto Q da reta OP, de modo que $d(0, Q) = 1/d(0, P)$. É fácil ver que esta função em coordenadas polares será dada por $I(r, \theta) = (1/r, \theta)$. Passando para coordenadas cartesianas, teremos $I(x, y) = \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right)$, e para números complexos, $I(z) = 1/\bar{z}$. Note que esta função deixa o (semi) círculo unitário invariante. A ação dela é semelhante a da homotetia, isto é, apesar de reduzir e ampliar coisas, esta função leva coisas de um lugar do plano para outro, onde as contribuições para a distância são maiores ou menores.

Algo que não foi comentado mas que é muito importante e fácil de ser provado, é de que todas estas funções são bijetivas, caso contrário, não poderiam ser isometrias.

Apenas com translações e homotetias já é possível ver que o plano hiperbólico é um espaço homogêneo:

Proposição 6.5. *O espaço métrico* $(\mathbb{H}, d)$ *é homogêneo.* □

Demonstração. Sejam $P = (x_0, y_0)$ e $Q = (x_1, y_1)$ pontos quaisquer de $\mathbb{H}$. Note que $\phi_1(0, 1) := (L_{x_0} \circ T_{y_0})(0, 1) = P$, donde $(0, 1) = \phi_1^{-1}(P)$. Analogamente, $\phi_2(0, 1) := (L_{x_1} \circ T_{y_1})(0, 1) = Q$, donde $(\phi_2 \circ \phi_{-1})(P) = Q$. Como inversas e compostas de isometrias também são isometrias, para qualquer par de pontos achamos uma isometria que leva um no outro, donde a afirmação segue. ■

Enunciaremos agora, sem provar, um lema que nos ajudará a estudar as geodésicas do plano.

Lema 6.6. *Sejam $\gamma : I \to \mathbb{H}$ uma curva $\mathcal{C}^\infty$ por partes e $\phi : \mathbb{H} \to \mathbb{H}$ uma isometria. Então $l(\gamma) = l(\phi \circ \gamma)$.* □

6.4.2 Geodésicas

Caracterizaremos agora as geodésicas do plano hiperbólico, isto é, informalmente, as curvas que desempenham o papel das retas na geometria euclidiana. Já vimos que retas também desempenham este papel para pontos com a mesma abscissa e veremos agora que os arcos de circunferência com centro no eixo real desempenham esta função para o outro caso. Antes disso, vejamos que dados dois pontos é único o arco de circunferência nestas condições.

Lema 6.7. *Sejam $P = (x_0, y_0)$ e $Q = (x_1, y_1)$ pontos de $\mathbb{H}$. Se $x_0 \neq x_1$, então existe e é única a circunferência de centro no eixo real (isto é, centro $(c, 0)$) que passa pelos dois pontos.* □

Demonstração. Tome o ponto $(c, 0)$ tal que:

$$c = \frac{x_0^2 + y_0^2 - x_1^2 - y_1^2}{2(x_0 - x_1)}$$

e $r = ||P-(c, 0)||$ e veja que o círculo de centro $(c, 0)$ e raio r satisfaz as condições especificadas. Agora, igualando $||P-(c, 0)|| = ||Q-(c, 0)||$, é possível ver que c necessariamente deve ter a fórmula descrita acima, donde vem a unicidade. ■

A prova de que arcos de circunferência minimizam distância utiliza o lema 6.7 e o fato que existem isometrias que levam retas verticais nos arcos. Para ver isto usaremos o lema a seguir.

Lema 6.8. *Sejam $P = (x_0, y_0)$ e $Q = (x_1, y_1)$ pontos de $\mathbb{H}$. Se $x_0 \neq x_1$, então existe $\phi : \mathbb{H} \to \mathbb{H}$ isometria tal que $\phi(P)$ e $\phi(Q)$ possuem a mesma abscissa. Além disso, se $\gamma : [p, q] \to \mathbb{H}$ é o arco de semicírculo ligando P e Q, então a imagem por ϕ de $\gamma([p, q])$ é o segmento de reta entre $\phi(P)$ e $\phi(Q)$.* □

Demonstração. Seja C o semicírculo que passa por P e Q. Sejam c e r a abscissa do centro e o raio de C, respectivamente. Tome então a translação L_{r-c}. Seja α uma parametrização em coordenadas polares de $L_{r-c}(C)$, de forma que o raio de α coincida com seu centro e a circunferência se torne tangente ao eixo imaginário. Deveremos ter, necessariamente:

$$(\rho(t)\cos(\theta(t)) - r)^2 + \rho(t)^2 \sin^2(\theta(t)) = r^2 \implies$$

$$\rho(t)^2 \cos^2(\theta(t)) - 2r\rho(t)\cos(\theta(t)) + r^2 + \rho(t)^2 \sin^2(\theta(t)) = r^2$$

$$\implies \rho(t)^2 = 2r\rho(t)\cos(\theta(t)) \implies \rho(t) = 2r\cos(\theta(t))$$

Tomando então $\theta = t$, ficamos com a curva $\alpha : (0, \pi/2) \to \mathbb{H}$, $\alpha(t) = (2r\cos(t), t)$, em coordenadas polares. Seja I a inversão. Então

$$(I \circ \alpha)(t) = \left(\frac{1}{2r\cos(t)}, t\right)$$

.

Passando para coordenadas cartesianas, ficamos com:

$$(I \circ \alpha)(t) = \left(\frac{1}{2r}, \frac{\tan(t)}{2r}\right)$$

.

E é fácil ver que esta reta parametriza a semirreta $x = 0$. Como I e L_{r-c} são isometrias e $P, Q \in C$, segue que a isometria $I \circ L_{r-c}$ leva P e Q para o eixo imaginário. ■

Teorema 6.9. *Sejam $P = (x_0, y_0)$ e $Q = (x_1, y_1)$ pontos de $\mathbb{H}$. Se $x_0 \neq x_1$, então o arco de circunferência C_{PQ} ligando P a Q é a única curva com a propriedade de que $l(C_{PQ}) = d(P, Q)$.* □

Demonstração. Pelo lema anterior (6.8), sabemos que existe uma isometria ϕ que leva C_{PQ} em um segmento de reta no eixo imaginário. Assim, ϕ^{-1} é uma isometria que leva a um segmento de reta vertical no arco C_{PQ}. Uma vez que isometrias levam curvas de menor comprimento em curvas de menor comprimento (falta provar) e que segmentos de reta verticais são curvas minimizantes, segue que $l(C_{PQ}) = d(P, Q)$. ■

Até agora temos utilizado "curvas minimizantes de comprimento" e "geodésicas" como sinônimos. A próxima definição, entretanto, desfará a confusão e veremos que geodésicas são classes de curvas mais gerais, que só precisam minimizar comprimentos localmente.

Definição 6.1. Uma curva $\alpha : I \subset \mathbb{R} \longrightarrow \mathbb{H}$ que é C^1 por partes é dita uma geodésica se para todo $t \in I$, existe $\epsilon(t) > 0$ tal que:

$$d_{\text{hyp}}(\alpha(t), \alpha(s)) = l_{\text{hyp}}(\alpha|_{I_{st}})$$

Para todo $s \in I$ com $|s - t| < \epsilon(t)$, onde:

$$I_{st} = \begin{cases} [s, t], \text{ se } s < t; \\ [t, s], set \text{ ¡ s} \end{cases}$$

Uma curva $\alpha : I \subset \mathbb{H} \longrightarrow \mathbb{H}$ é dita uma geodésica completa se não existe uma extensão $\beta : J \longrightarrow \mathbb{H}$ de α, com $I \subset J$, i.e, se $\beta : J \longrightarrow \mathbb{H}$ é uma geodésica e $\beta|_I = \alpha$, então $I = J$ e $\alpha = \beta$.

♣

Não bastaria dizer que a curva é minimizante (isto é, que o comprimento entre dois pontos da curva dá no mesmo que a distância entre eles)? Por que introduzimos o critério do ϵ? Pense no caso de uma esfera. Imagine um grande círculo quase completamente ao redor da esfera. Como exemplo concreto, pense num grande círculo ao redor da terra que passe por São Paulo e pelo Rio de Janeiro, mas que não seja completo, isto é, há um buraco na curva entre as duas cidades (na parte mais curta ligando as duas), e elas só podem ser ligados dando toda a volta pelo globo no outro sentido, isto é, pela parte do grande círculo mais longa ligando as duas. Nesse caso, não vale que a distância entre São Paulo e o Rio de Janeiro é o comprimento da curva que passa entre os dois. O comprimento da curva é todo o diâmetro da Terra menos a distância entre Rio e São Paulo. Então, a curva não minimiza a distância entre quaisquer dois pontos dela, mas minimiza localmente, para pontos próximos o que motiva a definição acima.

O teorema abaixo estabelece um critério definitivo sobre quais curvas são geodésicas completas.

Teorema 6.10. *Seja $\gamma : I \to \mathbb{H}$ um geodésica completa. Então, ou γ é uma semi-reta vertical, ou γ é um semi-círculo centrado na reta real.* □

A proposição seguinte diz que, no sentido do programa de Erlagen, geodésicas são objetos de estudo da geometria.

Proposição 6.11. *Seja $\gamma : [a, b] \to \mathbb{H}$ uma geodésica e M uma isometria. Então a curva $\mu : [a, b] \to \mathbb{H}$, $\mu = M \circ \gamma$ também é uma geodésica.* □

Demonstração. Tome $t \in [a, b]$. Como γ é geodésica, existe $\epsilon > 0$ tal que

$$|s - t| < \epsilon \implies d(\gamma(t), \gamma(s)) = l(\gamma|I_{st})$$

Tomando o mesmo ϵ e $|s - t| < \epsilon$, como M é isometria, vem que:

$$d(\mu(t), \mu(s)) = d(M \circ \gamma(t), M \circ \gamma(s)) = d(\gamma(t), \gamma(s)) = l(\gamma|I_{st})$$

Como isometrias preservam comprimento de curvas (lema 6.6), temos que $l(\gamma|I_{st}) = l(M \circ \gamma\|I_{st}) = l(\mu|I_{st})$, donde:

$$|s - t| < \epsilon \implies d(\mu(t), \mu(s)) = l(\mu|I_{st})$$

Pela arbitrariedade de t concluímos que isometrias levam geodésicas em geodésicas. ■

O lema a seguir mostra uma propriedade técnica das geodésicas e nos auxiliará a achar o grupo de isometrias do semiplano.

Lema 6.12. *Sejam $P = (0, y_0)$ e $Q = (0, y_1)$ com $0 < y_0 < y_1$. Então para toda geodésica completa $g \subset \mathbb{H}$ com $P \in g$, são equivalentes:*

- $d_{hyp}(P,Q) = d_{hyp}(Q,g)$;
- *$g = g_0$, onde $g_0 = S_{y_0} \cap \mathbb{H}$, sendo S_{y_0} o círculo euclidiano de centro na origem e raio y_0.*

□

A demonstração deste lema pode ser vista em detalhes em [Bon09]. Veremos aqui apenas a ideia da prova. A parte mais fácil é provar que a segunda afirmação implica na primeira. Para tanto, note que P é o ponto da circunferência cuja ordenada é um máximo global. Assim, tomando qualquer outro ponto $S = (x,y)$, podemos estimar a distância dele a Q por baixo usando o ponto $(0,y)$ (proposição 6.3). Por outro lado, a distância de P a $(0,y)$ será maior que a de P a Q, uma vez que $y_1 > y$, pelo que já foi comentado.

