

VITOR MONTEIRO PEREIRA

ENSAIO SOBRE A DESCRIÇÃO CATEGORIAL DE ANYONS

Dissertação apresentada à Universidade Federal de Viçosa como parte das exigências do Programa de Pós-Graduação em Física para a obtenção do título de *Magister Scientiae*.

Orientador: Jakson Miranda Fonseca

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

P436e
2021

Pereira, Vitor Monteiro, 1997-
Ensaio sobre a descrição categorial de anyons / Vitor
Monteiro Pereira. – Viçosa, MG, 2021.
278 f. : il. ; 29 cm.

Inclui apêndices.

Orientador: Jakson Miranda Fonseca.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 218-221.

1. Anyons. 2. Categorias (Matemática). 3. Partículas
idênticas. 4. Fibonacci, Anyons de . I. Universidade Federal de
Viçosa. Departamento de Física. Programa de Pós-Graduação
em Física. II. Título.

CDD 22. ed. 530.12

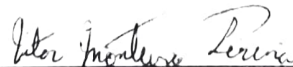
VITOR MONTEIRO PEREIRA

ENSAIO SOBRE A DESCRIÇÃO CATEGORIAL DE ANYONS

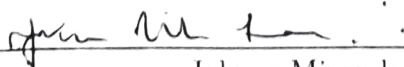
Dissertação apresentada à Universidade Federal de Viçosa como parte das exigências do Programa de Pós-Graduação em Física para a obtenção do título de *Magister Scientiae*.

APROVADA: 12 de abril de 2021.

Assentimento:



Vitor Monteiro Pereira
Autor



Jakson Miranda Fouseca
Orientador

Agradecimentos

Aos meus familiares, em particular meus pais pelo apoio.

Aos meus professores, em particular meu orientador, pelos ensinamentos.

A todes mis amigues, essenciais durante toda a minha formação e em particular, essenciais para manutenção da sanidade durante o mestrado. E à livraria do campus UFV onde nos encontrávamos todos os dias, quase religiosamente, após o almoço para um café, cappuccino ou chá e conversar.

À Universidade Federal de Viçosa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“Girando em torno deste sol, a uma distância de cerca de 148 milhões de quilômetros, há um planetinha verde-azulado absolutamente insignificante, cujas formas de vida, descendentes de primatas, são tão extraordinariamente primitivas que ainda acham que relógios digitais são uma grande ideia.”

- Douglas Adams

Resumo

PEREIRA, Vitor Monteiro, M.Sc., Universidade Federal de Viçosa, abril de 2021. **Ensaio sobre a descrição categorial de anyons**. Orientador: Jakson Miranda Fonseca.

Anyons são partículas que vivem em duas dimensões espaciais e possuem comportamentos estatísticos intermediários entre férmions e bósons. Neste trabalho fazemos uma exposição de como anyons podem ser adequadamente descritos usando mecânica quântica básica a partir de um processo de quantização adequado. Este processo revela a conexão entre as propriedades estatísticas de anyons e as propriedades topológicas do espaço de configurações do sistema. Em seguida, fazemos uma introdução elementar a teoria de categorias e a usamos para construir uma descrição categorial de anyons. A descrição categorial possui utilidades para aplicação de anyons em computação quântica e, nesta linha, resolvemos um exemplo, o modelo de anyons de Fibonacci para ilustrar como esta descrição funciona e apontar suas aplicação em computação quântica com estas partículas. Por fim, seguindo a natureza ensaística deste trabalho, concluímos por levantar vários tópicos de pesquisa sobre a assunto para os quais este trabalho pode servir como um ponto de partida.

Palavras-chave: Anyons. Categorias. Partículas idênticas. Anyons de Fibonacci.

Abstract

PEREIRA, Vitor Monteiro, M.Sc., Universidade Federal de Viçosa, April, 2021. **Essay on the categorical description of anyons.** Adviser: Jakson Miranda Fonseca.

Anyons are particles that exist in two spacial dimensions and present intermediate statistical behavior between that of fermions and that of bosons. In this work we present an exposition of how anyons may be properly described using basic quantum mechanics based on a proper quantization procedure. This procedure also reveals a connection between the statistical behavior and the topological properties of the system's configuration space. Following that, we present an elementary introduction to category theory and use it to construct a categorical description of anyons. This categorical description is useful in applications of anyons to quantum computation, and we solve an example, that of Fibonacci anyons to show how it works and to point the applications of this description in quantum computation with these particles. Finally, as this is an essay, we conclude by mentioning various related research topics for which this work may serve as a stating point.

Keywords: Anyons. Categories. Identical particles. Fibonacci anyons.

Lista de Figuras

2.1	Os dois tipos de trajetórias possíveis para as duas partículas na colisão.	20
2.2	Recipiente contendo um fluido separado em dois compartimentos.	43
2.3	Rotação adiabática de uma partícula em torno de outra.	47
2.4	Diagrama ilustrando possíveis processos de permutação de duas partículas.	48
2.5	Identificação dos pontos no semiplano inferior com pontos do semiplano superior.	51
2.6	Processo de identificação do subconjunto U	52
2.7	Dois loops distintos em \mathcal{C}_2/Z_2	52
2.8	Loops no hemisfério superior de uma esfera com pontos antipodais identificados.	56
2.9	Representação esquemática do fibrado de linhas e de uma seção.	61
2.10	Ação do gerador σ_i do grupo de tranças à esquerda e seus inversos, à direita.	71
2.11	Representação visual da relação de Yang-Baxter.	72
2.12	Comparação entre a aplicação de $\sigma_1^2 \neq 1$ e $\tau_1^2 = 1$	73
2.13	Trançamento de três partículas segundo $\tau = \sigma_2\sigma_1\sigma_2$	78
3.1	Exemplo de um diagrama de flechas.	82
3.2	Relação entre tipos de flechas em uma categoria. A flecha com o gancho indica inclusão.	106
3.3	Relação completa entre os tipos de flechas em uma categoria.	107
3.4	Grafo direcionado subjacente à categoria de três objetos.	127
4.1	Representação visual da torção $\theta_{A \otimes B}$	156
A.1	Diferentes topologias para um conjunto de três elementos.	224
A.2	Nenhuma destas coleções é uma topologia para este conjunto.	224
A.3	$B_3 \subset B_1 \cap B_2$	226
A.4	Bases \mathcal{B} e \mathcal{B}'	227
A.5	Exemplos de conjuntos abertos em $\mathbb{R} \times \mathbb{R}$ na topologia da ordem lexicográfica.	233
A.6	Conjunto aberto na topologia produto.	234
A.7	Conjunto aberto na topologia produto determinado a partir das funções projeção.	236
A.8	Comparação das topologias do subespaço e da ordem em um subespaço.	239
A.9	Digramas ilustrando A e sua relação com X e Y	242
A.10	Diagrama ilustrando a relação entre os abertos de X e Y sob o homeomorfismo f	247

A.11	Processo de construção de um toro a partir de um retângulo.	252
A.12	Processo de formação de uma casca esférica a partir de um disco.	252
A.13	Representação da topologia quociente de A	253
A.14	Ilustração dos conjuntos que compõe $[a, b]$	258
B.1	Diagrama representando uma homotopia de caminhos.	263
B.2	Processo de construção de uma homotopia de caminho a partir duas outras, F e G	264
B.3	Diagrama mostrando a construção da homotopia H entre $f * g$ e $f' * g'$	265
B.4	Mapeamento de uma homotopia de caminhos de X em uma de caminhos de Y	268
B.5	Caminhos em I mapeados em um único caminho em X por f	269
B.6	Caminhos em I mapeados no caminho constante, e_{x_0} e $f * \bar{f}$	270
B.7	Caminho $p : I \rightarrow I$	271
B.8	Loop gerado pelo mapa $\hat{\alpha}$	273

Sumário

1	Introdução	11
2	Sobre anyons	17
2.1	Estatística quântica	17
2.1.1	Permutações	25
2.1.2	Postulado de simetricidade	34
2.2	Crítica da indistinguibilidade	38
2.2.1	Espaço de configurações	44
2.3	Anyons	59
2.3.1	Quantização	59
2.3.2	Formulação de Feynman	66
2.3.3	Representações e o grupo de tranças	71
3	Sobre categorias	82
3.1	Categorias	89
3.1.1	Flechas especiais	99
3.1.2	Objetos especiais	107
3.1.3	Princípio de dualidade	110
3.2	Funtores	113
3.3	Transformações naturais	120
3.4	Universalidade e limites	126
3.4.1	Produto e coproduto	133
3.4.2	Equalizador e coequalizador	138
4	Sobre anyons e categorias	144
4.1	Monoidalidade	149
4.2	Rigidez e extensão	154
4.3	Semissimplicidade	158
4.3.1	Categorias aditivas	159
4.3.2	Categorias abelianas	169
4.3.3	Categorias semissimples	176
4.4	Modularidade	178
4.5	Simplificações	181
4.6	Exemplo: Anyons de Fibonacci	185
4.6.1	Associatividade	191
4.6.2	Trançamentos	203
4.7	Computação quântica	208

5	Conclusão	211
5.1	Síntese	212
5.2	Sobre escopo	213
5.3	Considerações finais	216
	Referências	217
A	Sobre topologia	222
A.1	Espaços topológicos	222
A.2	Topologia da ordem	229
A.3	Topologia produto	233
A.4	Topologia do subespaço	236
A.5	Conjuntos fechados e pontos aderentes	240
A.6	Funções contínuas	245
A.7	Topologia quocientes	251
A.8	Espaços conexos	254
B	Sobre homotopia	261
B.1	Homotopia de caminhos	262
B.2	Grupo fundamental	272

Capítulo 1

Introdução

Anyons são partículas que vivem em duas dimensões espaciais e apenas esta mudança de dimensionalidade confere a elas comportamentos exóticos e bastante ricos. Naturalmente, por conta desta condição, não é possível encontrar anyons “fundamentais” na natureza, no entanto, é possível encontrá-los em certos materiais. Apesar de poderem ter havido menções à ideia antes, julgamos que a história da pesquisa em anyons começa realmente na década de 1970 com os trabalhos de Laidlaw e Dewitt [1] em 1971, Leinaas e Myrheim [2] em 1977, indo para a década de 1980 com Wilczek [3, 4], Wu [5] e continuando até os dias de hoje, década de 2020 com um interesse especial nestas partículas em sua possível aplicação na construção de computadores quânticos. Em particular, em 2020 surgiram evidências bastante fortes da detecção experimental de anyons [6, 7].

Paralelamente ao começo de estudos sobre anyons, a física de sistemas bidimensionais foi fortemente impulsionada pela descoberta do efeito Hall quântico por Klitzing, Dorda e Pepper [8] em 1980. Pouco tempo depois Wilczek [3] mostrou, baseado nos trabalhos de Laughlin [9] que as excitações presentes no efeito Hall quântico, especificamente um caso particular chamado efeito Hall quântico fracionário, são, teoricamente, anyons. Desde esse momento então, o estudo teórico e experimental de anyons prosseguem lado a lado. Não trataremos, neste trabalho da questão experimental ou mesmo do efeito Hall quântico. Para o leitor interessado, um bom ponto de partida é o artigo de Stormer e Tsui [10] sobre o assunto.

Como se a década de 1980 já não estivesse agitada o suficiente, em 1980, Paul Benioff [11] introduziu a ideia de computadores quânticos. Essencialmente a ideia de um computador quântico é usar de sistemas quânticos para realizar computações. Em particular, a ideia surgiu como um método para calcular ou simular a física de sistemas quânticos, baseado na ideia de que computadores que funcionam segundo as regras da mecânica quântica, explicitamente, seriam mais eficientes em simular fenômenos quânticos. O conceito foi posteriormente levado mais adiante por Feynman [12] em 1982, por Deutsch [13] em 1985 e vários outros. Eventualmente, na década de 1990, descobriu-se, com a criação do algoritmo de Shor [14], que estes computadores podem ter aplicações além do reino da simulação de fenômenos quânticos. Isto despertou interesse generalizado no assunto.

Em particular, a importância deste algoritmo, devido a Peter Shor, é que ele resolve um

problema, a fatoração de inteiros em seus componentes primos em tempo polinomial. Isto significa que o tempo gasto para fatorar um número n cresce polinomialmente com n . Por outro lado, o melhor algoritmo clássico para a mesma tarefa gasta um tempo que cresce exponencialmente com n . A significância disto é que talvez, computadores quânticos, não só sejam muito mais rápidos que computadores clássicos para certas tarefas, mas sejam superiores. Essencialmente, acontece que problemas são classificados em termos de níveis de complexidade, por exemplo, problemas que podem ser resolvidos em tempo polinomial e problemas que não. Se um dado problema pode ser resolvido em tempo polinomial tanto por computadores quânticos quanto clássicos, e o computador quântico é mais rápido, então podemos “alcançar” o computador quântico “meramente” construindo computadores clássicos mais poderosos, por exemplo, acumulando mais unidades de processamento.

Por outro lado, se um problema não pode ser resolvido em tempo polinomial classicamente e pode, quanticamente, como o algoritmo de Shor sugere, então computadores quânticos são, neste sentido, simplesmente superiores. Isto porque os recursos para o algoritmo clássico crescem exponencialmente, então mesmo se obter computadores mais poderosos fosse simplesmente uma questão de conectar N computadores, os recursos do planeta rapidamente seriam esgotados e atingiríamos um limite. Outro motivo pelo qual resolver estes tipos de problema é importante é porque se for possível resolver um, então todos os outros serão resolvidos e isto forneceria avanços em diversas áreas. O que apresentamos aqui foi uma versão extremamente simplificada de um dos maiores problemas abertos em ciência da computação, chamado $P = NP?$. Mesmo hoje, que já existem computadores quânticos, ainda não há uma resposta definitiva sobre a capacidade destes na perspectiva de complexidade.

Bem, já estabelecemos que existe uma relação de anyons e o efeito Hall quântico. Agora, estabelecemos a relação de anyons com computação quântica. A propriedade definidora de anyons é sua estatística. Diferentemente do que ocorre em três dimensões, onde partículas são ou férmions ou bósons, anyons podem ter um comportamento estatístico intermediário. Assim, se ao permutar dois férmions, a função de onda adquire uma fase igual a -1 , e dois bósons, 1 , permutar anyons acrescenta uma fase igual a $\phi \in \mathbb{C}$, tal que $|\phi|^2 = 1$. Esta fase depende apenas do tipo de anyon. Por exemplo, permutar dois anyons de um dado tipo pode resultar em $\phi = 1/\sqrt{2} + i/\sqrt{2}$, enquanto permutar anyons de outro tipo pode resultar em $\phi = -1$. Neste sentido, o comportamento estatístico de anyons é uma generalização dos comportamentos estatísticos presentes em três dimensões.