Para a recíproca, a prova será por absurdo, isto é, suporemos que a geodésica não está centrada na origem mas que mesmo assim a distância entre Q e a geodésica empata com $d(P,Q)$, isto é, não há outro ponto da geodésica com distância a Q menor que P. Chame a abscissa do centro da geodésica de a. Usaremos o Teorema da Função Implícita na função $F(x,y) = x^2 - 2ax + y^2 - y_0^2$ (certifique-se que $\partial_y F(0,y_0) \neq 0$) para obter uma função $s : (-\epsilon,\epsilon) \to$ tal que $t \mapsto (t, s(t))$ parametriza um trecho da geodésica. Para cada t, podemos calcular explicitamente a o comprimento da reta ligando Q a $(t,s(t))$, fazendo uma função $l(t)$. Agora, basta calcular $l'(0)$ e será possível ver que $l'(0) \neq 0$, donde $l(t)$ é estritamente monótona numa vizinhança de 0, e portanto há t_0 tal que $d(P,Q) = l(0) > l(t_0) \geqslant d(Q,(t_0,s(t_0)))$. No cálculo de $l'(0)$ usamos que $s'(0) \neq 0$, que pode ser visto usando a relação que o próprio Teorema da Função Implícita nos dá para derivada de s.

Corolário 6.12.1. *O lema vale ainda que as abscissas de P e Q não sejam zero.* □

Basta tomar translações horizontais e lembrar que são isometrias.

6.4.3 O Grupo de Isometrias

Nesta seção finalmente acharemos o grupo de isometrias do plano. De antemão, informo que este grupo tem ligações bem fortes com um outro grupo já conhecido por nós da seção 4.2: : $\text{Möb}_+(\mathbb{R})$. Recomendo que a seção citada seja lida antes desta. Relembro desta seção o fato importante que qualquer elemento de $\text{Möb}_+(\mathbb{R})$ pode ser escrito de forma que $ad - bc = 1$. Relembro também a notação f_M, $M \in SL(2,\mathbb{R})$ para as transformações de Möbius que estava sendo utilizada.

Primeiramente, vejamos que elementos desde grupo estão bem definidos tendo o plano hiperbólico como domínio e contradomínio.

Proposição 6.13. *Sejam $\phi \in Möb_+(\mathbb{R})$ e $z \in \mathbb{H}$. Então $\phi(z) \in \mathbb{H}$.* □

Demonstração. Como $\phi \in \text{Möb}_+(\mathbb{R})$, existem $a,b,c,d \in \mathbb{R}$ tais que $ad - bc = 1$ e:

$$\phi(z) = \frac{az+b}{cz+d}$$

Já sabemos que $\phi(z) \in \mathbb{C}$. Basta agora provarmos que $\mathrm{Im}(\phi(z)) > 0$. De fato:

$$\mathrm{Im}(\phi(z)) = \mathrm{Im}\left(\frac{(az+b)(c\bar{z}+d)}{|cz+d|^2}\right) = \mathrm{Im}\left(\frac{ac|z|^2 + adz + bc\bar{z} + bd}{|cz+d|^2}\right)$$

$$= \frac{(ad-bc)\mathrm{Im}(z)}{|cz+d|^2} = \frac{\mathrm{Im}(z)}{|cz+d|^2} > 0$$

■

Além disso, é possível mostrar que os elementos de $\text{Möb}_+(\mathbb{R})$ são biholomorfismos de $\mathbb{H}$ preservando portanto a orientação.

Há dois tipos especiais de transformações de $\text{Möb}_+(\mathbb{R})$, que são isometrias do plano hiperbólico já bem conhecidas. Dados $\lambda > 0$ e $r \in \mathbb{R}$, considere:

$$M_\lambda = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & 1/\sqrt{\lambda} \end{pmatrix}$$

$$N_r = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$$

Não é difícil ver que a primeira matriz corresponde a homotetias T_λ e a segunda a translações L_r. Note que não há nenhuma transformação em $\text{Möb}_+(\mathbb{R})$ que corresponde à inversão I, pois ela muda a orientação do espaço. Agora, aproveitaremos a semelhança de $\text{Möb}_+(\mathbb{R})$ com $SL(2,\mathbb{R})$ para reciclar um teorema importante:

Proposição 6.14. *Sejam $a, b, c, d \in \mathbb{R}$ tais que $ad - bc = 1$. Então as funções $\phi, \psi : \mathbb{H} \to \mathbb{H}$ tais que:*

$$\phi(z) = \frac{az+b}{cz+d}$$

$$\psi(z) = \frac{c\bar{z}+d}{a\bar{z}+b}$$

São isometrias. □

Demonstração. Caso $a \neq 0$, basta notar que

$$\psi(z) = (L_{\frac{c}{a}} \circ T_{a^2} \circ I \circ L_{\frac{b}{a}})(z)$$

Além disso, $\phi(z) = (I \circ \psi)(z)$. A afirmação então segue do fato que as funções citadas são isometrias como já visto e que composição de isometrias é isometria.

Se, por outro lado, $a = 0$, devemos ter necessariamente $c \neq 0$. Sejam $a' = c' = c$, $d' = d$ e $b' = b + d$. Daí, $a' \neq 0$. Além disso, $a'd' - b'c' = cd - cd - bc = -bc = ad - bc = 1$. Assim, podemos usar o resultado anterior e notar que a função $\zeta : \mathbb{H} \to \mathbb{H}$

$$z \mapsto \frac{a'z + b'}{c'z + d'} = \frac{cz + b + d}{cz + d}$$

É uma isometria. Agora, basta ver que $\phi(z) = (L_{-1}) \circ \zeta)(z)$, e portanto ϕ é isometria. O fato de ψ ser isometria vem do fato que $\psi = I \circ \phi$. ■

Lema 6.15. *Seja γ uma geodésica completa. Então existe $M \in SL(2, \mathbb{R})$ tal que $f_M \circ \gamma$ é a reta $L = \{z \in \mathbb{H}; Re(z) = 0\}$.* □

Demonstração. Sabemos, pelo teorema 6.10, que geodésicas completas podem ser ou retas verticais, ou semicírculos centrados na reta real. Se for o primeiro caso, a transformação que estávamos procurando será trivialmente uma translação, que já vimos pertencerem a Möb$_+(\mathbb{R})$. Suponha, do contrário, que seja o semicírculo. Seja então γ este semicírculo, intersectando o eixo real nos pontos a e b e suponha sem perda de generalidade que $a < b$. Afirmo que a seguinte transformação transforma γ numa reta vertical:

$$f_M(z) = \frac{\frac{z}{\sqrt{b-a}} - \frac{b}{\sqrt{b-a}}}{\frac{z}{\sqrt{b-a}} - \frac{a}{\sqrt{b-a}}}$$

Pelo lema 6.6, sabemos que $(f_M \circ \gamma)(t)$ é uma geodésica, isto é, ou é um trecho de reta vertical ou é um semicírculo centrado nalgum ponto do eixo real. Se t_0 é tal que $\gamma(t_0) = a$ (ou $\lim_{t \to t_0} \gamma(t) = a$), é fácil ver que $(f_M \circ \gamma)(t_0) = f_M(a) = \infty$ (caso se sinta desconfortável, transforme esta igualdade num limite). Como semicírculos não tomam valor no infinito (ou, como são limitado, não pode haver algum limite tendendo ao finito), só nos resta assumir que $(f_M \circ \gamma)(t)$ é o trecho de uma reta vertical. Podemos fazer o limite pelo outro lado dando zero. Por continuidade, só nos resta assumir que $(f_M \circ \gamma)(t)$ é uma reta vertical. Agora, recaímos no caso anterior e podemos realizar uma translação. Lembrando do fato de que Möb$_+(\mathbb{R})$ é um grupo, sabemos que a composição delas ainda está no conjunto, o que conclui nossa demonstração. ■

Agora iremos caracterizar definitivamente o grupo de isometrias do plano hiperbólico. Antes disso, veremos um lema que nos servirá de auxílio para encontrarmos o grupo de isometria. O lema diz que as únicas isometrias que deixam todos os pontos do eixo imaginário fixos são a identidade e a reflexão em torno deste eixo.

Lema 6.16. *Seja ϕ uma isometria de $\mathbb{H}$ tal que $\phi(iy) = iy$, $\forall y > 0$. Então $\phi(z) = z$ ou $\phi(z) = -\bar{z}$, $\forall z \in \mathbb{H}$.* □

Novamente, detalhes em [Bon09]. A ideia da demonstração consiste em notar primeiro que as geodésicas completas centradas na origem são fixadas pela isometria. Para ver isso, usamos que a imagem por ϕ de uma geodésica é uma geodésica (lema 6.6), tomamos um ponto P no eixo imaginário acima da geodésica e, usando o fato que ϕ é uma isometria e que $\phi(P) = P$, pela unicidade do lema 6.7, concluímos que a imagem da geodésica pela ϕ deve ser a própria geodésica.

Depois, dado um ponto z, tomamos a geodésica que contém z, centrada na origem e que atravessa o eixo y em Q. Devemos ter que $d(Q, z) = d(Q, \phi(z))$. É possível mostrar que isso só acontece se $\phi = z$ ou $= \bar{z}$. Isso pode ser feito usando fórmulas explícitas para a distância entre pontos ou podemos nos convencer que z e $\phi(z)$ devem ter a mesma ordenada. De qualquer forma, basta agora usarmos a continuidade para mostrar que $\phi(z)$ ou $\phi(z) = -\bar{z}$, $\forall z \in \mathbb{H}$.

Veremos agora que as funções ϕ e ψ definidas na proposição 6.14 são as únicas isometrias do plano hiperbólico

Teorema 6.17. *Seja Isom*$(\mathbb{H})$ *o grupo de Isometrias de* $\mathbb{H}$. *Então:*

$$Isom(\mathbb{H}) = \{f_M : M \in SL(2)\} \cup \{I \circ f_M : M \in SL(2)\}$$

□

A ideia da demonstração é tomar uma isometria que preserva orientação e mostrar que é possível compô-la com um transformação de Möbius f_M de tal forma que ela fixe a semirreta imaginária. Pelo lema 6.16 e usando a hipótese que a transformação preserva orientação, devemos ter necessariamente que esta composição é a identidade, e portanto que a isometria é a função inversa de f_M, donde segue que é uma transformação de Möbius.

A parte chata da demonstração é mostrar a existência da transformação de Möbius satisfazendo a propriedade citada acima. Na verdade, pelo lema 6.15, podemos achar uma transformação de Möbius que compondo com a isometria torna a semirreta imaginária invariante. Podemos ainda compor com uma homotetia para garantir que i fique invariante. Então, a parte realmente chata é mostrar que, dado que esta composição deixa i invariante, então também deixa todo outro ponto da semirreta imaginária invariante para usarmos o lema 6.16.

Vejamos uma expressão para o diferencial das isometrias.

Proposição 6.18. *Seja* $\phi : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ *dada por:*

$$\phi(z) = \frac{az + b}{cz + d}$$

Então:

$$d\phi_z(v) = \frac{ad - bc}{(cz + d)^2} v$$

. □

Proposição 6.19. *Isometrias preservam a métrica. Isto é, se* ϕ *é uma isometria, então* $g_{\phi(z)}(d\phi_z(v), d(\phi_z(w))$ *=* $g_z(v, w)$, *onde* v *e* w *estão em* $T_z\mathbb{H}$.

$\square$

Corolário 6.19.1. *Isometrias preservam a norma, isto é, se ϕ é uma isometria, então* $||d\phi_z(v)||_{\mathbb{H}} = ||v||_{\mathbb{H}}$ $\square$

Já vimos que o semiplano é homogêneo, isto é, dados dois pontos P, Q, existe uma isometria que leva um no outro. Agora, veremos que o semiplano é isotrópico, isto é, dados dois pontos p e q quaisquer e vetores quaisquer $v \in T_p\mathbb{H}$, $w \in T_q\mathbb{H}$, existe uma isometria que ao mesmo tempo leva p em q e cuja diferencial leva v em w.

Proposição 6.20. $\mathbb{H}$ *é isotrópico.* $\square$

6.4.4 Área

Em acordo com o que foi comentado na subseção 6.2.4, a área "hiperbólica"de um subconjunto $E \subset \mathbb{H}$ pode ser definida por:

$$\mu_{\mathbb{H}}(E) = \iint\limits_E \frac{1}{y^2} dxdy$$

E temos que:

Proposição 6.21. *Áreas são preservadas por isometrias. Isto é, se ϕ é uma isometria do semiplano, então* $\mu_{\mathbb{H}}(E) = \mu_{\mathbb{H}}(\phi(E))$. $\square$

Ao comparar com a área usual do plano, observamos características semelhantes àquelas comentadas quando comparamos as distâncias. Ponto com ordenada menor contribuem muito mais para a área.

6.4.5 Algumas expressões úteis

Aqui iremos expor algumas relações práticas para a distância hiperbólica.

Proposição 6.22. *Vale as seguintes expressões referentes à distância hiperbólica no modelo do semiplano superior:*

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

$$\tanh(\frac{1}{2}d(z,w)) = \frac{|z-w|}{|z-w^*|}$$

$$d(z,w) = \log\left(\frac{|z-w^*|+|z-w|}{|z-w^*|-|z-w|}\right)$$

□

6.5 O Disco de Poincaré

Começamos aqui o nosso segundo modelo do espaço hiperbólico. O leitor poderá notar claramente que todas as construções aqui serão muito mais fáceis e diretas. Isto porque basta acharmos uma isometria entre este modelo e o do semiplano e importar todos os resultados que já conhecemos de lá. Este será o argumento ubíquo nesta seção.

Nosso objeto desta vez será o disco unitário $\mathbb{B} := \{(x,y) \in \mathbb{R}^2; x^2+y^2 < 1\} \approx \{z \in \mathbb{C}; |z|^2 < 1\}$. A métrica será dada por:

$$g_z(v,w) = \frac{4}{(1-|z|^2)^2}\langle v,w\rangle$$

Apesar de ser uma expressão um pouco mais desengonçada, ela ainda é da forma $g_z(v,w) = f(z)\langle\cdot,\cdot\rangle$, com:

$$f(z) = \frac{4}{(1-|z|^2)^2}$$

O que já garante que a distância induzida por esta métrica de fato está bem definida, pelo lema 6.1. A norma ficará:

$$||v||_{\mathbb{B}} = \frac{2}{1-|z|^2}||v||$$

E o comprimento de uma curva γ:

$$l(\gamma) = \int \frac{2\sqrt{x'(t)^2+y'(t)^2}}{1-x(t)^2-y(t)^2}dt$$

O disco unitário munido da distância obtida tomando o ínfimo destes comprimentos é chamado de Disco de Poincaré.

Se formos comparar esta distância com a euclidiana, vemos que pontos com maior módulo, isto é, mais próximos (no sentido euclidiano) da fronteira do disco contribuem mais para a distância hiperbólica. Analogamente, esperamos que, se dois pontos estão muito próximos da fronteira, então a geodésica tende a se afastar da borda e depois voltar. Veremos que de fato isto ocorre e as geodésicos serão arcos de circunferência que incidem perpendicularmente na fronteira do disco de Poincaré. Isto não ocorre, por outro lado, com pontos que estão no eixo y, por exemplo. Pois, para qualquer que seja o lado que a curva penda, ficará mais próxima da borda. De modo geral, se dois pontos são colineares com o centro do disco então a geodésica é uma reta de fato.

O próximo lema nos dará uma transformação de Möbius que é isometria entre o semiplano e o disco de Poincaré:

Lema 6.23. *A função:*

$$\Phi(z) = \frac{-z+i}{z+1}$$

É uma isometria entre $\mathbb{H}$ e $\mathbb{B}$. □

Com isto também podemos mostrar que a distância induzida pela métrica também está bem definida. A inversa é:

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

Calcularemos a função para alguns pontos para ter uma ideia geral de como estas isometrias agem. É possível ver, por exemplo, que Φ leva o círculo no infinito na fronteira do disco e i na origem do disco. Além disso, achemos os pontos que vão nos quatro pontos cardeais do disco. Os pontos 0, 1, ∞ e -1 são mapeados em 1, i, -1 e $-i$ respectivamente. Vemos então que há como se fosse uma rotação de 90 graus entre o semiplano e o disco.

Da proposição 6.22, podemos tirar como resultado:

Proposição 6.24. *Se d é a métrica do disco, então vale que:*

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

□

6.5.1 Isometrias

Teorema 6.25. *O grupo de isometrias de $\mathbb{B}$ é o subgrupo de Möb$(\mathbb{C})$ tal que $d=\bar{a}$, $c=\bar{b}$ e $|a|^2-|b|^2=1$.* □

Este grupo é homomorfo ao grupo $SU(1,1)$, que é o grupo das matrizes

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}$$

Com $|\alpha|^2 - |\beta|^2 = 1$.

Uma classe importante de isometrias do disco de Poincaré são as rotações. Para uma rotação de ângulo θ, tome $\alpha = \exp(i\theta/2)$ e $\beta = 0$

6.5.2 Geodésicas

Proposição 6.26. *As geodésicas de $\mathbb{B}$ são os círculos ortogonais a $\partial\mathbb{B}$, incluindo as linhas retas.* $\square$

6.5.3 Área

Conforme foi discutido, podemos definir área no Disco de Poincaré como:

$$\mu_{\mathbb{B}}(E) = \iint \frac{4}{(1 - x^2 - y^2)^2} dxdy$$

É possível mostrar que esta definição é consistente com a área do semiplano. Isto é:

Proposição 6.27. *Seja $E \subset \mathbb{B}$. Então $\mu_{\mathbb{B}}(E) = \mu_{\mathbb{H}}(\Phi^{-1}(E))$.* $\square$

6.6 O hiperboloide

Para este modelo, o conjunto que consideraremos é uma esfera no espaço de Minkowski. Refresquemos a memória.

Definição 6.2. Espaço de Minkowski é um espaço vetorial real (ou espaço afim) de dimensão finita munido de uma forma bilinear simétrica não degenerada de índice 1. ♣

Onde:

Definição 6.3. Se B é uma forma bilinear, então o índice de B é o número da dimensão do subespaço de maior dimensão possível onde a forma é negativa definida, isto é, onde $v \neq 0 \implies B(v, v) < 0$, para todo v do subespaço. ♣

Espaços de Minkowski são sempre isométricos ao $\mathbb{R}^n$ com a forma bilinear $h(v, w) = v_1w_1 + ... + v_{n-1}w_{n-1} - v_nw_n$, onde $v = (v_1, ..., v_n)$ e $w = (w_1, ..., w_n)$. Esta forma será chamada neste texto de *produto de Minkowski*. Este produto não é um produto interno, caso exigirmos que produtos internos sejam positivos definidos. (Isto é, nem sempre vale que $v \neq 0 \implies h(v, v) > 0$). Tome, por exemplo $v = (0, ..., 1)$. Assim, ao falarmos espaço de Minkowski estaremos sempre nos referindo ao $\mathbb{R}^n$ munido do produto de Minkowski, por ser menos abstrato. Mais especificamente, estaremos interessados no $\mathbb{R}^3$.

Note que o conjunto $\{v \in \mathbb{R}^3$ tal que $h(v, v) = -1\}$ é uma esfera "de raio -1"com relação ao produto de Minkowski. Mas trata-se do conjunto $x^2 + y^2 - z^2 = -1$, que é um hiperboloide, com duas folhas (de revolução em torno do eixo z). O conjunto que estaremos interessados nesta seção é especificamente a folha superior do hiperboloide: $\{\mathcal{H} = (x, y, z); z > 0, x^2 + y^2 - z^2 = -1\}$, o que justifica o nome do modelo (modelo do hiperboloide).

Para as pessoas menos familiarizadas, vale saber que o espaço de Minkowski é o espaço físico segundo a teoria da relatividade restrita de Einstein. As pessoas mais familiarizadas sabem que este espaço não é uma variedade riemanniana pois, como vimos, $h(\cdot, \cdot)$ não é um produto interno. Entretanto, a restrição de h à $\mathcal{H}$ constitui-se sim uma métrica riemanniana.

Isso, entretanto, não será tão importante pois definiremos distância de um jeito diferente do que estávamos acostumados a definir. Definiremos:

$$d_{\mathcal{H}}(p, q) = \operatorname{arcosh}(-h(p, q))$$

Capítulo 7

Grupos Fuchsianos

7.1 Definição

Um grupo fuchsiano é um tipo particular de subgrupo de $PSL(2,\mathbb{R})$. A condição extra requerida é que a ação deste subgrupo no espaço hiperbólico não produza órbitas contínuas, em algum sentido. A maneira precisa de formular esta condição pode aparecer das mais diversas formas e esta seção irá se dedicar a sondar tais formas. Entretanto, podemos deixar claro logo de início que o fato importante, e que deve ser satisfeito por todas as definições, é que não haja pontos do espaço hiperbólico que sejam de acumulação para alguma órbita.

Mas qual é a motivação para tal definição? Por que a descontinuidade mencionada é importante? Posso citar dois motivos. O primeiro deles é que tal condição garante a regularidade do espaço quociente. Explico. Uma vez que tivermos um grupo fuchsiano Γ agindo num espaço X, cada órbita formará uma classe de equivalência. Por diversos motivos, iremos querer estudar o espaço formado pelas classes de equivalência, isto é, o espaço em que cada ponto corresponde a uma órbita inteira. Este espaço será denotado por X/Γ. A descontinuidade da ação garante, por exemplo, que o espaço quociente seja métrico, ou que seja uma variedade (mais precisamente, uma superfície de Riemann). Esta aplicação nos fará estudar espaços métricos quocientes, tesselações e, portanto, polígonos e regiões fundamentais, que serão temas das próximas seções. A segunda motivação vem do conceito de funções automórficas, mais detalhadas no apêndice histórico. Brevemente, uma função automórfica é uma função holomórfica que dá os mesmos valores em todos os pontos de uma órbita. Se essa órbita acumulasse, pela analiticidade da função e o fato dela ser constante numa órbita, ela teria que ser constante no espaço inteiro, o que, convenhamos, não é uma situação nada interessante.

Voltando para a definição de grupo fuchsiano, podemos dividir as definições que usualmente são dadas em duas categorias (equivalentes). A primeira, foca-se no grupo de maneira mais intrínseca, enquanto espaço topológico. Ela diz que grupo fuchsiano é um subgrupo discreto de $PSL(2,\mathbb{R})$. Esta maneira exigirá que reflitamos qual seria a topologia "natural"de $PSL(2,\mathbb{R})$. A segunda, mais focada na ação do grupo, propriamente dita, diz que grupo fuchsiano é um grupo que age descontinuamente em $\mathbb{H}$. Por sua vez, uma ação ser descontínua é alvo de inúmeras definições diferentes.