A origem do nome anyons deriva justamente deste comportamento, *any* é a palavra inglesa para qualquer, representando qualquer estatística e *-on* é um sufixo usualmente usado para nomes de partículas. Em particular, os anyons que descrevemos são ditos **abelianos**. Existem também anyons **não abelianos**. Para estes, ao invés de uma fase, permutar dois anyons pode gerar uma rotação, ou de modo mais geral, uma transformação unitária em um subespaço degenerado de estados. A conexão com computação quântica

é que podemos armazenar informação nos vetores deste subespaço degenerado e realizar computações apenas permutando os anyons entre si, pois realizar computações, neste sentido, corresponde a realizar uma série de transformações unitárias específicas sobre o estado que guarda a informação. No entanto, anyons ainda possuem uma propriedade extra.

Esta propriedade é comumente referida como robustez topológica. Essencialmente, o comportamento estatístico de anyons é de natureza topológica. Isto significa, essencialmente, que os detalhes específicos do caminho que usamos para permutar anyons não importa. Dois caminhos que podem ser continuamente deformados um no outro produzirão a mesma transformação na função de onda. Isto é particularmente útil para a computação quântica devido à sensibilidade de sistemas quânticos. É necessário muito pouco para perturbar estes sistemas e isto pode introduzir inúmeros erros no processo de computação, se não o destruir completamente. Isolar estes sistemas é um sério problema técnico e neste sentido, anyons são uma solução. Não importa se o processo físico de permutar os anyons contenha imprecisões, o resultado será sempre o mesmo. A robustez topológica ainda se revela de outras formas, como veremos. Alertamos, porém, que nem tudo são flores. Anyons possuem certas desvantagens, a principal delas, hoje, sendo as enormes dificuldades experimentais para se obter anyons. Manipulá-los, será outro problema.

A exposição anterior é uma brevíssima exposição sobre o desenvolvimento histórico de anyons e sua relação com computação quântica. Passamos agora às informações sobre este trabalho. Como é de se esperar de partículas emergentes como anyons, sua descrição não é particularmente simples. Geralmente é feita por meio de teorias topológicas de campos ou teorias de campos conformes [15]. Estas descrições surgem explicitamente para modelar anyons enquanto excitações em um dado material. Apesar disto, os primeiros trabalhos sobre anyons os investigavam da perspectiva de mecânica quântica básica. Essencialmente tratando do que mudaria se alterássemos a dimensão do sistema em questão. Os trabalhos de Laidlaw e Dewitt e Leinaas e Myrheim [1, 2] seguem esta abordagem.

Esta abordagem é particularmente interessante pois permite estudar anyons e suas propriedades separadamente de dificuldades oriundas do fato de serem partículas emergentes em um material. Assim, se presta particularmente bem, não só para introduzir anyons sem que sejam necessários conhecimentos muito avançados. No primeiro capítulo do desenvolvimento deste trabalho, seguiremos esta abordagem para fazer uma exposição sobre anyons. Começamos com uma revisão de como sistemas de partículas idênticas são tratados em mecânica quântica simples — aquela presente nos estágios finais de cursos de graduação em física — e terminaremos com uma descrição quântica de anyons. Durante o processo, pontuaremos falhas presentes na abordagem usual que a torna incapaz de descrever sistemas anyônicos. Corrigindo estas falhas, chegaremos à quantização de anyons, que faremos de dois modos: um será uma quantização canônica de modo a obter uma

descrição na formulação de Schrödinger, baseada nos trabalhos de Leinaas e Myrheim [2] e outra será uma quantização usando a formulação de Feynman, segundo Laidlaw e Dewitt e Wu [1, 5].

O propósito deste capítulo é descrever anyons de maneira rigorosa e capturar suas propriedades fundamentais. Essencialmente, capturar a fenomenologia de anyons. A partir deste ponto, existem duas escolhas possíveis: podemos ir em direção às teorias de campo ou podemos ir em direção à descrição categorial de anyons. A descrição categorial é aquela feita por meio de uma teoria matemática chamada teoria de categorias. Como o nome do trabalho sugere, escolheremos esta opção. Escolhemos abordar anyons por meio de teoria de categorias porque assim como a descrição por mecânica quântica básica captura apenas o essencial de anyons, a teoria de categorias também o faz. Com a vantagem que a primeira encontra dificuldades ao tentar descrever anyons não abelianos enquanto teoria de categorias serve igualmente bem para anyons abelianos e não abelianos.

Em particular existe uma escassez de trabalhos que tratam de anyons não abelianos usando apenas mecânica quântica simples. Se isto é uma consequência histórica do fato da pesquisa em anyons ser muito ligada ao desenvolvimento experimental ou se é uma incapacidade inerente da formulação quântica, não está claro. Veremos alguns resultados que sugerem que é possível ao menos obter anyons não abelianos neste formalismo. Ainda neste tópico, notamos que a formulação por teorias de campos, como dissemos, são as mais usuais para tratar deste assunto, em particular teorias topológicas de campos. Em particular acreditamos que isto se dá porque é uma descrição mais própria para atacar problemas “práticos” sobre anyons. A formulação categorial por outro lado é bastante abstrata e um tanto quanto removida do reino de aplicações físicas.

Para justificar, então, o porquê de usar a formulação categorial, primeiro precisamos falar sobre o que é teoria de categorias. Teoria de categorias é uma área relativamente jovem da matemática tendo sido criada na década de 1940 por Eilenberg e Mac Lane [16]. Poderíamos descrever esta teoria essencialmente como uma visão panorâmica da matemática. Olhando de longe, muitos dos detalhes deixam de ser visíveis, em particular, realizar certos cálculos não é nada prático usando teoria de categorias. Por outro lado, a estrutura global das coisas se torna mais nítida e é possível enxergar comunalidades que antes não estavam visíveis. É importante mencionar também que teoria de categorias surgiu a partir de estudos em topologia algébrica.

Isto é importante justamente porque a descrição usual de anyons é feita por teorias **topológicas** de campos. A descrição categorial, então, é uma que surgiu a partir de estudos de natureza mais matemática sobre as estruturas algébricas das teorias de campos que descrevem anyons. Além disto a descrição categorial se aproveita da natureza de teoria de categorias e essencialmente “descarta” detalhes particulares, apresentando uma visão global de anyons. O resultado final é uma descrição completa do comportamento de anyons, isto é, sua estatística que não é afetada por detalhes particulares de como ou onde

estes anyons surgem. Esta descrição não é particularmente útil para investigações físicas como tentar determinar em que tipo de material um dado tipo de anyon pode aparecer, mas é útil para aplicações em computação quântica com anyons.

A própria teoria de categorias possui aplicações em alguns ramos de ciência da computação [17] como construção de linguagens de programação, uma tarefa necessária para que possamos usar computadores quânticos e estudo de algoritmos, o que pode ser interessante para estudar questões de complexidade em se tratando de algoritmos quânticos. Além disto, teoria de categorias tem aplicações em lógica e teoria de tipos, que também possuem aplicação em ciência da computação. Não afirmamos que a descrição categorial de anyons pode ser alavancada para todos estes assuntos, no entanto, ter uma linguagem comum para um conjunto de problemas é bastante útil.

Além disso, as ideias de teoria de categorias são relativamente simples sob certo aspecto. Elas são razoavelmente simples de se compreender, mas devido ao nível de abstração, são bastante profundas. Por este motivo e, pelo fato de que este não é um assunto comumente presente em currículos de física, dedicamos o segundo capítulo do desenvolvimento a introduzir o vocabulário categorial. Começando do básico, introduzimos os principais conceitos necessários para este trabalho acompanhados de alguns exemplos para amenizar um pouco a profundidade que cada conceito categorial carrega.

Por fim, no terceiro e último capítulo do desenvolvimento combinamos o conteúdo dos dois capítulos anteriores: a fenomenologia de anyons e a linguagem categórica para obter a descrição categorial de anyons. Como é de se esperar a descrição categorial requer conceitos além da teoria de categorias básica. Introduziremos estes conceitos conforme necessário, contextualizando-os segundo sua aplicação à descrição de anyons. Por fim, solucionaremos um modelo de anyons particularmente simples porém bastante rico chamado modelo dos anyons de Fibonacci e o usaremos para apontar como a descrição categorial é útil para propósitos de computação quântica.

Isto concluirá o desenvolvimento. Para o capítulo final de conclusões, despertaremos a natureza ensaística deste trabalho, como o título sugere. Existem dois objetivos para este trabalho:

- Reunir um conjunto básico de informações sobre anyons, desde seus fundamentos em mecânica quântica básica até sua descrição em termos de categorias e apresentá-los de maneira coerente.
- Realizar a tarefa acima de maneira didática e consistente. De modo que este trabalho possa servir como uma introdução ao assunto com pré-requisitos relativamente mínimos.

Neste sentido, acreditamos que ensaio seja o gênero textual mais adequado para descrever este trabalho, pois ele é uma exposição sobre o assunto, mas não tem pretensões de cobrir todos os seus aspectos e porque não se propõem a tirar conclusões mas levantar ainda mais

questões e possibilidades de pesquisa sobre o assunto. O capítulo de conclusão tratará precisamente desta última parte, abordando, em particular, questões que acreditamos ser um grande potencial da descrição categórica e que pretendemos investigar em pesquisas futuras. Por fim, em concordância com o segundo objetivo listado, existem dois apêndices ao final que tratam sobre topologia básica e sobre o grupo fundamental, assuntos necessários para a exposição no capítulo um do desenvolvimento.

Capítulo 2

Sobre anyons

Neste capítulo será introduzido o conceito de anyons seguindo primeiro uma ideia de complexificação gradual, isto é, começando com uma explicação mais simples e então considerando-se nuances e explorando o conceito mais a fundo, levando a uma maior complexidade e, segundo, uma abordagem histórica, passando por autores como Leinaas e Myrheim, Laidlaw e DeWitt e Wu [2, 1, 5]. Felizmente estas duas abordagens são aproximadamente equivalentes. Neste sentido, iniciamos com uma exposição sobre a mecânica quântica de partículas idênticas como apresentada nos livros-texto introdutórios da disciplina [18, 19, 20]. Em seguida, explorando o que poderia-se dizer uma falha na exposição usual, introduziremos o conceito de anyons e algumas de suas características fundamentais.

2.1 – Estatística quântica

Começemos com a seguinte definição [18]:

Definição 2.1: Duas partículas são ditas idênticas se todas as suas propriedades intrínsecas – massa, carga, spin, etc – são iguais. Isto é, não existe experimento capaz de distinguir estas propriedades em quaisquer duas partículas.

Como, no momento da escrita, não existe nenhum experimento capaz de distinguir um elétron de outro, ou um próton de outro, tem-se que todos os elétrons são idênticos, assim como todos os prótons, todos os átomos de hidrogênio e assim por diante. Note, no entanto que duas partículas serem idênticas não necessariamente implica que são indistinguíveis na prática.

Para ver isto, considere o caso de duas partículas clássicas idênticas. Suponha que no instante inicial t_0 , a configuração do sistema é definida especificando-se a posição e velocidade das duas partículas: $\{\mathbf{r}_0, \mathbf{v}_0\}$ e $\{\mathbf{r}'_0, \mathbf{v}'_0\}$. Para calcular a evolução do sistema, numeramos as partículas. \mathbf{r}_1 e \mathbf{v}_1 denotam a posição e velocidade da partícula (1) e \mathbf{r}_2 e \mathbf{v}_2 denotam a posição e velocidade da partícula (2). Note que, a princípio, esta numeração não possui nenhuma fundamentação física.

Segue que o estado inicial pode ser descrito, matematicamente, de duas formas distintas:

$$\begin{aligned} \mathbf{r}_1(t_0) &= \mathbf{r}_0 & \mathbf{v}_1(t_0) &= \mathbf{v}_0; \\ \mathbf{r}_2(t_0) &= \mathbf{r}'_0 & \mathbf{v}_2(t_0) &= \mathbf{v}'_0, \end{aligned} \quad (2.1)$$

ou

$$\begin{aligned} \mathbf{r}_1(t_0) &= \mathbf{r}'_0 & \mathbf{v}_1(t_0) &= \mathbf{v}'_0; \\ \mathbf{r}_2(t_0) &= \mathbf{r}_0 & \mathbf{v}_2(t_0) &= \mathbf{v}_0. \end{aligned} \quad (2.2)$$

Agora, suponha que a solução para as equações de movimento, tomando 2.1 como condições iniciais, seja:

$$\mathbf{r}_1(t) = \mathbf{r}(t) \quad \mathbf{r}_2(t) = \mathbf{r}'(t)$$

onde $\mathbf{r}(t)$ e $\mathbf{r}'(t)$ são funções vetoriais. Como as duas partículas são idênticas, deve-se ter que tanto a lagrangiana do sistema, $\mathcal{L}(\mathbf{r}_1, \mathbf{v}_1; \mathbf{r}_2, \mathbf{v}_2)$ quanto a hamiltoniana, $\mathcal{H}(\mathbf{r}_1, \mathbf{p}_1; \mathbf{r}_2, \mathbf{p}_2)$, são invariantes sob permutação dos índices (1) e (2). Portanto, as soluções para as equações de movimento correspondentes ao estado inicial 2.2 são:

$$\mathbf{r}_1(t) = \mathbf{r}'(t) \quad \mathbf{r}_2(t) = \mathbf{r}(t)$$

onde $\mathbf{r}(t)$ e $\mathbf{r}'(t)$ representam as mesmas funções nas soluções para cada uma das condições iniciais.

Ambas as descrições do sistema físico são, portanto, inteiramente equivalentes, já que levam as mesmas previsões físicas. A partícula que começou no estado $\{\mathbf{r}_0, \mathbf{v}_0\}$ está em $\mathbf{r}(t)$, com velocidade $\mathbf{v}(t) = d\mathbf{r}/dt|_t$ no instante t e aquela que começou no estado $\{\mathbf{r}'_0, \mathbf{v}'_0\}$ está em $\mathbf{r}'(t)$, com velocidade $\mathbf{v}'(t) = d\mathbf{r}'/dt|_t$ no mesmo instante. Note que, nesta situação, tudo o que precisamos fazer é escolher, no instante inicial uma das descrições – 2.1-2.2 – e ignorar a existência da outra. Ao fazer isto, tratamos ambas as partículas como se não fossem idênticas ou, melhor dizendo, tratamo-las como distinguíveis. Os números (1) e (2) passam a atuar **como se fossem** propriedades intrínsecas das partículas, efetivamente é como se pintássemos uma partícula de uma cor e a outra de outra cor. Como podemos seguir a trajetória de cada partícula passo-a-passo, podemos determinar a posição de cada partícula em qualquer instante.

Se usamos a trajetória para distinguir partículas clássicas idênticas, segue de imediato que a situação é diferente no caso de partículas quânticas já que estas não possuem trajetórias bem definidas. Duas partículas quânticas idênticas, em geral, também são indistinguíveis. Isto se dá porque, mesmo supondo que em um instante inicial, t_0 , os pacotes de onda de ambas as partículas estejam completamente separados, se houver

algum instante posterior onde eles se misturam, isto é, houver alguma posição, \mathbf{r} e instante, t , onde a probabilidade de se encontrar ambas as partículas seja não nula, então não poderemos distinguir qual das duas partículas, (1) ou (2) seria detectada lá. A exceção ocorre quando ao longo de toda a evolução considerada, as funções de onda das duas partículas nunca são não nulas na mesma posição e no mesmo instante. Como este caso não costuma ser de muito interesse, geralmente, ignora-se essa possibilidade e diz-se, prontamente que partículas quânticas idênticas são indistinguíveis. Para estudar o caso quântico em mais detalhes, analisemos dois exemplos concretos: o de colisão de partículas e duas partículas de spin meio.

Colisão de partículas idênticas

Considere a colisão de duas partículas idênticas tratada no referencial do centro de massa. Cada uma das partículas é descrita por uma coleção de números quânticos, representados por q_i e q'_i , que correspondem a um conjunto completo de observáveis. Assim, o estado das partículas é $|q_i\rangle$ e $|q'_i\rangle$. No instante inicial, ambas estão completamente separadas, isto é, não existe superposição de suas funções de onda. Deste modo, poderíamos numerá-las (1) e (2) tal como no caso clássico. Aqui surge a primeira distinção. No caso clássico, a numeração das partículas, que a princípio é sem fundamentos, torna-se real por ser possível acompanhar a trajetória de cada partícula. Em outras palavras, torna-se real devido à informação da trajetória. No caso quântico, não só não possuímos informação sobre a trajetória de cada partícula como também toda a informação possível de se conhecer sobre a partícula está codificada no conjunto de observáveis dados, aos quais correspondem os estados $|q_i\rangle$ e $|q'_i\rangle$. Portanto, mais que uma numeração arbitrária, ela é, em princípio, sem significado desde o início, exceto no caso onde não existe superposição das funções de onda.

Apesar disso, mantêm-se a numeração como um modo de auxiliar na compreensão do problema e assim o faremos por hora. Adiante veremos como esta numeração introduz alguns problemas de conceituação e como descrever sistemas de partículas idênticas sem menção a rótulos deste tipo. De todo modo, assim como no caso clássico então podemos atribuir o estado $|q_i\rangle$ à partícula (1) e $|q'_i\rangle$, à (2) ou vice-versa. Deste modo, os estados iniciais possíveis são:

$$|q_i\rangle|q'_i\rangle \quad |q'_i\rangle|q_i\rangle.$$

Aqui, a justaposição $|\bullet\rangle|\bullet\rangle$ é uma abreviação do produto tensorial: $|\bullet\rangle \otimes |\bullet\rangle$. Note que, estes dois estados são distintos a menos que tenhamos $q_i = q'_i$. No instante inicial, quando as duas partículas estão completamente separadas, podemos distingui-las e assim escolhemos um dos dois estados acima segundo uma numeração arbitrária. Observe que neste instante, esta numeração é fundamentada pelo fato de não haver superposição das funções de onda das duas partículas, assim, poderíamos extrair a informação a partir de

algum observável. Por exemplo, poderíamos denominar as regiões onde a função de onda de cada partícula é não nula de \mathcal{R}_1 e \mathcal{R}_2 respectivamente. Um detector em \mathcal{R}_1 detectaria a partícula que rotulamos de (1) e um na região \mathcal{R}_2 , a partícula (2).

A situação muda, no entanto, após a colisão. A partir deste instante, deixamos o caso de exceção e devemos encarar a indistinguibilidade das partículas. Denominando a região da colisão por \mathcal{R} , centrada na posição do centro de massa e suponha que coloquemos um detector em algum lugar na fronteira de \mathcal{R} . Se os estados finais das partículas, após a colisão são $|q_f\rangle$ e $|q'_f\rangle$ e o detector encontrar uma partícula, digamos, no estado $|q_f\rangle$, não poderemos dizer se esta partícula é aquela que originalmente numeramos por (1) ou por (2). A partir do instante que houve uma superposição das funções de onda das duas partículas, a possibilidade de distingui-las deixou de existir. Na perspectiva do exemplo anterior, de usar a posição para diferenciá-las, por um período após a colisão, não existem nenhuma região \mathcal{R}' onde apenas a função de onda de uma partícula seja não nula. Deste modo, não pode-se saber se uma partícula detectada nesta região é aquela originalmente numerada (1) ou (2).

Posteriormente, um longo tempo após a colisão, podemos voltar a ter uma separação das funções de onda, mas mesmo assim, não podemos determinar os rótulos originais atribuídos a cada partícula pois “as perdemos de vista” por algum tempo após a colisão. Tomando a ideia de trajetórias segundo Feynman, podemos dizer que existem dois tipos de trajetórias possíveis para as partículas, representadas na figura 2.1 seguir.

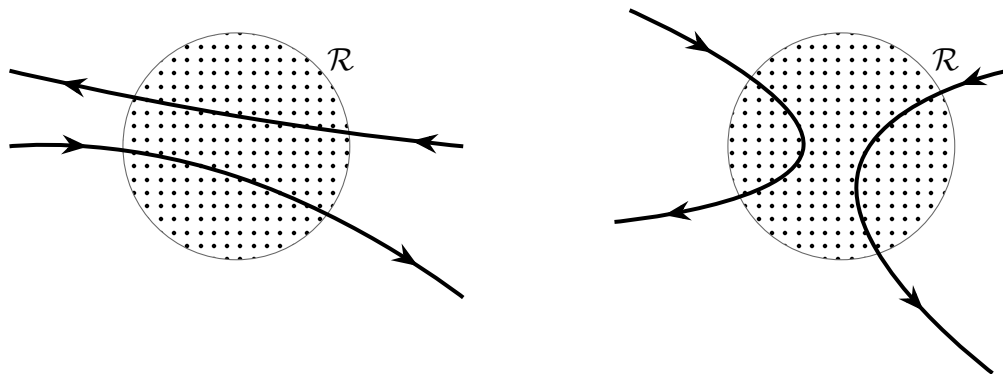


Figura 2.1: Os dois tipos de trajetórias possíveis para as duas partículas na colisão.

Podemos convencionar que o detector se encontra à esquerda da região \mathcal{R} e que a partícula direcionada para a esquerda é numerada (1) e a direcionada para a direita, (2). Assim, no primeiro tipo de trajetória, o detector mede a partícula (1) e no segundo, a (2). Os dois estados finais possíveis são:

$$|q_f\rangle|q'_f\rangle \quad |q'_f\rangle|q_f\rangle.$$

Como não é possível extrair destes kets informação o suficiente para determinar a qual partícula pertence cada número, diferentemente da situação antes da colisão, não

podemos mais simplesmente escolher um destes kets. Ambos devem ser levados em conta. Pensando segundo Feynman, todas as trajetórias possíveis contribuem, de modo que não podemos simplesmente ignorar uma classe de caminhos. A questão que permanece é: deveríamos somar as probabilidades relativas a cada uma dos dois tipos de trajetórias? Ou deveríamos somar as amplitudes de probabilidade? Responderemos esta pergunta mais adiante, antes, vejamos outro exemplo.

Duas partículas de spin 1/2

Considere duas partículas idênticas de spin 1/2. Confinar-nos-emos a estudar apenas os graus de liberdade de spin. Suponha que uma partícula possua, inicialmente, a componente z do spin igual a $+\hbar/2$ e a outra, $-\hbar/2$. Representaremos estes estados por $|+\rangle$ e $|-\rangle$ respectivamente. Deste modo, podemos escolher numerar as partículas e obtemos, a depender da numeração, duas descrições possíveis:

$$|+\rangle|-\rangle \quad |-\rangle|+\rangle.$$

Suponha, agora, que realizássemos uma medição da componente z do spin das duas partículas. Sabemos que o resultado da medida seria que uma partícula possui a componente z do spin igual a $+\hbar/2$ e outra igual a $-\hbar/2$. No entanto não podemos saber se a este estado físico correspondem os kets $|+\rangle|-\rangle$ ou $|-\rangle|+\rangle$, ou mesmo uma combinação linear:

$$\alpha|+\rangle|-\rangle + \beta|-\rangle|+\rangle$$

onde a condição de normalização dos kets impõe a condição de que $|\alpha|^2 + |\beta|^2 = 1$.

Isto se dá por uma inconsistência em qualquer numeração que escolhamos e as informações que nos são acessíveis. Qualquer numeração que escolhamos será baseada no seguinte predicado: “a partícula cujo spin possui componente z positiva será numerada (1) e a outra (2)”. Porém, uma medição, como vimos, não é capaz de nos revelar qual partícula é qual, por serem idênticas, qualquer uma pode estar em qualquer um destes estados. Nossa numeração só poderia ser válida se possuíssemos mais informação. Por exemplo, se os pacotes de onda das duas partículas estivessem separados, poderíamos afirmar, que a partícula (1) está à esquerda e a (2) à direita da origem do sistema de coordenadas e então, realizando uma medição do spin, poderíamos dizer que a partícula cuja componente z do spin é positiva é aquela numerada (2), por exemplo, donde teríamos certeza que o ket que representa o estado do sistema é:

$$|-\rangle|+\rangle.$$

Como confinamo-nos ao estudo apenas dos graus de liberdade de spin, com o propósito de explorar mais as implicações da indistinguibilidade, vamos assumir, por exemplo, que

as duas partículas estão confinadas em uma mesma região de modo que não se pode separar suas funções de onda.

O desconhecimento sobre o ket correspondente ao estado físico dos sistema produz um fenômeno conhecido como **degenerescência de troca**. Para investigarmos este conceito, suponha que agora desejássemos medir a componente x do spin das partículas, em particular, que desejássemos saber qual a probabilidade das duas partículas serem encontradas com a componente x do spin igual a $+\hbar/2$. Sabe-se, dos resultados básicos da mecânica quântica que, na base da componente z , o ket $|+\rangle_x$ é escrito como:

$$|+\rangle_x = \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle].$$

Assim, desejamos saber a probabilidade do sistema se encontrar no estado:

$$\begin{aligned} |+\rangle_x |+\rangle_x &= \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle] \otimes \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle] \\ &= \frac{1}{2} [|+, +\rangle + |+, -\rangle + |-, +\rangle + |-, -\rangle]. \end{aligned}$$

Onde a notação $|+, +\rangle$ é apenas uma forma concisa de escrever $|+\rangle|+\rangle$. Note que a posição dos símbolos em relação à vírgula importa. $|+, -\rangle = |+\rangle|-\rangle$ enquanto $|-, +\rangle = |-\rangle|+\rangle$.

Segundo os postulados da mecânica quântica e tomando como estado inicial a superposição dos kets $|+, -\rangle$ e $|-, +\rangle$, a probabilidade de encontrarmos a componente x do spin de ambas as partículas igual a $+\hbar/2$ é:

$$\left| \frac{1}{2} (\alpha + \beta) \right|^2.$$

Ou seja, a probabilidade depende do estado que escolhemos. Deste modo, não é possível descrever o sistema físico em questão usando o ket $\alpha|+, -\rangle + \beta|-, +\rangle$. A degenerescência de troca — isto é, a existência de múltiplos kets que descrevem a mesma situação física — deve ser removida. Em outras palavras, é necessário construir um ket que descreva adequadamente o sistema, garantindo a boa definição da probabilidade.

Caso geral

Antes de prosseguirmos, no entanto, é interessante notar que o conceito de degenerescência de troca é generalizado facilmente para sistemas com mais de duas partículas. Para tanto, consideremos um sistema de três partículas idênticas e indistinguíveis. Seja \mathcal{E} o espaço de estados de uma destas partículas individualmente e suponhamos que um operador \mathcal{B} sozinho forme um conjunto completo de observáveis compatíveis em \mathcal{E} e que

seu espectro seja discreto, por conveniência. Assim, uma tal partícula é descrita por kets $|b_i\rangle$ onde $i \in \mathbb{N}$.

O espaço de estados de N tais partículas juntas seria, então, simplesmente o produto tensorial de N cópias de \mathcal{E} :

$$\mathcal{E}^N = \bigotimes_{i=1}^N \mathcal{E}.$$

A base deste espaço de estados, conseqüentemente, é o produto tensorial de todos os ket base de \mathcal{E} entre si N vezes:

$$\mathbf{B}(\mathcal{E}^N) = \left\{ \bigotimes_{i=1}^N |b_j\rangle \mid |b_j\rangle \in \mathbf{B}(\mathcal{E}) \right\}$$

onde $j \in \mathbb{N}$ e $\mathbf{B}(\mathcal{E})$ é uma base de \mathcal{E} . Agora, é importante notar um detalhe sobre o produto tensorial acima, especialmente porque desejaremos numerar as partículas logo em seguida. Estritamente falando, reordenações distintas dos fatores do produto tensorial do espaço de estados de uma partícula fornecem espaços que são idênticos. Porém, quando numeramos as partículas, numerando os espaços correspondentes, as mesmas reordenações fornecem espaços distintos porém trivialmente isomorfos. Isto se dá basicamente porque a numeração atua como uma informação extra sobre as partículas, tratando-as efetivamente como se fossem distinguíveis, como já discutido.