Antes de nos debruçarmos sobre as questões remanescentes sobre a definição de grupo fuchsiano, resta uma observação. Lembro que $PSL(2, \mathbb{R})$ é justamente o grupo de isometrias que preservam orientação do modelo hiperbólico do semiplano superior. Entretanto, como nossa motivação principal ao estudarmos grupos fuchsianos será estudar uma medida definida na fronteira do espaço hiperbólico, o modelo mais apropriado será o do disco de Poincaré, pois este trata todos os pontos da fronteira de maneira igualitária. Assim, usaremos a correspondência entre o semiplano e o disco para induzir uma correspondência entre $PSL(2, \mathbb{R})$ e $PSU(1, 1)$, de modo que, na prática, um grupo fuchsiano será para nós um subgrupo de $PSU(1, 1)$. Uma maneira alternativa de definir grupo fuchsiano de modo a contornar este problema seria de definir como um subgrupo do grupo de isometrias (de algum modelo) do espaço hiperbólico.

Começando pela mais fácil, um grupo fuchsiano é simplesmente um subgrupo discreto de $PSL(2, \mathbb{R})$, com a topologia sendo aquela discutida no capítulo 4, isto é, a topologia quociente de $SL(2, \mathbb{R})$. Como também foi visto, vale lembrar que esta topologia admite uma métrica.

Lema 7.1. *A ação de $PSL(2, \mathbb{R})$ em $\mathbb{H}$ é contínua.* □

7.1.1 Ações descontínuas

Como dito, um conceito importantíssimo no estudo de grupos fuchsianos é de uma ação de grupo ser descontínua. Este conceito, nada tem a ver com continuidade topológica. Na verdade, iremos nos preocupar somente com ações que no sentido topológico sejam contínuas, mas... que agem descontinuamente. Esta última expressão tem um sentido próprio que veremos a seguir. Peço desculpas de antemão ao leitor pelo fato das nossas ações serem ditas "ações contínuas que agem descontinuamente", mas infelizmente esta é a terminologia padrão, apesar de algumas tentativas de mudança.

Na verdade, há dois conceitos de descontinuidade. Diremos que a ação de um grupo é livremente descontínua quando satisfizer o conceito mais forte (que acabará não sendo tão interessante por ser muito restritivo). Já quando a ação satisfizer o conceito mais fraco, diremos que ela é propriamente descontínua. Este último conceito possui diversas e diversas definições parecidas. Elas diferem, muitas vezes, pelo objeto que tratam (algumas são para espaços métricos, outras para espaços topológicos, etc.) e outras muitas vezes são equivalentes ou equivalentes assumindo algumas hipóteses.

O conceito de descontinuidade livre é relativamente mais unânime e começaremos por ele.

Definição 7.1. Seja X um espaço topológico e G um grupo agindo em X. Dizemos que a ação é livremente descontínua em $x \in X$ quando existe uma vizinhança V de x tal que $gV \cap V \neq \emptyset \implies g = e$, onde e é o elemento identidade do grupo. ♣

Observo que esta definição é usada principalmente quando G é um grupo de homeomorfismos de X ou quando a ação é contínua no sentido topológico. Se considerarmos $\mathbb{R}^n$, a transformação $x \mapsto x + 1$ seria livremente descontínua em todo ponto. Mas cuidado! Se acrescentarmos o infinito, como na esfera de Riemann (vide seção 5.1.1), a ação não é livremente descontínua no infinito, pois é um ponto fixo da transformação ($\infty + 1 = \infty$). De fato, é evidente que uma ação nunca pode ser livremente descontínua num ponto fixo.

Antes de começarmos a investigar as várias definições de descontinuidade própria, vejamos um conceito que será muito importante - o de finitude local. Este mesmo conceito possui duas definições diferentes que são equivalentes satisfeitas algumas hipóteses.

Definição 7.2. Um subconjunto F de um espaço topológico X é dito localmente finito se, para todo $x \in X$, existir uma vizinhança V de x tal que $V \cap F$ é um conjunto finito. F é dito K-localmente finito se, para todo conjunto compacto $K \subset X$, $K \cap X$ é finito. ♣

Proposição 7.2. *Seja X um espaço topológico. Se $F \subset X$ é localmente finito, então é K-localmente finito. Se X for localmente compacto, vale a recíproca.* □

Demonstração. Suponha que F é localmente compacto e seja K um conjunto compacto. Para cada $x \in K$, tome V_x a vizinhança que satisfaz a definição de F ser localmente finito. Como V_x é uma vizinhança de x, existe um aberto U_x tal que $x \in U_x \subset V_x$ e é fácil ver que $U_x \cap F$ também é finito para todo x. Note que (U_x) será uma cobertura aberta de K, donde admite uma cobertura finita. Chamemos de (U_n). A intersecção de cada aberto da subcobertura com F é finita, e como há um número finito deles, vale que a interseção de F com a união de (U_n) é finita. Como tais conjuntos é uma subcobertura, vale que a intersecção de K com F é menor que a interseção com a subcobertura e portanto também finita.

Para a recíproca, suponhamos que X é localmente compacto. Dado x, vai existir K uma vizinhança compacta de x. Por ser F um conjunto K-localmente finito, a intersecção com K é finita e daí existe uma vizinhança tal que a intersecção com F é finita. Pela arbitrariedade de x, prova-se que F é localmente finito ■

Os espaços tratados neste texto serão sempre localmente compactos, então o leitor não precisará se preocupar com esta hipótese.

É hora de fazer um alerta bem importante. Quando estamos estudando ações de um grupo num espaço, frequentemente nos deparamos com elementos distintos g, h do grupo, mas que, entretanto, $gx = hx$ para algum x (lembre-se que uma ação ser livre é justamente dizer que isso não acontece). Apesar do ponto do espaço ser o mesmo, ao estudarmos a órbita de x, gx e hx contam como elementos diferentes. Este entendimento será crucial para o bem esclarecimento do que vem a ser o conceito de descontinuidade própria. Iremos, baseados nessa reflexão, introduzir o conceito de família localmente finita e K-localmente finita. De fato, ações de grupo livres estarão longes de ser as únicas tratadas. Os resultados da proposição anterior são naturalmente estendidos para este novo entendimento, e deixados como exercício para o leitor fanático.

Definição 7.3. Seja $(y_i)_{i \in I}$ uma família de elementos de um espaço topológico X indexada por um conjunto I arbitrário. Esta família é dita localmente finita se, para todo $x \in X$, existir uma vizinhança V de x tal que existe apenas um número finito de índices em I tais que $y_i \in V$. F é dito K-localmente finito se, para todo conjunto compacto $K \subset X$, o conjunto de índices tais que $y_i \in K$ é finito. ♣

Antes de vermos as definições, vamos estabelecer uma notação. $(A|B)_G := \{g \in G; g(A) \cap B \neq \emptyset\}$. As seguintes definições serão listadas por número e, quando quisermos nos referir a alguma em específica, diremos "ação propriamente descontínua do tipo n", onde n será o número.

Definição 7.4. Seja X um espaço topológico e G um grupo agindo em X. A ação é dita propriamente descontínua em X se:

1. $(K|K)_G$ for finito para todo K compacto;
2. existir uma vizinhança V de x tal que $(V|V)_G$ for finito, para todo x;
3. Gx for localmente finita, para todo x;
4. Gx for K-localmente finita, para todo x;
5. não existir um ponto $y \in X$ e uma sequência (g_n) de elementos distintos de G tais que $g_n y \to x$, para todo x ♣
6. Gx for discreta e G_x é finito, para todo x;

Um esclarecimento: na condição 5, acima, queremos dizer que x é um ponto limite da *sequência* $(g_n y)$, isto é, se há infinitos elementos distintos do grupo que fixam x, então x é um ponto limite desta sequência.

Note que na definição acima, descontinuidade é um conceito global. Entretanto, podemos definir descontinuidade pontualmente através das condições 2 e 5.

Se adicionarmos a hipótese das órbitas serem fechadas na condição 6, ela se torna equivalente a 5 sob fracas hipóteses:

Lema 7.3. *Seja G um grupo agindo em um espaço topológico X, $T1$. Então as seguintes afirmações são equivalentes:*

1. *Gx é discreta, fechada e G_x é finito, para todo x;*
2. *não existe um ponto $y \in X$ e uma sequência (g_n) de elementos distintos de G tais que $g_n y \to x$, para todo x.* □

Demonstração. Suponha, por absurdo, que exista um ponto $y \in X$ e uma sequência de elementos distintos (g_n) tais que $g_n y \to x$. Em primeiro lugar, note que devemos ter $g_n y = x$, para algum n. De fato, se não fosse o caso, x seria um ponto de acumulação para Gy no sentido de conjunto e teríamos que Gy não seria fechado, já que não teria um dos seus pontos de acumulação. Entretanto, como a órbita é discreta, deve haver alguma vizinhança U de x tal que $U \cap Gy = \{x\}$. Porém, isto implica que há infinitos elementos g_n tais que $g_n y = x$, donde o estabilizador não é finito e caímos numa contradição.

Reciprocamente, se o estabilizador não fosse finito, conseguiríamos obter infinitos elementos distintos do grupo com x sendo ponto de acumulação para a sequência $(g_n x)$. Além disso, suponha que alguma órbita não seja discreta. Então existe um elemento $x \in Gy$, para algum y, de maneira que, para toda vizinhança U de x, existe g tal que $gy \in U$ e $gy \neq x$. Assim, x é ponto de acumulação de Gy. Como estamos num espaço $T1$, isso significa que, em toda vizinhança de x, existem infinitos elementos de Gy. Com tais elementos, construiríamos uma sequência que acumula em x. Assim, toda órbita deve ser discreta e não pode possuir pontos de acumulação enquanto conjunto. Como um conjunto é fechado se, e somente se, contém todos seus pontos e acumulação, temos que as órbitas são fechadas. ■

Teorema 7.4. *Valem as seguintes implicações:*

a) $1 \implies 2$*, se* X *for localmente compacto*

b) $1 \implies 4$.

c) $2 \implies 5$ *se* X *for* $T1$

d) $3 \implies 4$*;*

e) $4 \implies 3$ *se* X *for localmente compacto;*

f) $4 \implies 5$ *se* X *for localmente compacto e* $T1$*;*

g) $5 \implies 4$

h) $5 \implies 6$

Demonstração. a) Dado x, como X é localmente compacto, existe uma vizinhança compacta K de x. Mas aí, sabemos que há apenas um número finito de elementos de G tais que $gK \cap K$ é não-vazia, donde existe uma vizinhança de x tal que $gV \cap V$ seja não-vazia apenas para um número finito de elementos, a saber, $V = K$.

b) Suponha por absurdo que haja K compacto e x tal que $gx \in K$ para infinitos elementos. Seja S o conjunto de todos os elementos de G tais que $gx \in K$. Por hipótese, S é infinito. Além disso, seja $y = hx \in K$ para algum $h \in S$. Então, Sh^{-1} é um conjunto de elementos de G que mandam $y \in K$ em K. em particular, para qualquer elemento $g \in Sh^{-1}$, temos $g(K) \cap K \neq \emptyset$. O resultado segue do fato que Sh^{-1} é infinito pois a multiplicação num grupo por um elemento é uma operação injetiva.

c) Dado x, como existe V vizinhança de x tal que $(V|V)_G$ é finito, segue de imediato que G_xé finito. Agora, quero mostrar que Gx é um conjunto discreto e fechado. Pela proposição 2.1, basta mostrar que Gx não tem pontos de acumulação. Suponha, por absurdo, que tenha e denotemos o ponto por y. Como o espaço é $T1$, existem infinitos elementos do grupo em qualquer vizinhança de y. Tome uma vizinhança U de y arbitrária. Tome também g_0 tal que $g_0x \in U$, que existe pelo que foi dito. Para qualquer outro ponto $x_i = g_ix \in U$, note que $g_0g_i^{-1}x_i = g_0x \in U$, para todo i, donde $g_0g_i^{-1}U \cap U \neq \emptyset$ para um número infinito de elementos, a saber $g_0g_i^{-1}$, o que contradiz a condição 2.