Para ver isto, considere o caso de três partículas por exemplo. Podemos escolher uma numeração arbitrária para as partículas e então numerar os espaços de estado correspondente a cada partículas. O espaço de estados do sistema então poderia ser, por exemplo:

$$\mathcal{E}^{123} = \mathcal{E}^1 \otimes \mathcal{E}^2 \otimes \mathcal{E}^3$$

Neste caso, os vetores da base seriam da forma:

$$|1 : b_i, 2 : b_j, 3 : b_k\rangle$$

onde $i, j, k \in \mathbb{N}$. Este ket indica que a partícula numerada 1 está no estado b_i , aquela numerada 2 no estado b_j e, 3 no estado b_k . Note que este ket representaria, em princípio — pois não podemos realmente distinguir as partículas — um estado diferente daquele representado pelo ket

$$|2 : b_i, 1 : b_j, 3 : b_k\rangle$$

que corresponderia a um ket da base do espaço

$$\mathcal{E}^{213} = \mathcal{E}^2 \otimes \mathcal{E}^1 \otimes \mathcal{E}^3.$$

Naturalmente estes espaços são isomorfos, como afirmado anteriormente, basta permutar

tar os índices das partículas. Por exemplo, usando os dois kets anteriores, o isomorfismo seria um mapa, M , tal que:

$$\begin{aligned} M|1 : b_i, 2 : b_j, 3 : b_k\rangle &= |1 : b_j, 2 : b_i, 3 : b_k\rangle \\ &= |2 : b_i, 1 : b_j, 3 : b_k\rangle \end{aligned}$$

A importância deste detalhe surge durante o processo de medição que abordamos agora, retornando ao caso geral. Os autokets do observável \mathcal{B} formam uma base para \mathcal{E} . O próprio observável \mathcal{B} pode ser usado para definir inúmeros observáveis em \mathcal{E}^N . Primeiro, podemos simplesmente estendê-lo para \mathcal{E}^N tomando o produto tensorial de \mathcal{B} com $N - 1$ vezes a identidade:

$$\mathcal{B}_N^1 = \mathcal{B} \otimes \mathbb{1}_{N-1}.$$

Onde N indica meramente que este é o operador \mathcal{B} em \mathcal{E}^N e 1 indica que ele atua apenas em uma partícula. Pela definição de \mathcal{B}_N^1 vê-se prontamente que ele atua apenas em uma das N partículas, em particular sobre aquela na primeira posição no ket em questão. Existem N tais operadores de uma partícula, um para cada posição possível de \mathcal{B} no produto. No caso de $N = 3$, por exemplo, tem-se:

$$\mathcal{B} \otimes \mathbb{1} \otimes \mathbb{1}; \quad \mathbb{1} \otimes \mathcal{B} \otimes \mathbb{1}; \quad \mathbb{1} \otimes \mathbb{1} \otimes \mathcal{B}.$$

Apesar destes observáveis revelarem o estado da partícula que ocupa uma determinada posição dada pela ordenação que escolhemos no produto dos espaço de estados, não possuem nenhuma utilidade para identificar qual partícula é qual, caso contrário não se tratariam de partículas indistinguíveis afinal. A consequência disso é que ao usar qualquer um destes operadores, podemos apenas afirmar que **uma** partícula está em determinado estado.

Podemos, no entanto, usá-los em sequência para medir o estado de todas as partículas, o que é equivalente a realizar uma medida do observável \mathcal{B}_N^N dado por:

$$\mathcal{B}_N^N = \bigotimes_{i=1}^N \mathcal{B}.$$

A degenerescência de troca surge, porque mesmo que aplicássemos \mathcal{B}_N^N a um ket da base de \mathcal{E}^N , por exemplo, poderíamos obter apenas a informação de que uma partícula está no estado b_i , outra no estado b_j , duas no estado b_k e assim por diante, sem nenhuma informação sobre qual é qual. Naturalmente isto era esperado, afinal se existisse algum observável capaz de revelar informações sobre qual partícula é qual, segundo uma numeração arbitrária escolhida, elas não seriam indistinguíveis. No caso de $N = 3$ partículas, uma medida de \mathcal{B}_3^3 poderia revelar o seguinte conjunto de autovalores: $\{b_i, b_j, b_k\}$ aos quais

poderiam corresponder os seguintes seis kets:

$$\begin{array}{ll} |1 : b_i, 2 : b_j, 3 : b_k\rangle & |1 : b_i, 3 : b_j, 2 : b_k\rangle \\ |2 : b_i, 1 : b_j, 3 : b_k\rangle & |2 : b_i, 3 : b_j, 1 : b_k\rangle \\ |3 : b_i, 2 : b_j, 1 : b_k\rangle & |3 : b_i, 1 : b_j, 2 : b_k\rangle \end{array}$$

É interessante notar, no entanto, que todos estes kets são equivalentes entre si segundo os isomorfismos que relacionam os seis espaços de estados distintos para um sistema de três partículas idênticas: $\mathcal{E}^{123}, \mathcal{E}^{132}, \mathcal{E}^{213}, \mathcal{E}^{231}, \mathcal{E}^{321}$ e \mathcal{E}^{312} . Uma permutação dos índices das partículas pode levar qualquer um destes kets em qualquer outro. Isto nos sugere abandonar a numeração das partículas, que desde o início não possui significado, ou mesmo realidade física e buscar uma representação onde todos estes seis kets, por exemplo, sejam identificados com um único ket. Note que a identificação aqui é um processo diferente relacioná-los por via de um isomorfismo. Não faremos isto ainda, continuaremos por mais algum tempo usando o formalismo que é usualmente apresentado em livros didáticos, explorando agora o papel que permutações desempenham na descrição de um sistema de partículas idênticas.

2.1.1 - Permutações

O último exemplo da seção anterior — **Caso geral** — deixa especialmente claro que ação de permutar os números atribuídos às partículas desempenha um papel importante na teoria. Como veremos nesta e na próxima seção, é através de permutações que pode-se construir os kets adequados para descrever sistemas de partículas idênticas e indistinguíveis. Nesta seção abordaremos permutações removidas do contexto particular da mecânica quântica de partículas idênticas e na seção seguinte abordaremos sua utilização no problema em questão.

Dado $k \in \mathbb{Z}_+$, definiremos o conjunto \mathbb{C}_k por $\mathbb{C}_k = \{1, \dots, k\}$. Estes conjuntos serão usados para definir e estudar permutações, servindo como conjuntos arbitrários de k elementos. Apesar de representarmos os elementos dos conjuntos \mathbb{C}_k como $1, \dots, k$, é interessante pensar nestes elementos como sendo arbitrários. Por este motivo, ocasionalmente poderemos representar os elementos como letras indexadas: a_1, \dots, a_k , por exemplo.

Definição 2.2: Uma **permutação**, σ , é uma bijeção

$$\sigma : \mathbb{C}_k \rightarrow \mathbb{C}_k.$$

Definição 2.3: Uma permutação $\sigma : \mathbb{C}_k \rightarrow \mathbb{C}_k$ é dita **cíclica de comprimento r** ou um **r-ciclo** e representada por $(c_1, c_2 \cdots c_r)$ se

$$\sigma(c_1) = c_2, \quad \sigma(c_2) = c_3, \quad \cdots \quad \sigma(c_{r-1}) = c_r, \quad \sigma(c_r) = c_1$$

vale e todos os outros $k - r$ elementos são mantidos fixos. Uma **transposição** é um ciclo de comprimento dois.

O conjunto de todas as permutação $\sigma : \mathbb{C}_k \rightarrow \mathbb{C}_k$ para um dado k é denotado por \mathcal{S}_k . É possível definir um produto em \mathcal{S}_k dado pela composição usual de funções. Assim, dada duas permutações $\sigma, \tau : \mathbb{C}_k \rightarrow \mathbb{C}_k$, o produto delas, $\sigma\tau$ é dado pela composição $\sigma \circ \tau$, nesta ordem, primeiro aplica-se τ e depois σ .

Proposição 2.1: O conjunto \mathcal{S}_k equipado com o produto definido acima forma o **grupo simétrico de grau k**.

Prova: A prova requer que verifiquemos se \mathcal{S}_k com o produto definido satisfaz os quatro axiomas de grupo:

i Se $\sigma, \tau \in \mathcal{S}_k$ implica que $\sigma\tau \in \mathcal{S}_k$.

ii Existe $\mathbf{1} \in \mathcal{S}_k$ tal que para todo $\sigma \in \mathcal{S}_k$ vale

$$\mathbf{1}\sigma = \sigma = \sigma\mathbf{1}.$$

iii Para todo $\sigma \in \mathcal{S}_k$ existe $\sigma^{-1} \in \mathcal{S}_k$, única, tal que vale

$$\sigma\sigma^{-1} = \mathbf{1} = \sigma^{-1}\sigma.$$

iv Para todo $\sigma, \tau, \rho \in \mathcal{S}_k$ vale

$$\sigma(\tau\rho) = (\sigma\tau)\rho.$$

- Pela definição de permutações, uma permutação é uma bijeção em \mathbb{C}_k , como a composição de duas bijeções também é uma bijeção, o produto de duas permutações também é uma permutação e vale o primeiro axioma.
- A permutação identidade, $\mathbf{1}$ é simplesmente a função identidade em \mathbb{C}_k . Como a identidade não permuta nenhum elemento, compor qualquer permutação com a identidade de qualquer um dos lados resulta na permutação original, de modo que o segundo axioma também é satisfeito.

- Como cada permutação é uma bijeção e toda bijeção possui uma inversa única, o terceiro axioma é satisfeito.
- Para verificar o quarto axioma considere a ação de $\sigma(\tau\rho)$ em um elemento $c \in \mathbb{C}_k$:

$$[\sigma(\tau\rho)](c) = \sigma[(\tau\rho)(c)] = \sigma(\tau(\rho(c))) = (\sigma\tau)(\rho(c)) = [(\sigma\tau)\rho](c).$$

Como isto vale para qualquer $c \in \mathbb{C}_k$, tem-se: $\sigma(\tau\rho) = (\sigma\tau)\rho$.

□

Qualquer permutação $\sigma : \mathbb{C}_k \rightarrow \mathbb{C}_k$ pode ser descrita de duas maneiras: [21]

- Como uma matriz:

$$\begin{bmatrix} 1 & 2 & \cdots & k \\ \sigma(1) & \sigma(2) & \cdots & \sigma(k) \end{bmatrix}.$$

- Como um produto de ciclos disjuntos:

$$(a_1 \cdots a_{r_1})(b_1 \cdots b_{r_2}) \cdots$$

Note que o fato de que permutações podem ser decompostas em ciclos disjuntos não é óbvio porém não apresentaremos a prova desta afirmação aqui. Em particular, afirmamos também sem prova que qualquer ciclo pode ser decomposto em um produto de transposições. Esta decomposição não é única, no entanto, dada quaisquer duas decomposições de um dado ciclo, ambas são um produto ou de um número par ou de um número ímpar de transposições [22]. Este fato sustenta a utilidade da segunda representação e a definição a seguir:

Definição 2.4: Uma permutação $\sigma \in \mathcal{S}_k$ é dita **par** ou **ímpar** se é o produto de um número par ou ímpar de transposições, respectivamente. O **signal** de uma permutação σ , $sgn \sigma$, é definido de modo que:

$$sgn \sigma = \begin{cases} +1, & \text{se } \sigma \text{ for par;} \\ -1, & \text{se } \sigma \text{ for ímpar.} \end{cases}$$

e

$$sgn(\sigma\tau) = sgn(\sigma)sgn(\tau).$$

para $\sigma, \tau \in \mathcal{S}_k$.

Uma nota antes de prosseguir: apresentamos aqui a definição da paridade de uma

permutação um tanto quanto desprovida dos resultados usuais no estudo de permutações. O conceito de paridade é sustentado pela prova do resultado de que as fatorações de uma dada permutação são sempre o produto ou de um número par ou ímpar de fatores. Assim, faz-se necessário apresentar, sem prova o seguinte resultado: toda transposição é ímpar [22].

Com estas informações sobre permutações, passamos a estudar a aplicação destes conceitos na mecânica quântica de várias partículas idênticas e indistinguíveis. Começamos por definir operadores de permutação da seguinte maneira. Considere um sistema de N partículas idênticas e indistinguíveis e seja $\mathcal{L}(\mathcal{H})$ o conjunto de todos os operadores lineares $L : \mathcal{H} \rightarrow \mathcal{H}$, onde \mathcal{H} é o espaço de Hilbert do sistema em questão. Definimos o mapa R tal que para cada permutação $\sigma \in \mathcal{S}_N$ é atribuída um operador $P_\sigma = R(\sigma) \in \mathcal{L}(\mathcal{H})$:

$$\begin{aligned} R : \mathcal{S}_N &\longrightarrow \mathcal{L}(\mathcal{H}) \\ \sigma &\longrightarrow (P_\sigma : \mathcal{H} \rightarrow \mathcal{H}) \end{aligned}$$

Definimos dois modos de denominar os operadores em questão. Usaremos $R(\sigma)$ quando estivermos interessados em falar do mapa R e usaremos P_σ quando estivermos interessados em falar dos operadores. Impomos ainda, as seguintes propriedades ao mapa R e aos operadores P_σ :

i Dados $\sigma, \tau \in \mathcal{S}_N$, tem-se que:

$$R(\sigma\tau) = R(\sigma)R(\tau).$$

ii Os operadores $R(\sigma)$, $\sigma \in \mathcal{S}_k$ são bijeções.

iii Dado um ket $|1 : u_i, 2 : u_j, \dots, N : u_k\rangle \in \mathcal{H}$, onde os u_i são multi-índices representando os números quânticos relevantes das partículas e $P_\sigma \in R(\mathcal{S}_N)$, tem-se:

$$P_\sigma |1 : u_i, 2 : u_j, \dots, N : u_k\rangle = |\sigma(1) : u_i, \sigma(2) : u_j, \dots, \sigma(N) : u_k\rangle.$$

As duas primeiras propriedades definem o mapa R como uma representação do grupo \mathcal{S}_k em \mathcal{H} , o que fornece várias propriedades úteis como veremos a seguir. A terceira propriedade meramente define a ação de cada um dos operadores. Eles simplesmente permutam os índices das partículas. Como estas propriedades estão sendo impostas é interessante verificar se elas são compatíveis. Em particular, dada a atuação dos operadores P_σ nos kets do espaço de Hilbert, verifiquemos se eles satisfazem às propriedades i e ii.

Primeiro, considere a aplicação sucessiva dos operadores P_σ e P_τ , $\sigma, \tau \in \mathcal{S}_N$ em um

ket arbitrário:

$$\begin{aligned}
P_\sigma P_\tau |1 : u_i, 2 : u_j, \dots, N : u_k\rangle &= P_\sigma |\tau(1) : u_i, \tau(2) : u_j, \dots, \tau(N) : u_k\rangle \\
&= |\sigma(\tau(1)) : u_i, \sigma(\tau(2)) : u_j, \dots, \sigma(\tau(N)) : u_k\rangle \\
&= |(\sigma\tau)(1) : u_i, (\sigma\tau)(2) : u_j, \dots, (\sigma\tau)(N) : u_k\rangle \\
&= P_{\sigma\tau} |1 : u_i, 2 : u_j, \dots, N : u_k\rangle.
\end{aligned}$$

Vê-se que a ação dos operadores de permutação, como definida, é compatível com a primeira propriedade. Verifiquemos se esta ação fornece operadores são bijetivos. Para provar a injetividade, considere dois kets distintos: $|1 : u_i, 2 : u_j, \dots, N : u_k\rangle$ e $|1 : u'_i, 2 : u'_j, \dots, N : u'_k\rangle$ pertencentes a \mathcal{H} . Dada uma permutação $\sigma \in \mathcal{S}_N$, aplicamos o operador P_σ em ambos os kets:

$$\begin{aligned}
P_\sigma |1 : u_i, 2 : u_j, \dots, N : u_k\rangle &= |\sigma(1) : u_i, \sigma(2) : u_j, \dots, \sigma(N) : u_k\rangle; \\
P_\sigma |1 : u'_i, 2 : u'_j, \dots, N : u'_k\rangle &= |\sigma(1) : u'_i, \sigma(2) : u'_j, \dots, \sigma(N) : u'_k\rangle.
\end{aligned}$$

Os kets resultantes serão iguais apenas se $u_i = u'_i, u_j = u'_j, \dots, u_k = u'_k$, ou seja, apenas se os kets iniciais forem iguais. Portanto, P_σ é injetiva para qualquer $\sigma \in \mathcal{S}_N$. Para provar a sobrejetividade, deve-se mostrar que para cada permutação $\sigma \in \mathcal{S}_N$ dado um ket arbitrário $|1 : u_i, 2 : u_j, \dots, N : u_k\rangle \in \mathcal{H}$ existe pelo menos um ket $|1 : u'_i, 2 : u'_j, \dots, N : u'_k\rangle \in \mathcal{H}$ pertencentes a \mathcal{H} tal que:

$$P_\sigma |1 : u'_i, 2 : u'_j, \dots, N : u'_k\rangle = |1 : u_i, 2 : u_j, \dots, N : u_k\rangle.$$

Tal ket existe e é dado por:

$$P_{\sigma^{-1}} |1 : u_i, 2 : u_j, \dots, N : u_k\rangle = |\sigma^{-1}(1) : u_i, \sigma^{-1}(2) : u_j, \dots, \sigma^{-1}(N) : u_k\rangle.$$

De fato, pela propriedade i, tem-se:

$$\begin{aligned}
P_\sigma [P_{\sigma^{-1}} |1 : u_i, 2 : u_j, \dots, N : u_k\rangle] &= P_\sigma |\sigma^{-1}(1) : u_i, \sigma^{-1}(2) : u_j, \dots, \sigma^{-1}(N) : u_k\rangle \\
&= |(\sigma\sigma^{-1})(1) : u_i, (\sigma\sigma^{-1})(2) : u_j, \dots, (\sigma\sigma^{-1})(N) : u_k\rangle \\
&= |1 : u_i, 2 : u_j, \dots, N : u_k\rangle
\end{aligned}$$

Como \mathcal{S}_N é um grupo toda permutação $\sigma \in \mathcal{S}_N$ possui uma inversa, de modo que é sempre possível construir este ket para qualquer permutação e qualquer ket de \mathcal{H} . Logo P_σ é, também sobrejetivo e portanto, bijetivo. Ou seja, não existe nenhuma inconsistência entre as três propriedades de modo que podemos continuar o estudo dos operadores P_σ .

Devido a \mathbb{R} ser uma representação, existem alguns resultados bastante úteis que podemos aproveitar.

Proposição 2.2: Decorrem da propriedade i e ii os seguintes resultados:

- $R(\mathbf{1}) = \mathbb{1}$, onde $\mathbf{1} \in \mathcal{S}_N$ e $\mathbb{1} \in \mathcal{L}(\mathcal{H})$.
- $R(\sigma^{-1}) = [R(\sigma)]^{-1}$.

Prova: Primeiro provamos o resultado auxiliar:

$$R(\sigma\tau^{-1}) = R(\sigma)[R(\tau)]^{-1} \quad \text{e} \quad R(\tau^{-1}\sigma) = [R(\tau)]^{-1}R(\sigma)$$

Tem-se:

$$R(\sigma\tau^{-1})R(\tau) = R[(\sigma\tau^{-1})\tau] = R(\sigma).$$

Como os operadores $R(\tau)$ são bijeções, eles possuem uma inversa única. Assim, multiplicando-se os dois lados por $[R(\tau)]^{-1}$:

$$R(\sigma\tau^{-1})R(\tau)[R(\tau)]^{-1} = R(\sigma\tau^{-1}) = R(\sigma)[R(\tau)]^{-1}.$$

A prova do outro resultado é análoga:

$$R(\tau)R(\tau^{-1}\sigma) = R[\tau(\tau^{-1}\sigma)] = R(\sigma).$$

O que implica que:

$$[R(\tau)]^{-1}R(\tau)R(\sigma\tau^{-1}) = R(\tau^{-1}\sigma) = [R(\tau)]^{-1}R(\sigma).$$

Deste resultado, decorrem diretamente as afirmações. Tomando $\tau = \sigma$, tem-se:

$$R(\mathbf{1}) = R(\sigma\sigma^{-1}) = R(\sigma)[R(\sigma)]^{-1} = \mathbb{1}.$$

E tomando-se $\tau = \mathbf{1}$:

$$R(\sigma^{-1}) = R(\mathbf{1}\sigma^{-1}) = \mathbb{1}[R(\sigma)]^{-1} = [R(\sigma)]^{-1}.$$

□

Além destes resultados, podemos usar o fato de que qualquer permutação σ pode ser escrita como um produto de transposições para estudar os operadores P_σ pois, segundo a

propriedade i, tem-se:

$$\begin{aligned}
 R(\sigma) &= R(\tau_1\tau_2\cdots\tau_n) \\
 &= R(\tau_1)R(\tau_2\cdots\tau_n) \\
 &\quad \vdots \\
 &= R(\tau_1)R(\tau_2)\cdots R(\tau_n)
 \end{aligned}$$

onde τ_i são transposições. Ou seja, cada operador P_σ pode ser escrito como um produto de operadores P_{τ_i} . Por este motivo, os principais resultados acerca destes operadores podem ser obtidos estudando-se apenas operadores P_τ onde τ é uma transposição. A propriedade mais interessante de transposições no estudo dos operadores de permutação é:

Proposição 2.3: Dada uma transposição τ , tem-se que $\tau^2 = \mathbf{1}$, ou seja, transposições são suas próprias inversas.

Prova: A prova é simples, considere uma transposição $\tau = (ab)$. Sabemos que $\tau(a) = b$ e $\tau(b) = a$ enquanto todos os outros elementos permanecem fixos. Assim, tem-se:

$$\begin{aligned}
 \tau^2(a) &= \tau(\tau(a)) = \tau(b) = a; \\
 \tau^2(b) &= \tau(\tau(b)) = \tau(a) = b.
 \end{aligned}$$

Como todos os outros elementos permanecem fixos, tem-se que $\tau^2 = \mathbf{1}$.

□

Desta propriedade segue diretamente que $[P_\tau]^2 = \mathbb{1}$ para τ uma transposição. Além disso, é simples mostrar que para toda transposição $\tau = (ab)$, P_τ é hermitiano. Para ver isto, considere dois kets base de \mathcal{H} que escreveremos, por conveniência, como: $|a : u_i, b : u_j, \dots\rangle$ e $|a : u_{i'}, b : u_{j'}, \dots\rangle$. Note que apenas reordenamos os termos no ket por conveniência, já que a transposição mantém todos os outros termos inalterados. Aplicando-se o operador P_τ , obtém-se:

$$\begin{aligned}
 \langle a : u_{i'}, b : u_{j'}, \dots | P_\tau | a : u_i, b : u_j, \dots \rangle &= \langle a : u_{i'}, b : u_{j'}, \dots | b : u_i, a : u_j, \dots \rangle \\
 &= \delta_{i'j} \delta_{j'i}.
 \end{aligned}$$

E, aplicando-se o operador P_τ^\dagger :

$$\begin{aligned} \langle a : u_{i'}, b : u_{j'}, \dots | P_\tau^\dagger | a : u_i, b : u_j, \dots \rangle &= \langle b : u_{i'}, a : u_{j'}, \dots | a : u_i, b : u_j, \dots \rangle \\ &= \delta_{i'j} \delta_{j'i}. \end{aligned}$$

Logo, $P_\tau^\dagger = P_\tau$. Aliando-se a hermiticidade com o fato de que $[P_\tau]^2 = \mathbb{1}$, tem-se que P_τ é, também, unitário para qualquer transposição $\tau \in \mathcal{S}_N$. A hermiticidade também garante que os autovalores de P_τ sejam reais. Isto, aliado novamente com $[P_\tau]^2 = \mathbb{1}$, garante que os autovalores possíveis de P_τ são ± 1 .

Assim, dado um autoket $|\psi\rangle$ de P_τ , para algum $\tau \in \mathcal{S}_N$, tem-se:

$$\begin{aligned} P_\tau |\psi\rangle &= +1 |\psi\rangle, \\ P_\tau |\psi\rangle &= -1 |\psi\rangle. \end{aligned}$$

Kets com autovalor $+1$ são ditos **simétricos** e kets com autovalor -1 são ditos **antisimétricos**. Esta propriedade é chamada da paridade do ket em questão. É importante notar que estes valores ± 1 não estão relacionados com o sinal ou paridade de uma transposição, em particular porque qualquer transposição é ímpar. Ao invés disto, trata-se de uma propriedade do ket, ou mais especificamente, uma propriedade das partículas em questão. Partículas fermiônicas estão associadas ao autovalor -1 e partículas bosônicas ao autovalor $+1$. Falaremos mais sobre isso adiante. Além disso, note que isto só se aplica aos autokets de P_τ pois existem kets que não possuem uma paridade bem definida.

Voltamos a atenção agora para o caso de uma permutação arbitrária, σ . Como os operadores P_σ podem ser escritos como um produto de operadores P_τ onde τ é uma transposição, os operadores P_σ são unitários por serem um produto de operadores unitários. Apesar disto, não são, em geral, hermitianos pois as transposições não necessariamente comutam. Em particular, duas transposições que trocam elementos em comum não comutam. Para ver isto, considere $\tau_1 = (ab)$ e $\tau_2 = (bc)$ e a atuação dos operadores: P_{τ_1} e P_{τ_2} em um ket: $|a : u_i, b : u_j, c : u_k, \dots\rangle$:

$$\begin{aligned} P_{\tau_1} P_{\tau_2} |a : u_i, b : u_j, c : u_k, \dots\rangle &= P_{\tau_1} |a : u_i, c : u_j, b : u_k, \dots\rangle \\ &= |b : u_i, c : u_j, a : u_k, \dots\rangle, \end{aligned}$$

enquanto a aplicação na ordem contrária fornece:

$$\begin{aligned} P_{\tau_2} P_{\tau_1} |a : u_i, b : u_j, c : u_k, \dots\rangle &= P_{\tau_2} |b : u_i, a : u_j, c : u_k, \dots\rangle \\ &= |c : u_i, a : u_j, b : u_k, \dots\rangle. \end{aligned}$$

Todas as outras transposições, isto é, transposições disjuntas, geram operadores P_τ que comutam entre si.

Deste modo, os operadores de permutação além de não serem hermitianos, em geral também não comutam entre si. Portanto, não é possível, na maioria dos casos encontrar uma base formada pelos autovalores comuns dos operadores de permutação. No entanto, existem alguns kets, assim como no caso das transposições, que são autokets dos operadores de permutação. Estamos particularmente interessados em kets que são autokets simultâneos de todos os operadores de permutação. Chamaremos a propriedade de ser um autoket simultâneo de todos os operadores de permutação de **simetricidade**. Esta palavra é um neologismo significando a qualidade de ser simétrico em relação a alguma propriedade, aqui usaremos como a propriedade de ser simétrica sob permutações.

Existem dois tipos de kets simétricos satisfazendo, respectivamente a uma das duas propriedades a seguir:

- Dado $|\psi\rangle \in \mathcal{H}$, se

$$P_\sigma |\psi\rangle = +1 |\psi\rangle$$

para toda permutação $\sigma \in \mathcal{S}_N$, este ket é dito **completamente simétrico**.

- E, se um $|\psi\rangle \in \mathcal{H}$ satisfaz

$$P_\sigma |\psi\rangle = \text{sgn}(\sigma) |\psi\rangle$$

para toda permutação $\sigma \in \mathcal{S}_N$, este ket é dito **completamente antissimétrico**.

Aqui existe um inconveniente da linguagem: um ket simétrico pode se referir tanto à propriedade de simetricidade, isto é, ser um autoket simultâneo de todos os operadores de permutação ou à propriedade de ser completamente simétrico como enunciada acima. Em geral tentaremos usar a palavra “simetricidade” para tratar do primeiro caso e “simétrico” para o segundo.

Note a aparição do sinal da permutação no caso de kets completamente antissimétricos. Isto se dá porque a condição para ser completamente antissimétrico exige, em particular que o ket seja antissimétrico para todas as transposições, já que transposições são permutações ímpares. Assim, para uma permutação arbitrária, basta decompô-la em transposições e aplicar sucessivamente os operadores correspondentes, cada um dos quais irá contribuir com um fator de -1 . Isto corresponde precisamente ao método de determinar a paridade de uma permutação.

Note que para que um ket seja completamente simétrico ou antissimétrico é necessário que seja um ket simultâneo de todas as transposições, ou, mais especificamente, de um conjunto de transposições capaz de gerar o grupo simétrico de grau N . No entanto esta condição é apenas necessária. Além disso, o autovalor do ket deve ser o mesmo para todas as transposições. Assim, um ket que é simétrico para algumas transposições e antissimétrico para outras não será nem completamente simétrico nem completamente antissimétrico. Por adequação à literatura, diremos, então, que estes kets também não satisfazem à propriedade de simetricidade que definimos.

2.1.2 - Postulado de simetridade

Finalmente de posse dos resultados sobre permutações podemos finalizar a discussão sobre o problema da degenerescência de troca. Em essência, esta degenerescência surge porque existem múltiplos kets possíveis, matematicamente, que descrevem o mesmo estado físico. Como vimos, se $|\psi\rangle$ é um ket que descreve o estado de um sistema de N partículas, o ket $P_\sigma|\psi\rangle$ para qualquer permutação $\sigma \in \mathcal{S}_N$ também o é. Além disso, qualquer combinação linear de kets $P_\sigma|\psi\rangle$ também é uma escolha possível para descrever o sistema. Ou seja, existe um subespaço de kets possíveis correspondentes ao mesmo estado físico. Além disso, o resultado de certas medições pode depender de qual ket é usado de modo que não se pode simplesmente escolher um ket do subespaço.