d) segue diretamente da proposição 7.2

e) segue diretamente da proposição 7.2

f) Dado x, note que $\{x\}$ é compacto (e vale em qualquer espaço topológico). Da hipótese, temos que há apenas um número finito de elementos do grupo tais que $gx \in \{x\}$, donde o estabilizador é finito. Resta mostrar que Gx é fechada e discreta, isto é, não possui pontos de acumulação. Suponha que haja tal ponto e denote-o por y. Como o espaço é localmente compacto, existe uma vizinhança compacta de y. Como y é ponto de acumulação é o espaço é $T1$, deve haver um número infinito de elementos da órbita em K, o que contradiz nossa hipótese.

g) Usaremos a equivalência dada pelo lema 7.3. Tome um conjunto compacto K qualquer e suponha que haja um elemento x e um número infinito de elementos de G tais que $gx \in K$. Como o estabilizador de cada ponto é finito, devemos ter infinitos pontos distintos em K. Pela compacidade, segue que Gx possui um ponto de acumulação. Como a órbita é fechada, este ponto de acumulação está em Gx, o que contraria a hipótese que Gx é discreto.

h) segue diretamente do lema 7.3

■

Resumirei os principais fatos sobre este teorema. Em primeiro lugar, saliento que todas as condições, menos a 2 e a 6 implicam em 4. Assumindo a hipótese fraquíssima que X é $T1$, ganhamos que 2 implica 4. Se, em adição, assumirmos que X é localmente compacto, então 3, 4 e 5 são equivalentes, e portanto 1 e 2 implicam nelas. Assim, quando formos provar propriedades dos grupos fuchsianos, tentaremos sempre provar assumindo a descontinuidade do tipo 3, 4 ou 5, e valerá que a propriedade demonstrada é satisfeita independentemente de qual definição for usada. Faremos isto na proposição 7.5, por exemplo.

Proposição 7.5. *Se G é um subgrupo de $PSL(2,\mathbb{R})$ que age descontinuamente (do tipo 5) em um subconjunto e $\mathbb{R}^n$, então Γ é enumerável.* □

Demonstração. Pela proposição 1.23, existe uma bijeção entre G/G_x e Gx. Recordo que todo subconjunto não-enumerável de $\mathbb{R}^n$ tem ponto de acumulação. Com efeito, tome um cubo fechado que contenha infinitos elementos e use a compacidade do cubo. Assim, como a ação é descontínua, Gx não tem pontos de acumulação e portanto é enumerável, donde G/G_x é enumerável. Ora, como o estabilizador de cada ponto é finito (lema 7.3), por definição de G/G_x, cada classe de equivalência é finita, e portanto G é uma união enumerável de conjuntos finitos, portanto é enumerável. ■

A seguir provaremos o teorema que faz a ponte entre as duas definições. Como uma trata da natureza intrínseca do grupo e a outra trata da ação dele no espaço, é óbvio que teremos que assumir uma condição para a ação. Esta condição, que propicia a ligação entre os dois conceitos é a continuidade da ação.

Teorema 7.6. *Um subgrupo Γ de $PSL(2,\mathbb{R})$ é discreto se, e somente se age descontinuamente.* □

Demonstração. Primeiro, suponha que um grupo aja descontinuamente mas não seja discreto. Então, deve existir uma sequência $\gamma_n \to I$, com I a identidade. Isto decorre da proposição 3.4, do fato da topologia de $PSL(2,\mathbb{R})$ ser $T1$, e do fato de subespaços de espaços $T1$ serem $T1$. Agora, devemos notar que $\gamma_n \to I \implies \gamma_n z \to z$. Tal afirmação pode ser verificada diretamente, ou podemos usar o lema 7.1. Neste caso, como a ação $\alpha : PSL(2,\mathbb{R}) \times \mathbb{H} \to \mathbb{H}$ é contínua, ela é sequencialmente contínua. Isto é, $(\gamma_n, z_n) \to (\gamma, z) \implies \alpha(\gamma_n, z_n) = \gamma_n z_n \to \gamma z = \alpha(\gamma, z)$. Em particular, $\gamma_n z \to \gamma z = Iz = z$, donde a órbita de nenhum ponto é discreta e portanto a ação não pode ser propriamente descontínua.

Reciprocamente, suponha que Γ seja discreto, mas que não aja descontinuamente. Então existem um pontos z, z_0 e uma sequência (γ_n) tais que $\gamma_n z \to z_0$. Então, devemos ter $\mathrm{Im}(\gamma_n z) \to \mathrm{Im}(z_0) > 0$. Isso nos dá, usando um escólio da proposição 6.13:

$$\frac{\mathrm{Im}z}{|c_n z + d_n|^2} \to \mathrm{Im}z_0 \implies \frac{1}{|c_n z + d_n|^2} \to \frac{\mathrm{Im}z_0}{\mathrm{Im}z} > 0$$

Daí, $(|c_n z + d_n|^2)$ é limitado, digamos, por M_1 e disto podemos concluir que (c_n) e (d_n) são limitados. Além disso, temos: $|\gamma_n z|^2 \to |z_0|^2$. Ou seja:

$$\frac{|a_n z + b_n|^2}{|c_n z + d_n|^2} \to |z_0|^2$$

Como o denominador é limitado, é fácil ver que $(|a_n z + b_n|^2)$ também deve ser limitado e, analogamente, (a_n) e (b_n) são limitados. Assim, existe uma subsequência de (a_n) convergente, digamos, para a. Tomando a subsequência de (b_n) correspondente, ela também vai ser limitada e, portanto, vai haver uma subsequência da subsequência que será convergente para b, digamos. Repetindo este procedimento para todos os coeficientes, obteremos uma subsequência convergente de (γ_n), para γ, digamos. Como o determinante é uma função contínua, temos que $ad - bc = 1$. e portanto $\gamma \in PSL(2,\mathbb{R})$. Usando, por fim, o lema 3.5, o fato de $PSL(2,\mathbb{R})$ ser $T1$ e Γ ser discreto por hipótese, temos que Γ é fechado, e portanto $\gamma \in \Gamma$, mas γ passa a ser um ponto de acumulação de Γ contido em Γ, o que contraria o fato de Γ ser discreto. ■

7.2 Conjunto Limite

Definição 7.5. Seja Γ um grupo fuchsiano agindo em $\mathbb{B}$. Dizemos que $y \in \overline{\mathbb{B}}$ é um ponto limite de Γ quando existe um ponto $x \in \mathbb{B}$ e uma sequência (γ_n) de elementos de Γ tais que $\lim d(y, \gamma_n x) = 0$, onde neste caso d é a restrição da métrica euclidiana para $\overline{B}$. O conjunto limite do ponto x, $\Lambda_x(\Gamma)$ é definido como o conjunto de todos os pontos limites de x. O conjunto limite $\Lambda(\Gamma)$ é definido como o conjunto dos pontos limites para um x qualquer de $\mathbb{B}$.

Como grupos fuchsianos agem descontinuamente, está claro que:

Proposição 7.7. $\Lambda(\Gamma) \subset \partial\mathbb{B}$ □

Proposição 7.8. *Seja Γ um grupo fuchsiano. Então $\Lambda_x(\Gamma) = \Lambda_y(\Gamma)$, quaisquer que sejam $x, y \in \mathbb{B}$. Segue que $\Lambda(\Gamma) = \Lambda_x(\Gamma)$, para todo $x \in \mathbb{B}$.* □

Proposição 7.9. *Seja Γ um grupo fuchsiano. $\Lambda(\Gamma)$ é um conjunto fechado e invariante pela ação de qualquer elemento de Γ* □

7.3 Espaço Métricos Quocientes

Começaremos com um espaço métrico e queremos definir um segundo espaço após quocientar o primeiro.

Um modo natural de definir a métrica no espaço quociente é a seguinte:

$$\inf\left\{\sum_{i=1}^{n} d(P_i, Q_i)\right\}$$

onde $[Q_{i-1}] = [P_i]$.

Este é o ínfimo da soma das distâncias de uma sequência finita de pontos, intercalados com saltos dentro de uma mesma classe de equivalência. É fácil ver que a distância entre dois pontos do espaço quociente sempre será menor que a distância entre dois pontos quaisquer representantes destas classes no espaço original.

Na verdade, esta função nem sempre sempre uma métrica. Pode ser o caso, por exemplo, de haver uma sequência de pontos convergindo para P tal que, em cada classe de equivalência deles, há um ponto próximo de Q, de tal forma que eles convirjam para Q também. Isto faz com que a "distância"entre P e Q no espaço quociente seja zero. Apesar disto, a função será sempre uma semi-métrica. Quando a função é uma métrica, diremos que o espaço quociente é próprio.

Há duas situações de interesse ao tratarmos de espaços métricos quocientes. Uma delas é quando queremos "colar"lados de um polígono e a segunda é quando o quociente é dado pela ação de um grupo de isometrias. Em vista disso, a próxima seção comenta um pouco sobre polígonos. Eventualmente as duas aplicações irão se cruzar.

No caso específico em que o quociente é dado pela ação de um grupo de isometria (ou seja, as classes de equivalência sendo as órbitas de cada elemento do grupo, cf 1.16.1), há uma definição de distância equivalente a esta que acabamos de definir:

$$d([P],[Q]) = \inf\{d(P,Q), P \in [P], Q \in [Q]\} = \inf\{d(P,Q), Q \in [Q]\}$$

E além disso, há uma garantia de que a função seja de fato uma métrica desde que a ação seja descontínua (do tipo 2, por exemplo). A igualdade acima é expressão do fato de estarmos tratando de um grupo.

7.4 Regiões Fundamentais

Uma região fundamental pode ser entendida como um conjunto que possui um, e exatamente um, representante de cada classe de equivalência dada pela órbita de uma ação.

Apesar da ideia ser simples, novamente há bastante divergência entre os autores quanto à definição formal devido a algumas tecnicalidades. Uma das principais reside na questão do conjunto conter ou não as fronteiras, ou uma parte delas. Caso não contiver (e portanto o conjunto for aberto), iremos exigir que $\gamma_1(R) \cap \gamma_2(R) = \emptyset$ e que a união *dos fechos*: $\bigcup \overline{\gamma R}$ cubra todo o espaço. Caso contiver (e portanto o conjunto for fechado), iremos requerer que os próprios conjuntos cubram o espaço inteiro e que *apenas os interiores* deles não se justaponham.

Outra tecnicalidade é quanto a questão do conjunto ser conexo ou não. Em caso afirmativo, chamaremos de domínio fundamental.

Por fim, muitos autores ainda exigem que uma região fundamental seja um polígono. Neste caso, a chamaremos de polígono fundamental. Na verdade, uma região fundamental será um polígono em praticamente todos os casos de interesse. Para tanto, convém falarmos mais sobre o que entendemos por polígono. Novamente, a questão de incluir ou não a fronteira persiste (na verdade, alguns lugares tratam um polígono até mesmo como sendo apenas a fronteira. Pelo fato de ser inadequada aos nosso propósitos, não a iremos considerar). Uma maneira simples seria definir como uma intersecção de um número finito de semiplanos. Entretanto, esta definição não é muito útil para provar os resultados da nossa teoria. No lugar disso, podemos dar como exemplo a seguinte definição, que utiliza a convenção de um polígono incluir a borda.