No entanto, pela definição, se um ket é completamente simétrico ou antissimétrico tem-se:

$$P_\sigma|\psi\rangle = \begin{cases} +1|\psi\rangle & \text{para } |\psi\rangle \text{ completamente simétrico;} \\ \text{sgn}(\sigma)|\psi\rangle & \text{para } |\psi\rangle \text{ completamente antissimétrico.} \end{cases}$$

Ou seja, se um ket satisfaz à propriedade de simetridade, ele é único a menos de uma constante multiplicativa, isto é, o subespaço que daria origem à degenerescência de troca é unidimensional e todos os kets possíveis são colineares. Como neste caso o fator de fase é inessencial, podemos simplesmente eliminá-lo e afirmar que o ket em questão é aquele que representa adequadamente o estado do sistema. Denominamos este ket de **ket físico**. Este fato sustenta o seguinte postulado, ao qual atribuímos o nome de **postulado de simetridade**:

Postulado de simetridade

Em um sistema de N partículas idênticas e indistinguíveis, o ket que descreve o sistema, dito **ket físico** deve ser um ket simétrico. Isto é deve ser ou **completamente simétrico** ou **completamente antissimétrico**. Partículas para as quais o ket físico é completamente simétrico são chamadas **bósons** e aquelas para as quais o ket físico é completamente antissimétrico, **férmions**.

Este postulado possui implicações que chegam aos cantos mais longínquos da física. Destas, poder-se-ia dizer que a mais importante é o princípio da exclusão de Pauli, obedecido por partículas fermiônicas. É um exercício canônico de mecânica quântica mostrar que a antissimetria dos kets descrevendo partículas fermiônicas dá origem a uma repulsão efetiva quando estas partículas possuem todos os números quânticos idênticos. Ou seja, dá origem ao princípio de exclusão de Pauli. Além disso, o comportamento estatístico de partículas idênticas é governado majoritariamente por sua natureza, fermiônica ou bosônica, o que ajuda a explicar fenômenos como a supercondutividade de baixas tempe-

raturas e condensação de Bose-Einstein, por exemplo. É possível até mesmo relacionar a natureza da partícula, isto é, se ela é um férmion ou um bóson com seu spin por meio do **teorema spin-estatística**, um resultado da teoria quântica de campos. Consequentemente relacionando-se o spin à propriedade de simetridade imposta sobre a função de onda que descreve um sistema de partículas idênticas e indistinguíveis.

Apesar disto, o postulado, ou mais precisamente, as justificativas teóricas do postulado de simetridade, além de todo o tratamento de partículas indistinguíveis exposto anteriormente pode ser criticado, apesar desta ser a apresentação usual presente em livros didáticos. As primeiras críticas a este tratamento foram feitas por volta da década de 1970. [2, 23] Em particular, um dos efeitos destas críticas foi revelar que o postulado não se aplica a partículas em duas ou uma dimensão, apesar de não haver nenhuma menção explícita à dimensionalidade em todo o desenvolvimento mostrado para justificar o postulado. Na década de 1980, começaram a surgir evidências experimentais oriundas de estudos acerca do efeito Hall quântico que sustentam esta afirmação. [10, 9]

Isto significa, portanto, que existem hipóteses ocultas em algum dos argumentos utilizados para justificar o postulado de simetridade. Afinal, amplas evidências experimentais indicam que o postulado em si está correto. Estas críticas serão reproduzidas e discutidas na seção seguinte. Por completude, finalizaremos esta seção com o modo de construir os kets físicos usando dos resultados acerca de permutações obtidos. Veremos também que a crítica do tratamento apresentado até agora (re)interpreta este método sobre um ótica diferente e, possivelmente, mais natural.

Como o postulado estabelece que o ket físico deve satisfazer à propriedade de simetridade, a construção deste ket se resume em obter um ket com simetridade bem definida a partir de um dos inúmeros kets possíveis. Para isto definiremos dois operadores cuja função é produzir um ket que seja completamente simétrico ou completamente antissimétrico a partir de um ket arbitrário. Estes operadores são chamados **simetrizador** e **antissimetrizador** respectivamente e são definidos por:

$$S = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} P_{\sigma};$$

$$A = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \text{sgn}(\sigma) P_{\sigma}.$$

S é o simetrizador e A é o antissimetrizador.

Note que a ação destes operadores é formar uma combinação linear específica a partir de todas as permutações de índices de partícula do ket original. Estes dois operadores são, na verdade projetores, retornando a parte do ket original que é ou completamente simétrica ou completamente antissimétrica. Para ver isto considere primeiro o seguinte resultado:

Proposição 2.4: Para qualquer permutação $\sigma \in \mathcal{S}_N$, tem-se:

$$\begin{aligned} P_\sigma \mathbf{S} &= \mathbf{S} = \mathbf{S} P_\sigma; \\ P_\sigma \mathbf{A} &= \text{sgn}(\sigma) \mathbf{A} = \mathbf{A} P_\sigma. \end{aligned}$$

Prova: Primeiro, note que para quaisquer duas permutações $\sigma, \tau \in \mathcal{S}_N$, $P_\sigma P_\tau$ é também uma permutação. Isto se dá pois as permutações formam um grupo, de modo que $\sigma\tau = \rho$ para algum $\rho \in \mathcal{S}_N$. Pela definição dos operadores permutação, então, tem-se:

$$P_\sigma P_\tau = \mathbf{R}(\sigma)\mathbf{R}(\tau) = \mathbf{R}(\sigma\tau) = \mathbf{R}(\rho) = P_\rho$$

Em particular este resultado junto à proposição 2.2 garante que os operadores P_σ formem um grupo. O mapa \mathbf{R} é, na verdade um homomorfismo de grupos. Com este resultando, tem-se que:

$$P_\sigma \mathbf{S} = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} P_\sigma P_\tau = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} P_{\sigma\tau}.$$

Como $\sigma\tau = \rho$ para algum $\rho \in \mathcal{S}_N$, tem-se que $\tau = \sigma^{-1}\rho$, assim, como σ é fixo, conforme τ percorre todos os elementos de \mathcal{S}_N ρ também o faz, possivelmente em uma ordem distinta, de modo que podemos reescrever o somatório como:

$$P_\sigma \mathbf{S} = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} P_{\sigma\tau} = \frac{1}{N!} \sum_{\rho \in \mathcal{S}_N} P_\rho = \mathbf{S}.$$

A prova de que $\mathbf{S} P_\sigma = \mathbf{S}$ é análoga.

A prova para o antissimetrizador é análoga com a adição de um detalhe relativo ao sinal das permutações. Tem-se:

$$P_\sigma \mathbf{A} = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) P_\sigma P_\tau = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) P_{\sigma\tau}.$$

Usando o mesmo argumento podemos reescrever o somatório. Além disso, pela definição do sinal de uma permutação, tem-se que $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$, logo:

$$\begin{aligned} P_\sigma \mathbf{A} &= \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) P_{\sigma\tau} = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\sigma^{-1}\rho) P_\rho \\ &= \text{sgn}(\sigma^{-1}) \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\rho) P_\rho \\ &= \text{sgn}(\sigma^{-1}) \mathbf{A}. \end{aligned}$$

Por fim, como $sgn(\sigma^{-1}) = sgn(\sigma)$, pode-se ver isto pela condição de que $sgn(\sigma^{-1}\sigma) = sgn(\mathbf{1})$, obtêm-se:

$$P_\sigma \mathbf{A} = sgn(\sigma) \mathbf{A}.$$

A prova para $\mathbf{A}P_\sigma$ é análoga.

□

Este resultado prova que a aplicação do simetrizador ou do antissimetrizador gera um ket que é, respectivamente, completamente simétrico ou completamente antissimétrico. Por fim, para mostrar que eles são projetores basta lembrar que o grupo \mathcal{S}_N possui $N!$ elementos e considerar:

$$\begin{aligned} \mathbf{S}^2 &= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} P_\sigma \mathbf{S} \\ &= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \mathbf{S} \\ &= \mathbf{S}. \end{aligned}$$

E:

$$\begin{aligned} \mathbf{A}^2 &= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} sgn(\sigma) P_\sigma \mathbf{A} \\ &= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} [sgn(\sigma)]^2 \mathbf{A} \\ &= \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \mathbf{A} \\ &= \mathbf{A}. \end{aligned}$$

Note que $[sgn(\sigma)]^2 = 1$ pois o sinal de uma permutação só admite os valores ± 1 . Assim, fica demonstrado que os operadores \mathbf{S} e \mathbf{A} são projetores em subespaços de kets completamente simétricos e completamente antissimétricos, respectivamente. Deste modo, a obtenção do ket físico de um sistema é descrita nos três passos a seguir:

- i Escolha uma numeração arbitrária para as partículas.
- ii Aplique o \mathbf{S} ou \mathbf{A} a depender da natureza da partícula. \mathbf{S} para bósons e \mathbf{A} para férmions.
- iii Normalize o ket resultante.

Com isto finalizamos a exposição sobre o tratamento de partículas quânticas idênticas e indistinguíveis como apresentado em livros-texto clássicos sobre o assunto. Na seção

que segue discutiremos uma crítica a este tratamento e exploraremos um tratamento alternativo que revelará algumas falhas ou omissões na formulação apresentada.

2.2 - Crítica da indistinguibilidade

Nesta seção será feita uma revisão da discussão feita nas seções anteriores com o objetivo de explicitar e retificar algumas falhas. Em particular, iniciamos com uma crítica dos conceitos centrais do tratamento de partículas idênticas: identidade, indistinguibilidade e permutação. Muitas vezes estes conceitos, especialmente o de permutação de partículas são mal traduzidos matematicamente no formalismo da teoria além de muitas vezes serem usados de modo a não possuírem significado físico ou experimental ou as vezes, sem nenhuma carga semântica.

A crítica apresentada aqui é construída, inicialmente a partir daquela apresentada por R. Mirman em seu artigo cujo título, em tradução livre é: “significado experimental do conceito de partículas idênticas”.^[23] Em essência, a crítica de Mirman surge a partir do julgamento que as justificações teóricas do postulado de simetridade são insuficientes e o ponto principal é que conceitos relacionados ao tratamento de partículas idênticas, em particular o conceito de indistinguibilidade são “mal formulados, mal relacionados com os resultados experimentais e frequentemente usados de modo que são experimentalmente e as vezes, semanticamente, vazios de significado”.

Notamos aqui que Mirman diz, no artigo, que apesar de achar várias das análises feitas sobre o assunto insuficientes, ele não discorda dos resultados. Por exemplo, afirma nunca ter encontrado um argumento que, utilizando somente mecânica quântica não relativística, justifique de maneira satisfatória que o estado de um sistema de dois elétrons, inicialmente completamente antissimétrico permaneça completamente antissimétrico para sempre. Aqui introduz-se outro ponto da crítica: muitos dos resultados não podem possivelmente estarem errados devido ou a definições demasiadamente elásticas ou por decorrerem, também, de resultados de teoria de campo relativísticas.

Iniciamos, então, por considerar a seguinte situação: em uma galáxia existe um elétron e em outra, um pión. Em um dado instante, o pión decai produzindo um elétron. Considerando-se a função de onda inicial do elétron e do pión, vistos como um único sistema, ela não possui um estado de simetria definido, pois o elétron é um férmion e o pión, um bóson. Note que a simetridade não é exigida pois tratam-se de partículas não idênticas. Após o decaimento, uma aplicação ingênua do postulado de simetridade nos informa que a função de onda deve, agora, ser totalmente antissimétrica. A pergunta, então é: dado que a função de onda original não era totalmente antissimétrica, como os elétrons, tão distantes afetam um ao outro do modo a produzir este estado de simetridade definida? Ou, como o estado de simetridade surge conforme as funções de onda dos dois elétrons se misturam?

Na discussão das seções anteriores não abordamos como a simetridade da função de onda evolui no tempo e, a primeira vista, este problema parece ser desta natureza. Mirman chama este problema de um “aparente paradoxo”. Como indicado pela palavra “aparente”, este problema não gera um paradoxo em vista de uma sutileza aqui presente. Uma que Mirman aponta muito enfaticamente em seu artigo. Isto é, estes dois elétrons **não** são indistinguíveis.

Se relembarmos a definição de partículas idênticas que usamos neste trabalho, veremos que ela informa apenas que duas partículas são idênticas se não é possível diferenciar suas propriedades intrínsecas. Por exemplo, não há, no momento, nenhum experimento que possa afirmar que a massa de um elétron é diferente da de outro, ou seu spin ou sua carga. Isto pode parecer estranho, mas se pensarmos em objetos do dia a dia, consideramos vários objetos idênticos mas que na verdade possuem discrepâncias. Por exemplo, se comprarmos duas canetas azuis da mesma marca, antes de usá-las podemos considerá-las idênticas, no entanto, é plenamente possível que as massas das canetas sejam ligeiramente diferentes. Bem, note, agora, que a definição diz respeito apenas às propriedades intrínsecas das partículas e não exclui a possibilidade de existirem propriedades extrínsecas que podem ser usadas para distingui-las.

Como discutido no exemplo sobre a colisão de partículas, se as funções de onda de cada partícula estão confinadas a regiões disjuntas, então uma medida da posição das partículas nos permite distingui-las. Podemos chamar de partícula (1) aquela na região \mathcal{R}_1 e de (2), aquela em \mathcal{R}_2 . Enquanto estas regiões permanecerem disjuntas, estas partículas poderão ser distinguidas. Este é justamente o caso dos dois elétrons no problema acima. Dada a ordem de grandeza de distâncias intergalácticas é bastante razoável supor que a função de onda de cada elétron não se estenda além da galáxia local. Em termos mais gerais, podemos dizer que, enquanto a informação de que um novo elétron surgiu do decaimento do pión não alcançar o outro, ambos podem ser distinguidos, ao menos em princípio.

Frequentemente, no entanto, a distinção entre os conceitos de indistinguibilidade e identidade de partículas é mencionada apenas de modo passageiro, quando é mencionada. Alguns livros apresentam os dois conceitos como sendo o mesmo [20, 24]. Esta última situação parece estar mais alinhada com o contexto do artigo, quando foi escrito em 1973. Veremos, no entanto que esta distinção nos permite unificar o tratamento de partículas idênticas com aquele de partículas não idênticas e afrouxar um pouco a rigidez com que se trata a indistinguibilidade de partículas quânticas.

Por exemplo, Mirman menciona que o requerimento de que um termo de interferência na seção de choque de um processo de espalhamento de partículas idênticas — na verdade indistinguíveis — é tido como uma hipótese extra universalmente aceita. Investiguemos esta afirmação em vista do que foi discutido até agora. Se as partículas forem indistinguíveis, segundo o postulado de simetridade a função deve ser ou totalmente simétrica ou totalmente antissimétrica. Suponhamos então que após um experimento de espalha-

mento de duas partículas idênticas, uma foi detectada na posição x e outra na posição y . O estado final, segundo o postulado é:

$$|1 : x, 2 : y\rangle \pm |1 : y, 2 : x\rangle$$

a depender se as partículas são bosônicas ou fermiônicas. A probabilidade, P , de encontrarmos este resultado, então é:

$$P = ||1 : x, 2 : y\rangle|^2 + ||1 : y, 2 : x\rangle|^2 \pm 2\Re[\langle 1 : x, 2 : y | 1 : y, 2 : x \rangle].$$

O termo de interferência está presente, como esperado segundo esta dita hipótese extra. Vejamos o que acontece no caso onde as funções de onda das partículas (1) e (2) são diferentes de zero apenas se $x \in \mathcal{R}_1$ e $y \in \mathcal{R}_2$ respectivamente e $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$. Neste caso, escrevendo o termo de interferência de forma explícita:

$$\Re[\langle 1 : x, 2 : y | 1 : y, 2 : x \rangle] = \Re[\langle 1 : x | 1 : y \rangle \langle 2 : y | 2 : x \rangle]$$

pode-se ver que ele se anula, pois $|1 : y\rangle$ e $|2 : x\rangle$ são nulas. A probabilidade então é:

$$P = ||1 : x, 2 : y\rangle|^2 + ||1 : y, 2 : x\rangle|^2.$$

Que é precisamente a probabilidade em um espalhamento de duas partículas não idênticas se nenhuma tentativa de distingui-las é feita durante o experimento [25]. O segundo caso pode ser questionado com base no fato de que após o espalhamento as duas partículas não podem mais ser distinguidas. Mirman contesta esta afirmação. Antes disso, gostaríamos apenas de apontar que este resultado é válido fora do contexto de espalhamento. Por exemplo, o estado em questão poderia ser o estado descrevendo duas partículas **antes** do espalhamento, por exemplo. Como vimos, a degenerescência de troca existe também ao determinar o estado inicial de um sistema de partículas idênticas.

De todo modo, Mirman afirma que mesmo após um espalhamento, duas partículas idênticas não são completamente indistinguíveis, apesar de não podermos seguir suas trajetórias. Usualmente, a posição onde as partículas são detectadas é descrita em termos de ângulos. Em um caso com simetria azimutal, podemos escrever o estado acima como:

$$|1 : \theta, 2 : \pi - \theta\rangle \pm |1 : \pi - \theta, 2 : \theta\rangle.$$

Esta função de onda implica que uma partícula possui igual probabilidade de ser detectada a um ângulo de 0 ou π . Isto não é necessariamente verdade. No caso do espalhamento Coulomb, por exemplo, a amplitude de espalhamento é proporcional a $\sin^{-2}(\theta/2)$ o que fornece uma probabilidade muito maior para proespalhamento do que para retroespalhamento. Assim, uma partícula espalhada a pequenos ângulos é muito provavelmente

a partícula incidente. Notamos que em princípio isto não removeria a indistinguibilidade das partículas a menos que considerássemos que para um dado ângulo, ou intervalo angular, a amplitude de espalhamento vá a zero.

Analisemos agora o caso de espalhamento de duas partículas distintas, a partícula incidente i e a partícula alvo, t . Segundo os princípios da mecânica quântica, todas as possibilidades devem ser levadas em conta, então, o ket representando o estado final do espalhamento é:

$$|i : \theta, t : \pi - \theta\rangle + e^{i\phi} |i : \pi - \theta, t : \theta\rangle.$$

onde ϕ é um fator de fase arbitrário. Por simplicidade consideraremos $\phi = 0$. A probabilidade de detectarmos as partículas nesta configuração, então é:

$$||i : \theta, t : \pi - \theta\rangle|^2 + ||i : \pi - \theta, t : \theta\rangle|^2 + 2\Re[\langle i : \theta, t : \pi - \theta | i : \pi - \theta, t : \theta \rangle].$$

Note que o termo de interferência está presente. Então observamos que existem situações onde é teoricamente possível que o termo de interferência não esteja presente em espalhamento de partículas idênticas e que, segundo os princípios da mecânica quântica, o termo de espalhamento está presente para o caso de partículas não idênticas. Assim, podemos concluir que ao invés de tomar a presença do termo de espalhamento como uma hipótese extra, ela decorre dos princípios básicos da mecânica quântica e a pergunta que deve ser respondida é: por que o termo de interferência não está presente em espalhamento de partículas não idênticas?

Não entraremos a fundo nesta discussão e referimos ao artigo de Mirman, [23] e referências lá presentes. Porém, ele afirma que uma resposta possível é que os estados para partículas distintas são ortogonais. Em geral isto parece ser verdade, mas são mencionados casos onde isto falha. Em particular, mesmo para partículas não idênticas, estados diferentes não necessariamente são ortogonais e o termo de interferência pode estar presente.

Agora que tratamos da nuance cercando os conceitos de igualdade e indistinguibilidade de partículas, vamos à segunda parte, tratando do conceito de permutação e sua tradução matemática. Como vimos, o conceito de indistinguibilidade junto ao postulado de simetridade impõe a seguinte condição sobre os kets que descrevem o sistema:

$$|P_\sigma|\psi\rangle|^2 = ||\psi\rangle|^2$$

onde $|\psi\rangle$ é um ket descrevendo um sistema de N partículas indistinguíveis e $\sigma \in \mathcal{S}_N$. De modo mais explícito, usando os índices de partículas, tem-se:

$$||\sigma(1) : u_i, \sigma(2) : u_j, \dots, \sigma(N) : u_k\rangle|^2 = ||1 : u_i, 2 : u_j, \dots, N : u_k\rangle|^2. \quad (2.3)$$

Mirman afirma que a indistinguibilidade das partículas, ao invés de exigir esta identidade, não só esvazia-a de sentido experimental como torna a sua carga semântica nula. De fato, aceita-se que a indistinguibilidade de partículas requer que a situação permaneça inalterada caso duas partículas sejam trocadas de lugar. Além disso, é verdade que o estado do sistema pode ser descrito por um ket cujo valor absoluto é invariante sob permutações de índices de partículas. No entanto, a afirmação é que: o estado físico, não a função de onda, permanece inalterado após a permutação de partículas, não índices. Em essência, a origem do problema aqui é que a indistinguibilidade das partículas impede de atribuir qualquer significado físico aos índices. Isto, por sua vez, faz com que a operação de permutação não tenha nenhum significado físico.

Para evidenciar ainda mais a situação, considere um sistema composto por duas partículas não idênticas. Denominamos os tipos das partículas de t_1 e t_2 , assim, a quantidade:

$$|| t_1 : u_i, t_2 : u_j \rangle|^2$$

representa a probabilidade de encontrar uma partícula do tipo t_1 no estado u_i e uma do tipo t_2 no estado u_j . Se agora, fizermos $t_1 = t_2 = t$, podemos omitir a menção ao tipo de partícula e escrever:

$$|| 1 : u_i, 2 : u_j \rangle|^2.$$

Esta quantidade representa a probabilidade de encontrar uma partícula do tipo t no estado u_i e outra partícula do tipo t no estado u_j . Por outro lado, a quantidade:

$$|| 2 : u_i, 1 : u_j \rangle|^2$$

representa a probabilidade de encontrar uma partícula do tipo t no estado u_j e outra, também do tipo t no estado u_i .

Fica evidente, portanto que a identidade (2.3) não surge, realmente, devido a uma exigência experimental. Ela é, na verdade redundante em vista do fato da conjunção lógica, representada pelo conectivo “e”, ser comutativa. Isto é, na lógica proposicional, dadas duas proposições p e q , tem-se:

$$p \wedge q = q \wedge p.$$

Mirman resume isto como um caso onde as hipóteses — o sistema deve ser invariante sob permutação de partículas indistinguíveis — são traduzidas em uma manipulação sensata de símbolos, segundo as regras de manipulação de símbolos, mas que não podem ser enunciadas, em palavras, de nenhuma maneira que carregue significado.

Note que apesar disso a exigência de simetridade sobre a função de onda que descreve um sistema de várias partículas indistinguíveis leva a consequências observáveis, por

exemplo, descreve adequadamente níveis de energia. No entanto, esta discussão nos permite examinar o conceito de degenerescência de troca — a fonte original de dificuldades — sob uma nova ótica. A dificuldade criada pela degenerescência de troca seria ilusória, produto de um formalismo matemático inadequado que por sua vez é oriundo do uso de conceitos que não possuem nenhuma realidade física, os índices de partícula, arbitrários desde o início pois nenhuma medida pode extrair informação sobre eles.

Os símbolos $||1 : u_i, 2 : u_j \rangle|^2$ e $||2 : u_i, 1 : u_j \rangle|^2$ são distintos e aparentam representar coisas distintas mas na realidade referem-se ao mesmo objeto. Eles não nos informam nada sobre o comportamento de partículas indistinguíveis, apenas que a notação usada é redundante. O problema surge, usualmente, por serem usados como se possuíssem mais significado que realmente têm. Sobre a questão dos índices e o conceito de permutação, a crítica de Mirman termina aqui, porém, é continuada no trabalho de Leinaas e Myrheim [2] que é de particular interesse para este trabalho.

Leinaas e Myrheim, baseados no fato que o termo “permutação” como usualmente utilizado não carrega nenhum significado físico, argumentaram que um tratamento adequado de partículas indistinguíveis requer uma descrição adequada do espaço de configuração destas partículas. Eles iniciam com um tratamento clássico e então quantizam o sistema para obter a mecânica quântica de partículas indistinguíveis. Parte do motivo para começar pelo tratamento clássico é a constatação do fato que a indistinguíbilidade de partículas, ao contrário de como é muitas vezes descrita, também causa problemas no reino clássico, como constatado por Gibbs em seu famoso paradoxo.

Paradoxo de Gibbs

O paradoxo de Gibbs surge de uma então falha da mecânica estatística clássica em calcular a mudança de entropia ao se misturar dois fluidos idênticos e nas mesmas condições de temperatura, pressão e outras possíveis propriedades relevantes. A situação é representada na figura abaixo.

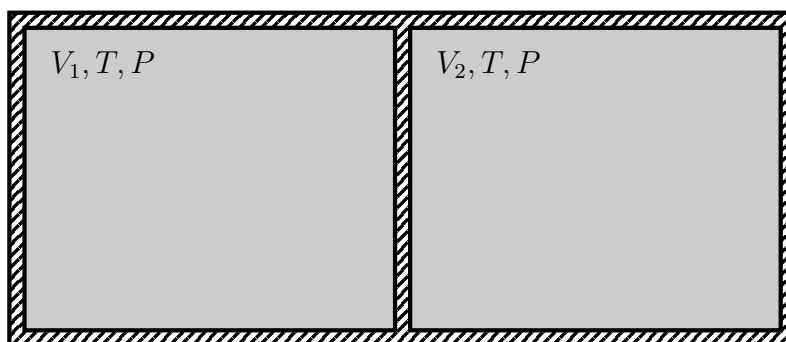


Figura 2.2: Recipiente contendo um fluido separado em dois compartimentos.

Considere um recipiente com paredes ideais sem qualquer tipo de troca com o exterior. Neste recipiente é posto um fluido, por exemplo, um gás ideal. Este fluido então é separado

em dois compartimentos por uma parede também ideal de modo que as duas partes estão à mesma temperatura e pressão. Este é o estado inicial. Agora, suponha que a parede que separa os dois compartimentos seja magicamente removida e o fluido de ambas as partes possa se misturar.

A identidade do fluido das duas partes, além do fato de ambas estarem à mesma temperatura e pressão indica que a mudança de entropia deve ser nula. Afinal, a mistura é um rearranjo, uma permutação, das moléculas que compõem o fluido. Não há nenhuma alteração na energia. Moléculas que estavam no compartimento 2 poderiam muito bem ter estado no compartimento 1 e vice-versa quando a parede primeiro foi colocada.

No entanto, a mecânica estatística da época fornecia um resultado diferente. A diferença de entropia entre o estado final e o inicial era não nula. Gibbs então postulou que estados que diferiam apenas por uma permutação de partículas idênticas — para a teoria clássica, partículas idênticas são sempre distinguíveis — não deveriam ser contados como distintos. Assim, se um sistema é composto de N partículas idênticas, a contagem usual de estados contaria o mesmo estado $N!$ vezes, fazendo-se necessário dividir o número de estados por $N!$ para corrigir a contagem.

O paradoxo de Gibbs é um problema clássico presente nos livros-texto de mecânica estatística. Não o exploraremos mais a fundo do que isto aqui. O propósito de apresentá-lo é apontar que a identidade de partículas pode causar efeitos até mesmo na mecânica clássica onde é sempre possível acompanhar a trajetória das partículas de modo a distingui-las. Argumentamos que em vista do tratamento estatístico usado estas partículas idênticas se tornam efetivamente indistinguíveis. Nenhuma tentativa de distingui-las é feita pela mecânica estatística. De fato, o contrário é verdade, a mecânica estatística abre mão da individualidade de cada partícula para tratá-las como grupos. Por exemplo, não importam quais partículas estão em uma dada região, apenas seu número médio, velocidade média e assim por diante.

2.2.1 - Espaço de configurações

O paradoxo de Gibbs nos chama atenção em particular para o espaço de configurações do sistema, dado que é um dos objetos fundamentais da mecânica estatística. A identificação de configurações que podem ser obtidas a partir de outra por uma permutação dos índices de partícula, ao que indica, faz com que o espaço de configurações de partículas idênticas seja distinto daquele de partículas não idênticas. Exploremos, portanto, esta diferença. Na discussão que segue serão usados conceitos básicos de topologia geral e topologia algébrica, em particular os conceitos de espaços quocientes, homotopia e grupo fundamental. Caso não esteja familiarizado com estes conceitos, é recomendado que leia os apêndices A e B antes de prosseguir.

Leinaas e Myrheim começam por estudar o espaço de configurações de partículas

clássicas e então realizam um processo de quantização, argumentando que a descrição adequada deste espaço leva, naturalmente ao postulado de simetricidade após a quantização. Usualmente, descreve-se o espaço de configurações de um sistema de partículas idênticas como o produto cartesiano do espaço de configurações de uma única partícula. Mas, como vimos, este não é o espaço mais adequado. Devemos identificar configurações que podem ser obtidas a partir de uma permutação.