Definição 7.6. Um polígono é uma região $X \subset \mathbb{R}^2$ tal que:

a) X é fechado

b) $\partial X = \bigcup_{1 \leqslant i \leqslant n} E_i$

Sendo que cada E_i é ou um segmento de reta (geodésica, no caso hiperbólico), ou uma reta ou uma semirreta. E satisfaz as seguintes propriedades:

i. Cada E_i é fechado

ii. Só se intersectam nas extremidades

iii. Apenas dois elementos de (E_i) podem se intersectar num dado ponto.

Cada E_i é chamado de aresta e o ponto onde dois deles se encontram é chamado de vértice. ♣

Note que, para nós, um polígono pode ser ilimitado. Será útil lembrar que um conjunto é fechado se, e somente se, possui todos seus pontos de fronteira. Além disso, um conjunto é aberto se, e somente se, é disjunto de todos os seus pontos de fronteira.

Se precisar, dotaremos os polígonos de uma métrica definida como o ínfimo dos comprimentos de curvas de classe $\mathcal{C}^\infty$ por partes contidas no polígono, à semelhança do que fizemos em grande parte do capítulo anterior. Caso o polígono for conexo, é fácil ver que esta métrica coincidirá com a restrição da métrica euclidiana ao polígono.

Da posse da definição de um polígono, podemos dizer, simplesmente, que um polígono fundamental é aquele que ladrilha o plano conforme varremos o grupo de isometrias. Vemos que este conceito de ladrilhamento casa bem com a nossa proposta de regiões fundamentais. O nome técnico para isto é *tesselação*, da qual forneceremos um exemplo rigoroso de definição, novamente usando a convenção que nossos polígonos contêm a fronteira.

Definição 7.7. Uma tesselação num plano hiperbólico X é uma família (X_n) tal que:

1. Cada X_n é um polígono conexo
2. Cada par X_n, X_m é isométrico
3. $\bigcup X_n = X$
4. A intersecção de dois elementos são apenas arestas ou vértices, que o devem ser de ambos.
5. (X_n) é localmente finito.

♣

Em vez de exigir que a intersecção da imagem da região fundamental por duas isometrias diferentes seja vazia, podemos simplesmente exigir que não haja dois pontos da mesma órbita (ou como é comumente dito, $\Gamma-$equivalentes ou $\Gamma-$congruentes) na região fundamental:

Proposição 7.10. *Se não existem pontos $x, y \in R$, R região fundamental para Γ, tais que $y \in \Gamma x$, então não existem $\gamma_1, \gamma_2 \in \Gamma$ tais que $\gamma_1 R \cap \gamma_2 R \neq \{O\}$*

Demonstração. Provaremos a contra-positiva. Suponha que hajam isometrias γ_1, γ_2 tais que $\gamma_1 R \cap \gamma_2 R \neq \{O\}$. Então existe $p \in \gamma_1 R \cap \gamma_2 R$, donde existem pontos em R, digamos, x, y tais que $\gamma_1 x = p$ e $\gamma_2 y = p$. Entretanto, teremos $y = \gamma_2^{-1} p = \gamma_2^{-1} \gamma_x y$, donde $y \in \Gamma x$. ■

Temos o seguinte teorema:

Proposição 7.11. *Se um grupo de isometrias Γ admite uma região fundamental, então Γ age descontinuamente (do tipo 2).* □

Somos deixados então com o problema de verificarmos a recíproca da proposição acima. Isto é, uma vez que nos é dado um grupo de isometrias que age descontinuamente, existe uma região fundamental para ele? Em caso afirmativo, qual seria ela? Existe uma maneira de construir regiões fundamentais, através de um conceito conhecido como *domínio de Dirichlet*. Esta ferramenta irá nos fornecer a recíproca da proposição 7.11, desde que permitamos polígonos com um número infinito de lados (mas localmente finitos).

Definição 7.8. Seja Γ um grupo de isometrias que age descontinuamente no espaço X. O domínio de Dirichlet para Γ, centrado em P, é o conjunto:

$$\Delta_\Gamma(P) = \{Q \in X; d(Q, P) \leqslant d(Q, \gamma(P)), \gamma \in \Gamma\}$$

Teorema 7.12. *Seja Γ um grupo de isometrias que age descontinuamente no plano hiperbólico. Então, para todo ponto $P \in X$, o domínio de Dirichlet $\Delta_\Gamma(P)$ é um polígono localmente finito que forma uma tesselação do plano. Além disso, se P não é fixado por nenhum elemento de Γ com exceção da identidade, então $\Delta_\Gamma(P)$ é um polígono fundamental para Γ*

A prova deste teorema é extensa e pode ser encontrada em [Bon09].

Capítulo 8

Teoria de Patterson-Sullivan e Medidas Conforme

Neste capítulo iremos construir a medida de Patterson [Pat76] que serviu de inspiração para construções semelhantes em diversas áreas da matemática.

Começaremos com um breve resumo. Em primeiro lugar, definiremos uma série que deve divergir e, para tanto, provaremos um lema que garante que isso ocorra sempre, com uma ligeira modificação da série original. Usando esta série como peso para medidas de Dirac, construiremos uma família de medidas com o intuito de extrair uma subsequência convergente desta família, cujo limite será a medida que procuramos. Esta medida, estará concentrada na borda do disco. Por isso, devemos considerar agora todo o disco. É possível demonstrar que não tem como estender a distância hiperbólica para a fronteira. Assim, quando tivermos tratando do disco fechado, trataremos apenas como espaço topológico, e não métrico, com a topologia induzida por $\mathbb{R}^2$, e que empata com a topologia advinda da métrica hiperbólica dentro do disco.

A série, por sua vez, depende de uma função auxiliar. No artigo original do Patterson, esta função é definida como:

$$h(z, w) = \frac{|1 - \overline{z}w|^2}{(1 - |z|^2)(1 - |w|^2)}$$

Esta função pode ser relacionada com a distância no disco de Poincaré da seguinte forma:

$$h(z, w) = \cosh^2\left(\frac{1}{2}d(z, w)\right)$$

Que pode ser obtido mediante 6.24. Noto então que a função h é proporcional a distância (no sentido que a função cosseno hiperbólico é crescente), mas $h(z, z) = 1$.

Entretanto, diversos outros lugares [Bor07, DS12] [outros] usam, como função auxiliar:

$$g(z,w) = \exp(d(z,w))$$

As propriedades são semelhantes, exponencial é crescente e quando os pontos são iguais temos 1.

A função, em si, não é tão importante. O que realmente importa é o seu expoente de convergência:

Definição 8.1. O expoente de convergência de uma série $\sum a_n^{-s}$ é definido como $\delta := \inf\{s \geqslant 0; \sum a_n^{-s} < \infty\}$. Em especial, para um grupo fuchsiano Γ, o expoente de convergência da função g é definido como o expoente de convergência da série cujos termos são $g(z, \gamma w)$, $\gamma \in \Gamma$ (lembre-se de que são enumeráveis). Neste caso, indicaremos o expoente como $\delta(\Gamma)$. ♣

Por definição temos que, para $s < \delta$, a série diverge e para $s > \delta$ a série converge. Para $s = \delta$, a série pode convergir ou divergir. Entretanto, será importante que a série sempre divirja para $s = \delta$, o que nos motiva a enunciar o seguinte lema:

Lema 8.1. *Seja $\delta > 0$ o expoente de convergência de uma série $\sum a_n^{-s}$. Então existe uma função k positiva e crescente tal que a série*

$$\sum_n k(a_n) a_n^{-s}$$

continua tendo expoente de convergência δ e diverge para $s = \delta$. Além disso, para todo $\epsilon > 0$ vai existir y_0 tal que, se $y > y_0$ e $x > 1$, então $k(xy) \leqslant x^\epsilon k(y)$.

□

Demonstração. Como a série converge para algum valor de s, temos necessariamente que $\frac{1}{a_n^s} \to 0$, o que implica que $a_n \to \infty$. Isso garante que possamos reescrever a sequência como uma sequência crescente, suposição a qual faremos a partir de agora. Tome $\epsilon_n = 1/n$. Vamos definir uma sequência X_n tal que $X_n \to \infty$ e k será construída por indução nos intervalos $[X_n, X_{n+1}]$. Começaremos pondo $X_1 = 1$ e $k(x) = 1$ para $x \in [0,1]$.

Para o passo indutivo, suponha que k está definida em $[0, X_n]$. Escolheremos X_{n+1} de forma que:

$$\sum_{X_n < a_p \leqslant X_{n+1}} a_p^{-(\delta - \epsilon_n)} \geqslant \frac{X_n^{\epsilon_n}}{k(X_n)} \tag{8.1}$$

Tal X_{n+1} existe pois a série em questão diverge. De fato, como $\epsilon_n > 0$, $\delta - \epsilon_0 < \delta$, que é um expoente tal que a série diverge. Para $x \in [X_n, X_{n+1}]$, defina:

$$k(x) = k(X_n) \left(\frac{x}{X_n}\right)^{\epsilon_n}$$

Temos enfim nossa função construída, bastando verificar que ela satisfaz todas as propriedades requeridas. Em primeiro lugar, é evidente que se trata de uma função positiva. Também vale frisar que a função é contínua, uma vez que a função se emenda nos extremos de cada intervalo da definição, como é fácil ver. Usando o fato que ela se emenda e que a função é crescente em cada intervalo, que vem direta da definição, temos também que a função é crescente.

Verificaremos agora que para $s = \delta$ a série diverge. Note que:

$$\sum k(a_p)a_p^{-\delta} = \sum_n \sum_{X_n < a_p \leqslant X_{n+1}} k(a_p)a_p^{-\delta}$$

Por definição da função k:

$$\sum k(a_p)a_p^{-\delta} = \sum_n \sum_{X_n < a_p \leqslant X_{n+1}} k(X_n) \left(\frac{a_p}{X_n}\right)^{\epsilon_n} a_p^{-\delta}$$

$$= \sum_n \frac{k(X_n)}{X_n^{\epsilon_n}} \sum_{X_n < a_p \leqslant X_{n+1}} a_p^{-(\delta-\epsilon_n)}$$

Por (8.1), temos que finalmente:

$$\sum k(a_p)a_p^{-\delta} \geqslant \sum_n 1 = \infty$$

A fim de verificar a próxima propriedade, vamos estudar $\log k(x)$. Note que, para $x \in [X_n, X_{n+1}]$, $\log k(x) = \log a_n + \epsilon_n u$, onde $u = \log x$ e

$$a_n = \log \frac{k(X_n)}{X_n^{\epsilon_n}}$$

De modo que $\log k(x)$ é uma função contínua e afim por partes com relação a variável u.

Agora, dado $\epsilon > 0$, tome n tal que $\epsilon > \epsilon_n$. Como (ϵ_n) é decrescente, temos que o coeficiente angular de cada componente de $\log k(x)$, para $x \geqslant X_n$, é menor de ϵ_n. Assim, dados $p_1, p_2 > X_n$ e $y_1 = \log p_1$, $y_2 = \log p_2$, temos:

$$\frac{\log k(p_1) - \log k(p_2)}{y_1 - y_2} < \epsilon_n < \epsilon$$

Onde supomos, sem perda de generalidade, que $p_1 > p_2$. Redefinindo $y = p_2$ e $x = p_1/p_2$, temos:

$$\frac{\log k(xy) - \log k(y)}{\log(xy) - \log(y)} < \epsilon$$

Note que o denominador é $\log(xy) - \log(y) = \log(x)$, donde:

$$\log k(xy) - \log k(y) < \epsilon \log(x)$$

Exponenciando temos $k(xy)/k(y) < x^\epsilon$ e, por fim, $k(xy) < x^\epsilon k(y)$, com $y_0 = X_n$ e $x > 1$, que é o que queríamos.

Por fim, queremos provar que δ ainda continua sendo o expoente de convergência da nova série. Como já vimos que $s = \delta$ faz com que a série divirja, é óbvio que para $s < \delta$, a série continuará divergindo. Assim, basta provarmos que para $s > \delta$ a série converge.

De fato, se $s > \delta$, existe $\epsilon > 0$ tal que $\delta + \epsilon < s$. Dado este ϵ, pelo que acabamos de provar, existe um y_0 tal que, para $y > y_0$, $a_p > 1$, $k(a_p.y) < a_p^\epsilon k(y)$. Tomando $y > 1$, $a_p < a_p.y$ e, pelo fato de k ser crescente, temos $k(a_p) < k(a_p.y) < a_p^\epsilon k(y)$. Resumidamente, dado $\epsilon > 0$, existe $y > 1$ tal que $k(a_p) < a_p^\epsilon k(y)$, para todo $a_p > 1$.

Assim, as duas séries abaixo convergem e divergem juntas:

$$\sum k(a_p) a_p^{-s}$$

$$\sum a_p^\epsilon a_p^{-s} = \sum a_p^{-(s-\epsilon)}$$

Como $\delta + \epsilon < s$, $s - \epsilon > \delta$, donde a segunda série acima converge e, portanto, a primeira. Isto completa a prova do lema.

■

Dito isto, definimos a função:

$$M(s) = \sum_{\gamma \in \Gamma} k(g(\alpha, \gamma(\beta))) g(\alpha, \gamma(\beta))^{-s}$$

Que corresponde ao valor da série em função do expoente.

Agora já estamos em condições de definir uma família de medidas, indexada pelo expoente:

$$\mu_s = M(s)^{-1} \sum_{\gamma \in \Gamma} k(g(\alpha, \gamma(\beta))) g(\alpha, \gamma(\beta))^{-s} \delta_{\gamma(\beta)}$$

Onde δ_x é a medida de Dirac em x. Note que cada medida é de probabilidade.

Por fim, toma-se uma sequência dentro desta família de maneira que $s_j \to \delta$.

É um resultado bem conhecido que o espaço das medidas definidas num conjunto compacto (no caso, o fecho do disco de Poincaré) é compacto, de maneira que podemos extrair uma subsequência convergente. Os detalhes disto serão feitos na próxima seção.

A medida obtida como limite desta subsequência é a medida que procuramos. Resta, entretanto uma dúvida. Esta medida é única? Isto é, qualquer que seja a subsequência convergente, a medida limite será a mesma? A resposta é: não necessariamente, mas acontece para muitos grupos. Provaremos agora que esta medida está concentrada no conjunto limite:

Teorema 8.2. *Seja Γ um grupo fuchsiano e μ uma medida associada a Γ conforme construída acima. Então vale que $\mu(\overline{\mathbb{B}}\backslash\Lambda(\Gamma)) = 0$* □

Demonstração. Seja $\beta \in \mathbb{B}$ o ponto cuja órbita $\Gamma\beta$ dá origem a medida. Seja $\overline{\Gamma\beta} := \Gamma\beta \cup \Lambda(\Gamma)$. Provarei primeiro que $A := \overline{B}\backslash\overline{\Gamma\beta}$ tem medida nula. Note, primeiramente, que $\overline{\Gamma z}$ é fechado, pois é a união de um conjunto com seus pontos limites. Então A é aberto. Sabendo disto, podemos utilizar o lema 8.3, que nos diz que $\mu(A) \leqslant \liminf \mu_{s_n}(A)$. Note, entretanto, que $\mu_s(A) = 0$ para todo s, pois A não possui nenhum ponto da órbita $\Gamma\beta$, por definição. E concluímos facilmente que $\mu(A) = 0$.

Agora, como a órbita é discreta, para cada $\gamma\beta$, tome uma bola aberta de centro em $\gamma\beta$ e raio de forma a não se intersectar com nenhum outro ponto da órbita. Definimos B como a união de todas essas bolas. Trivialmente, B é aberto. Como $\Gamma\beta$ é enumerável (escólio de 7.5, por exemplo). Para provar que $\mu(B) = 0$, basta provar que $\mu(\Delta_\gamma) = 0$, onde Δ é uma tal bola arbitrária. Apelamos novamente para o lema 8.3. Temos $\mu(\Delta_\gamma) \leqslant \liminf \mu_{s_n}(\Delta_\gamma)$. Agora, direto da definição das medida, temos:

$$\mu_{s_n}(\Delta_\gamma) = \frac{1}{M(s)} \# G_{\gamma(\beta)} k(g(\alpha, \gamma(\beta))) g(\alpha, \gamma(\beta))^{-s_n}$$

Quando $s_n \to \delta$, o numerador tende a uma constante. Basta então provar que $M(s) \to \infty$. Finalmente chegou o momento de vermos a importância do lema 8.1. O fato da medida estar suportada no conjunto limite é completamente dependente do fato da série divergir.

Um modo de ver que temos de fato $M(s) \to \infty$ é utilizando a teoria das séries de Dirichlet da forma:

$$M(s) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$$

Onde $a_n = k(g(\alpha, \gamma_n(\beta)))$ (lembre que Γ é enumerável) e $\lambda_n = d(\alpha, \gamma(\beta))$. Seguirá que $M(s)$ é analítica (veja [Tit39]), portanto contínua.

Com isso, é fácil ver que $\overline{B}\backslash\Lambda(\Gamma) = A \cup B$ e daí $\mu(\overline{B}\backslash\Lambda(\Gamma)) \leqslant \mu(A) + \mu(B) = 0$. ■

8.1 Convergência fraca de medidas

Esta seção é dedicada a uma investigação um pouco mais cuidadosa sobre o assunto de convergência de medidas que foi usado para construir a medida na seção anterior. O primeiro ponto relevante é estabelecer qual é o sentido de convergência. O sentido é o de convergência fraca que, ingenuamente, podemos definir da seguinte forma:

Definição 8.2. Seja $(X, d, \mathcal{B})$ um espaço métrico munido da $\sigma-$álgebra de Borel. Uma sequência de medidas μ_s converge fracamente para μ quando:

$$\lim_{s\to\infty} \int f d\mu_s = \int f d\mu$$

Para toda função $f : X \to \mathbb{R}$ contínua. ♣

Temos o seguinte lema:

Lema 8.3. *Seja $(X, \mathcal{B})$ um espaço métrico munida da $\sigma-$álgebra de Borel. Seja também (μ_n) uma sequência de medidas. Se (μ_n) converge fracamente para μ, então: $\limsup \mu_i(C) \leqslant \mu(C)$ e $\liminf \mu_i(U) \geqslant \mu(U)$, para todo fechado C e todo aberto U*

Um resultado importante é que é possível definir uma métrica no espaço das medidas de probabilidades de forma consistente com esta ideia de convergência. Antes, precisamos de um lema.

Lema 8.4. *Seja K compacto. Então, $C(K)$ é separável.* □

Teorema 8.5. *Seja $(K, \mathcal{B})$ um espaço métrico compacto munido da $\sigma-$álgebra de Borel e seja $M_1(K)$ o conjunto de todas as medidas de probabilidade definidas em $(K, \mathcal{B})$. Como K é compacto, $C(K)$ é separável e portanto há um conjunto de funções $\{f_n\}_n$ densas em $C(K)$. Tomando tal conjunto, a função $d : M_1(K) \times M_1(K) \to [0, +\infty)$ definida por:*

$$d(\mu, \nu) = \sum_{n=1}^{+\infty} \frac{1}{2^n ||f_n||_\infty} \left| \int f_n d\mu - \int f_n d\nu \right|$$

É uma métrica em $M_1(K)$ e μ_n converge fracamente para μ se, e somente se, $\lim d(\mu_n, \mu) = 0$
□

A topologia induzida em $M_1(K)$ pela métrica acima é chamada de topologia fraca-*. Quando um conjunto é compacto em relação a esta topologia, dizemos que ele é fracamente compacto.

Antes de partirmos para os resultados principais, devemos enunciar um teorema bastante importante e famoso:

Teorema 8.6. *(Teorema de Riesz-Markov) Seja X um espaço Hausdorff localmente compacto e seja L um funcional linear positivo no espaço das funções contínuas de suporte compacto. Então existe e é única uma medida de Borel regular μ em X tal que:*

$$L(f) = \int f d\mu$$

Para toda função f do domínio de L $\square$

Lembro que um funcional é positivo quando $f \geqslant 0 \implies L(f) \geqslant 0$. Além disso:

Definição 8.3. Seja $(X, \mathbb{A})$ um espaço mensurável. Uma medida μ em X é dita regular se:

1. $\mu(K) <= \infty$ para todo compacto $K \subset X$
2. Para todo $A \in \mathbb{A}$: $\mu(A) = \inf\{\mu(U) : A \subset U, U \text{ aberto}\}$
3. Para todo aberto $U \subset X$: $\mu(U) = \sup\{\mu(K) : K \subset U, K \text{ compacto}\}$ ♣

A prova para o teorema 8.6 pode ser encontrada em [Nor16]. E por fim:

Teorema 8.7. *Se K é compacto, então $M_1(K)$ é fracamente compacto.* $\square$

Demonstração. Provaremos que $M_1(K)$ é sequencialmente compacto, isto é, que toda sequência de medidas de probabilidade tem alguma subsequência que converge fracamente. Usaremos a notação $\mu(f) := \int f d\mu$ por comodidade (e porque medidas se relacionam com funcionais lineares de $C(X)$).

Seja $\{f_n\}$ um subconjunto denso de $C(K)$. Considere a sequência $\mu_n(f_1)$. Como as medidas são de probabilidade, temos que $|\mu_n(f_1)| \leqslant ||f_1||_\infty$, e portanto os números $\mu_n(f_1)$ são limitados. Bolzano-Weierstrass nos diz então que há uma subsequência $\mu_{1,n}(f)$ convergente. Agora, aplicamos esta sequência para a próxima função: $\mu_{1,n}(f_2)$ que, pelos mesmos motivos possui uma subsequência convergente $\mu_{2,n}$.

Assim, para cada i, obtemos uma sequência $\mu_{i,n}$ tal que $\mu_{i,n}(f_j)$ é convergente para $j \leqslant i$ e tal que $\{\mu_{i,n}\}_n \subset \{\mu_{i-1,n}\}_n$. Agora, tomaremos a sequência de medidas $\mu_{n,n}$. Pelo que acabamos de dizer, $(\mu_{n,n})$ é subsequência de $\mu_{i,n}$ para $n \geqslant i$, e daí $\mu_{n,n}(f_i)$ converge para todo i.

Usaremos a densidade de $\{f_n\}$ para provar que $\mu_{n,n}(f)$ converge para toda $f \in C(K)$. Antes, note que $||f - f_i||_{+\infty} < \epsilon \implies |\nu(f) - \nu(f_i)| < \epsilon$, para toda $\nu \in M_1(K)$. Com efeito:

$$|\nu(f) - \nu(f_i)| = |\nu(f - f_i)| \leqslant \nu(|f - f_i|) \leqslant \nu(||f - f_i||_{+\infty}) \leqslant \nu(\epsilon) = \epsilon\nu(K) = \epsilon$$

Dado um $\epsilon > 0$ arbitrário, tome f_i tal que $||f - f_i||_\infty < \epsilon$, que existe pela densidade das funções (f_n). Como $(\mu_{n,n}(f_i))$ converge, existe N tal que $n, m > N \implies |\mu_{n,n}(f_i) - \mu_{m,m}(f_i)| < \epsilon$. Tomando tal N temos que, para $n, m > N$:

$$|\mu_{n,n}(f) - \mu_{m,m}(f)| \leqslant |\mu_{n,n}(f) - \mu_{n,n}(f_i)| + |\mu_{n,n}(f_i) - \mu_{m,m}(f_i)| + |\mu_{m,m}(f_i) - \mu_{m,m}(f)|$$

O primeiro e o último termo são menores que ϵ pela desigualdade que acabamos de mostrar, enquanto o do meio é menor que ϵ pela convergência de $(\mu_{n,n}(f_i))$. Assim, mostramos que, para todo $\epsilon > 0$, existe N tal que $n, m > N \implies |\mu_{n,n}(f) - \mu_{m,m}(f)| < 3\epsilon$, donde $(\mu_{n,n}(f))$ converge, com $f \in C(K)$ arbitrária.

Com esta convergência em mãos, definimos $\mu(f) := \lim \mu_{n,n}(f)$. Usando o teorema da representação de Riesz, provaremos que $\mu(f)$ de fato se trata de uma medida. Assim, uma vez que checarmos as hipóteses do teorema teremos encontrado uma subsequência $(\mu_{n,n})$ de μ_n que converge fracamente para uma medida μ, isto é, $\lim \mu_{n,n}(f) = \mu(f)$, para toda $f \in C(K)$.

Em primeiro lugar, como estamos em um espaço métrico compacto, então ele é localmente compacto e Hausdorff. Além do mais, $\mu(\lambda f + g) = \lim \mu_{n,n}(\lambda f + g) = \lim(\lambda \mu_{n,n}(f) + \mu_{n,n}(g)) = \lambda \lim \mu_{n,n}(f) + \lim \mu_{n,n}(g) = \lambda\mu(f) + \mu(g)$, donde μ é um funcional linear. Mais que isso, se $f \geqslant 0$, então $\mu_{n,n}(f) \geqslant 0$, para todo n, donde $\mu(f) \geqslant 0$ e portanto μ é um funcional linear positivo. Por fim, $\mu(1) = \lim \mu_{n,n}(1) = \lim 1 = 1$, donde μ é de fato uma medida de probabilidade.

■

8.2 Série de Poincaré

A série de Poincaré surgiu, originalmente, com o fim de construir funções automórfica. Queremos uma função que seja invariante sob certo grupo. Para tanto, usaremos o fato que multiplicar todos os elementos de um grupo por um elemento recupera todos os elementos originais. Assim, nossa função será um somatório do tipo $F(z) = \sum_{\gamma \in \Gamma} H(\gamma z)$, de forma que $F(\gamma_i z) = \sum_{\gamma \in \Gamma} H(\gamma_i \gamma z)$. A segunda série é apenas um rearranjo da série original, de forma que $F(\gamma_i z) = F(z)$.

Apêndice A

Resumo Histórico

Muitos conceitos apresentados neste texto surgiram a partir do estudo das funções elíticas e modulares. Portanto, daremos uma breve introdução histórica acerca destas.

Integrais elípticas são uma classe de funções obtidas a partir de certas integral. Estas funções possuem este nome pois, historicamente, algumas delas eram relacionadas ao cálculo do comprimento de arco de uma elipse.

Há dois conjuntos importantes de integrais para os quais qualquer integral elíptica pode ser reduzida. O primeiro, mais clássico, é conhecido como formas de Legendre. São três tipos:

$$F(\phi, k) = \int_0^{\phi} \frac{d\theta}{\sqrt{1 - k^2 \sin^2(\theta)}}$$

$$E(\phi, k) = \int_0^{\phi} \sqrt{1 - k^2 \sin^2(\theta)}\,d\theta$$

$$\Pi(\phi, n, k) = \int_0^{\phi} \frac{d\theta}{(1 - n^2 \sin^2(\theta))\sqrt{1 - k^2 \sin^2(\theta)}}$$

Conhecidas como integrais elípticas incompletas de primeiro, segundo e terceiro tipo. As respectivas integrais elípticas completas apenas diferem por substituírem ϕ por $\pi/2$. O segundo conjunto é conhecido como forma simétrica de Carlson, e são:

$$R_F(x, y, z) = \frac{1}{2}\int_0^{\infty} \frac{dt}{\sqrt{(t + x)(t + y)(t + z)}}$$

$$R_J(x, y, z, p) = \frac{1}{2}\int_0^{\infty} \frac{dt}{(t + p)\sqrt{(t + x)(t + y)(t + z)}}$$

O último conjunto de formas é mais recomendado para se calcular numericamente seus valores.

Funções elípticas são, historicamente, associadas com funções inversas das Integrais elípticas. Estas também possuem duas formas canônicas: de Jacobi e de Weierstrass. A definição moderna de função elíptica é uma função que é duplamente periódica nos complexas.

A forma de Weierstrass será especialmente importante para nós e é dada por:

$$\wp(z|\omega_1, \omega_2) = \frac{1}{z^2} + \sum_{m\omega_1+n\omega_2\neq 0} \left(\frac{1}{(z + m\omega_1 + n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right)$$

.

Onde ω_1, ω_2 são complexos e $\omega_1/\omega_2 \notin \mathbb{R}$. Esta função é usualmente estudado quanto a sua dependência em relação a z, chamada de teoria das funções elípticas. O estudo quanto a dependência em ω_1 e ω_2 é chamado de teoria das funções elípticas modulares.

A função $\wp$ é duplamente periódica caso convirja (lembre-se de que a função está definida no plano complexo e aí o período pode ser dados por números complexos "linearmente independentes"). O conjunto $\Omega = \{m\omega_1 + n\omega_2; m, n \in \mathbb{Z}\}$ é chamado de reticulado. Estaremos interessados em achar outros dois números complexos α_1, α_2 tais que seu reticulado coincida com Ω. É possível mostrar que isto vale se, e somente se:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

com $ad - bc = \pm 1$. Considerando $\alpha_1/\alpha_2 > 0$ e $\omega_1/\omega_2 > 0$, temos que $ad - bc = 1$. Denotando uma tal matriz por V temos:

$$\tau' = \frac{\alpha_1}{\alpha_2} = \frac{a\omega_1 + b\omega_2}{c\omega_1 + d\omega_2} = \frac{a\tau + b}{c\tau + d} = V\tau$$

O conjunto das transformações $\frac{a\tau+b}{c\tau+d}$ forma um grupo, desde que identifiquemos M e $-M$ (formando classes de equivalência), cujo nome é grupo modular e denotado por $PSL(2, \mathbb{Z})$.

Estudaremos agora algumas funções relacionadas à função elíptica de Weierstrass e que são invariantes quanto a mudança de base, só dependem do reticulado em si. Para tanto, notaremos o fato que:

$$\wp'^2(z) = 4\wp^3(z) - g_2\wp(z) - g_3$$

Onde:

$$g_2(\omega_1, \omega_2) = 60 \sum (m\omega_1 + n\omega_2)^{-4}$$

$$g_2(\omega_1, \omega_2) = 140 \sum (m\omega_1 + n\omega_2)^{-6}$$

São funções conhecidas como invariantes da função elíptica de Weierstrass e é fácil ver que são invariantes quanto a base do reticulado. Estas funções também se relacionam com a série de Laurent em torno da origem da seguinte forma:

$$\wp(z) = z^{-2} + \frac{1}{20} g_2 z^2 + \frac{1}{28} g_3 z^4 + O(z^6)$$

Definindo $g_2(\tau) := g_2(\tau, 1)$, temos $g_2(\tau) = \omega_2^4 g_2(\omega_1, \omega_2)$, donde:

$$g_2(\tau') = (c\tau + d)^4 g^2(\tau)$$

Uma função que satisfaz a igualdade:

$$f(Vz) = (cz + d)^k f(z)$$

É chamada de forma modular de dimensão k (uma condição imposta sobre o crescimento da função pode ser requerido). Assim, vimos que g_2 é uma forma modular de dimensão -4. É possível ver, de forma análoga, que g_3 é uma forma modular de dimensão -6. As séries de g_2 e g_3 (sem considerar os coeficientes 60 e 140) são casos particulares das chamadas séries de Eisenstein.

Procuraremos agora formar uma forma modular de dimensão 0 (ou simplesmente forma modular). É fácil ver que a divisão de formas de mesma dimensão dá uma forma modular, por exemplo, g_2^3/g_3^2. Esta função, entretanto, tem a desvantagem de dar infinito, já que $g_3(z)$ é zero para $z = e^{2\pi i/3}$. Uma saída é considerar a função $\Delta_1(\tau) = g_2^3(\tau) - 27g_3^2(\tau)$, conhecida como discriminante modular, que é uma forma modular de dimensão -12 e nunca se anula e definir:

$$J(z) = \frac{g_2^3(\tau)}{\Delta_1(\tau)}$$

Esta função é chamada de função $j-$invariante de Klein ou invariante absoluta de Klein, que é modular, isto é: $J(V\tau) = \tau, \forall V \in PSL(2, \mathbb{Z})$. Esta função é muito importante para o estudo das funções modulares e vale que toda função modular é uma função racional de J.

É possível provar que um domínio fundamental para o grupo modular é o conjunto $\{\tau = (x, y) \in \mathbb{C}; -1/2 \leqslant x \leqslant 1/2, |\tau| \geqslant 1\}$.

Tanto a função J como o nome "função elíptica modular" aparecem pela primeira vez num paper de Dedekind em 1877 e um ano depois num paper de Klein. Eles mostraram que a função J assume qualquer valor do plano complexo num domínio fundamental. Assim nasce o estudo das funções automórficas, precisamente aquelas invariantes sob o grupo modular $PSL(2, \mathbb{Z})$.

Podemos dizer que o estudo das funções automórficas mais gerais, por outro lado, surgiu a partir do estudo da equação diferencial hipergeométrica por H. A. Schwarz, orientador de Lazarus Fuchs, que inclusive continuou o trabalho de Schwarz e inspirou Poincaré.

Bibliografia

[Bea95] Beardon, A. F.: The Geometry of Discrete Groups. (Springer-Verlag, New York, 1995)

[Bon09] Bonahon, F.: Low-Dimensional Geometry: *from Euclidean Surfaces to Hyperbolic Knots.* (AMS, Providence, 2009)

[Bor07] Borthwick, D.: Spectral Theory of Infinite-Area Hyperbolic Surfaces. Birkhauser, 2007.

[DS12] Denker, M. and Stratmann, O.: The Patterson Measure: *Classics, Variations and Applications..*

[JS87] Jones, G, A. and Singerman, D.: Complex Functions: *an algebraic and geometric viewpoint.* (Cambridge University Press, 1987)

[Kat92] Katok, S.: Fuchsian Groups. (University of Chicago Press, Chicago, 1992)

[Mar07] Marden, A.: Outer Circles: *An Introduction to Hyperbolic 3-Manifolds* (Cambridge University Press, Cambridge, 2007)

[Nic89] Nicholls, P.: The Ergodic Theory of Discrete Groups. (Cambridge University Press, Cambridge, 1989).

[Nor16] Norqvist, J.: The Riesz Representation Theorem for Positive Linear Functionals. Disponível em: http://umu.diva-portal.org/smash/get/diva2:953904/FULLTEXT01

[Pat76] Patterson, S. J.: The limit set of a Fuchsian group. Acta Math. **136** (1976), 241 - 273.

[Rat19] Ratcliffe, J.: Foundations of Hyperbolic Manifolds (Springer, New York, 2019)

[Sul79] Sullivan, D.: The density at infinity of a discrete group of Hyperbolic Motions, Publ. Math. IHES **50** (1979), 171-202.

[Tit39] Titchmarsh, E.: The Theory of Functions (Oxford University Press, Oxford, 1939)