Considere então o grupo simétrico de ordem n , \mathcal{S}_n e o espaço de configurações X^n de um sistema de n partículas idênticas. Denotamos por $x = (x_1, \dots, x_n)$ os elementos de X^n . Agora, defina uma operação $\star : \mathcal{S}_n \times X^n \rightarrow X^n$ por:

$$\sigma \star x = (x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Esta operação define a ação à esquerda de \mathcal{S}_n sobre X^n . De fato, tem-se:

$$\mathbf{1} \star x = x$$

para todo $x \in X^n$, além disso, para quaisquer $\sigma, \tau \in \mathcal{S}_n$ e $x \in X^n$, tem-se

$$\begin{aligned} \tau \star (\sigma \star x) &= \tau \star (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \\ &= (x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(n))}) \\ &= (x_{(\tau\sigma)(1)}, \dots, x_{(\tau\sigma)(n)}) = (\tau\sigma) \star x. \end{aligned}$$

A partir da operação \star , definimos a seguinte relação, \sim , em X^n :

$$x \sim y \iff \exists \sigma. y = \sigma \star x.$$

Esta relação será bastante útil para construir o espaço de configurações adequado para um sistema de n partículas idênticas. O resultado a seguir fundamenta toda a construção.

Lema 2.1: A relação \sim definida acima é uma relação de equivalência.

Prova: Devemos mostrar que \sim satisfaz à todas as propriedades exigidas.

- **Reflexividade:** para todo x , tem-se que $x \sim x$ pois $x = \mathbf{1} \star x$, onde $\mathbf{1}$ é o elemento identidade de \mathcal{S}_n .
- **Simetricidade:** se $x \sim y$, então existe $\sigma \in \mathcal{S}_n$ tal que $y = \sigma \star x$. Aplicando a ação de σ^{-1} aos dois lados obtém-se:

$$\sigma^{-1} \star y = (\sigma^{-1}\sigma) \star x = x.$$

Logo, $y \sim x$.

- **Transitividade:** se $x \sim y$ e $y \sim z$ então existem $\sigma, \tau \in \mathcal{S}_n$ tais que $y = \sigma \star x$ e $z = \tau \star y$. Substituindo-se y na segunda expressão, tem-se:

$$z = \tau \star (\sigma \star x) = (\sigma\tau) \star x.$$

Portanto, $x \sim z$.

□

Como \sim é uma relação de equivalência, suas classes de equivalência particionam o conjunto X^n . Denotamos por $[x]$ a classe de equivalência de x , isto é, o conjunto:

$$[x] = \{y \in X^n \mid y = \sigma \star x, \sigma \in \mathcal{S}_n\}.$$

E denotamos por X^n/\sim o conjunto de todas tais classes de equivalência. Como a relação \sim é oriunda da ação de \mathcal{S}_n , também podemos denotar este conjunto por X^n/\mathcal{S}_n . Em geral, se \mathcal{G} é um grupo, sua ação sobre um conjunto X induz uma relação de equivalência $\sim_{\mathcal{G}}$ em X definida de forma análoga a \sim . Isto é, se \star é a ação de \mathcal{G} sobre X , dados $x, y \in X$, definimos

$$x \sim_{\mathcal{G}} y \iff \exists \sigma. y = \sigma \star x.$$

onde $\sigma \in \mathcal{G}$. O lema anterior mostra que $\sim_{\mathcal{G}}$ é uma relação de equivalência. Para ajudar no entendimento nos ateremos à última notação neste trabalho e nomearemos o conjunto das classes de equivalência de X sob $\sim_{\mathcal{G}}$ por X/\mathcal{G} . Assim, fica sempre subentendido que a relação de equivalência é aquela induzida pela ação de \mathcal{G} em X .

Continuando, X^n/\mathcal{S}_n é o conjunto adequado para descrever o espaço de configurações de um sistema de n partículas idênticas. notamos no entanto que ele descreve as configurações de partículas imateriais, no sentido de que, é possível que duas ou mais partículas ocupem a mesma posição ao mesmo tempo. Para evitar este problema, subtrair de X^n o conjunto Δ , dado por:

$$\Delta = \{x \in X^n \mid x_i = x_j \text{ para algum } i \neq j\}.$$

O conjunto Δ é usualmente chamado de diagonal generalizada. Estabelecamos agora uma notação conveniente. Denotaremos por \mathcal{C}_n^d o espaço de configuração de n partículas em dimensão d . Escreveremos apenas \mathcal{C}^d quando o número de partículas estiver subentendido e \mathcal{C}_n quando a dimensionalidade o estiver. O espaço de configuração para n partículas de núcleo duro — que não podem atravessar uma a outra — e idênticas, independente da dimensão, então é $\mathcal{C}_n/\mathcal{S}_n$. Note que a prova que \star e \sim são uma ação a esquerda de \mathcal{S}_n e

uma relação de equivalência em \mathcal{C}_n é análoga.

Como descrito no apêndice A, $\mathcal{C}_n/\mathcal{S}_n$ pode ser transformado em um espaço quociente se for equipado com a topologia quociente obtida a partir da sobrejeção canônica:

$$p : \mathcal{C}_n \rightarrow \mathcal{C}_n/\mathcal{S}_n \\ x \mapsto [x].$$

Este é um detalhe importante pois, em geral, as propriedades topológicas de $\mathcal{C}_n/\mathcal{S}_n$ são diferentes das de \mathcal{C}_n . Leinaas e Myrheim afirmam que o fato do espaço de configurações adequado ser $\mathcal{C}_n/\mathcal{S}_n$ é frequentemente ignorado, ou inconsciente ou deliberadamente. No caso de partículas clássicas, isto não gera problemas pois a evolução das partículas dependem apenas de propriedades locais do espaço de configurações. Em particular, a trajetória do sistema é uma curva contínua em $\mathcal{C}_n/\mathcal{S}_n$. A esta curva, correspondem $n!$ curvas contínuas em \mathcal{C}_n . Em geral estas $n!$ curvas em \mathcal{C}_n não se interceptam e mesmo que isto ocorra, correspondendo a uma colisão de partículas, a condição de que as derivadas sejam contínuas nos permite escolher uma única curva em \mathcal{C}_n e chamá-la de trajetória do sistema. Essencialmente, é possível seguir as partículas ao longo da colisão mesmo que sejam idênticas.

No caso quântico, no entanto, as propriedades globais do espaço de configurações se tornam importantes. Em particular, o invariante topológico mais importante é o grupo fundamental do espaço de configurações. Primeiro, exploraremos o porquê do grupo fundamental ser importante para esta discussão e em seguida exploraremos alguns exemplos de espaços de configurações de sistemas de partículas idênticas e seus grupos fundamentais.

Recorde que Mirman [23] apontou que o termo “permutação” é frequentemente usado sem nenhum significado físico, pois se refere à permutação dos índices de partículas e não das partículas em si. Para corrigir isto, então, devemos permutar fisicamente as partículas, isto é, mover uma partícula, 1 para a posição da partícula 2 e vice-versa. Para evitar quaisquer outros efeitos desta movimentação no sistema, assumimos que o processo é feito adiabaticamente. Primeiro, suponha que ao rotacionarmos uma partícula em relação a outra, a função de onda do sistema adquira uma fase proporcional ao ângulo de rotação. Por exemplo, o seguinte processo:

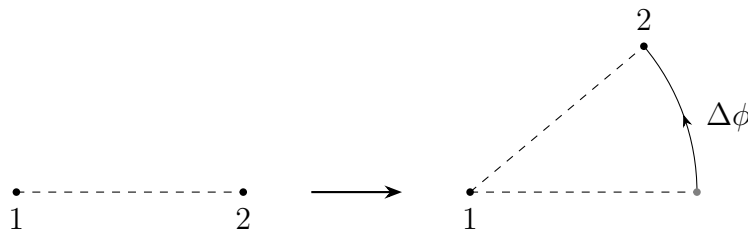


Figura 2.3: Rotação adiabática de uma partícula em torno de outra.

Faria com que a função de onda deste sistema de duas partículas idênticas se transformasse do seguinte modo:

$$\psi(1, 2) \rightarrow \psi'(1, 2) = e^{i\nu\Delta\phi}\psi(1, 2).$$

O parâmetro ν é usualmente chamado de **estatística** destas partículas. Seu significado completo, no entanto, só é perceptível quando permutamos as duas partículas. A permutação pode ser feita de dois modos distintos, rotacionando-se uma partícula por um $\Delta\phi = \pi$ ao redor da outra e então realizar uma translação rígida do centro de massa. Ou, rotacionando-se uma ao redor da outra por $\Delta\phi = -\pi$ e então realizando-se uma translação rígida do centro de massa. As duas rotações estão representadas na figura abaixo.

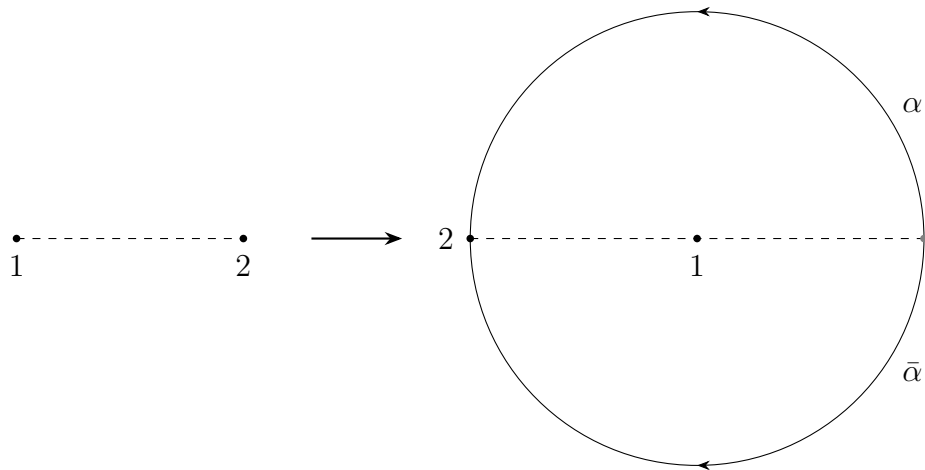


Figura 2.4: Diagrama ilustrando possíveis processos de permutação de duas partículas.

Aqui existe uma questão importante. A hipótese é de que permutar as duas partículas mantêm o sistema inalterado. O que sugere a igualdade:

$$e^{i\pi\nu} = e^{-i\pi\nu}. \quad (2.4)$$

De fato, esta igualdade vale desde que ν assumam valores $0, 1 \pmod{2}$, correspondendo a estatísticas bosônicas e fermiônicas, respectivamente. No entanto, pode-se perguntar, porque ambos os métodos de permutar as partículas são equivalentes. A resposta, é dada pelo conceito de homotopia [def. B.1]. Considerando-se a trajetória das duas partículas como curvas em \mathbb{R}^3 parametrizadas pelo tempo, tem-se que as curvas α e $\bar{\alpha}$ são caminhos homotópicos. Isto é, existe um mapa $F : I \times I \rightarrow \mathbb{R}^3$ tal que

$$\begin{aligned} F(t, 0) = \alpha(t) \quad \text{e} \quad F(t, 1) = \bar{\alpha}(t); \\ F(0, s) = x_2^i \quad \text{e} \quad F(1, s) = x_2^f. \end{aligned}$$

Onde x_2^i e x_2^f são os pontos inicial e final da partícula 2 e $I = [0, 1]$, o intervalo fechado. Poderíamos imaginar esta homotopia como simplesmente, levantar o caminho α na terceira dimensão, levá-lo até o caminho $\bar{\alpha}$ e então sobrepô-los.

O interessante aqui é que isto só vale para partículas em espaços com dimensão maior ou igual a três. Para partículas no espaço euclidiano de duas dimensões, por exemplo, não é possível levantar um dos dois caminhos e passar sobre a partícula 1. Deste modo, não existe homotopia entre estes dois caminhos e eles são inequivalentes. Isto, por sua vez implica que não há motivo suficiente para supor que a igualdade 2.4 valha. Isto significa, a princípio, que partículas em duas dimensões espaciais podem ter estatística arbitrária. Este resultado é consistente com o fato de que partículas em duas dimensões podem, também, ter spin arbitrário, dado que o grupo $\text{SO}(2)$ é abeliano, isto é, não existem relações de comutação a serem quantizadas. Dado que estas quantidades estão relacionadas pelo teorema spin-estatística em três dimensões, parece razoável dizer que sejam consistentes.

Estas partículas que vivem em duas dimensões foram chamadas, posteriormente, de anyons, aglutinação da palavra *any*, significando qualquer e o sufixo “-on”, usualmente usado no nome de partículas, por Frank Wilczek em seu artigo de 1982 [3]. O motivo, naturalmente, é por poderem ter quaisquer valores de spin e estatística. A situação é ainda mais dramática se considerarmos duas partículas idênticas confinadas a uma linha. Neste caso, é impossível permutá-las pois isto exigiria que atravessassem uma a outra e o conceito de estatística perde o sentido.

Note que, se permutássemos apenas os índices, ao invés de fisicamente alterar a posição das partículas adiabaticamente, não seria possível perceber como a dimensionalidade do espaço contribui para o conceito de estatística. Antes de passarmos a discutir anyons propriamente, vejamos alguns exemplos de espaços de configurações de sistemas de partículas idênticas e seus grupos fundamentais. Examinemos o caso de duas partículas em \mathbb{R}^2 e \mathbb{R}^3 .

Duas partículas idênticas em \mathbb{R}^2

O espaço de configurações adequado para um sistema de duas partículas idênticas é, conforme discutido anteriormente, dado por $\mathcal{C}_2/\mathcal{S}_2 = (X^2 - \Delta)/\mathcal{S}_2$, com $X^2 = \mathbb{R}^2$. Uma maneira de escrever este espaço, como argumentado por Leinaas e Myrheim [2] é primeiro decompor \mathcal{C}_2 em um produto de um espaço correspondente às configurações do centro de massa, no caso \mathbb{R}_2 e um espaço relativo, r_2^2 , que captura os graus de liberdade remanescentes. Neste caso, escrevemos os elementos dos dois espaços como:

$$\mathbb{R}_2 = \left\{ R \mid R = \frac{1}{2}(r_1 + r_2), r_1, r_2 \in \mathbb{R}^2 \right\}$$

e

$$r_2^2 = \left\{ r \mid r = r_1 - r_2, r_1 \neq r_2 \in \mathbb{R}^2 \right\}.$$

Note a condição $r_1 \neq r_2$ para garantir que as duas partículas não possam ocupar o mesmo lugar ao mesmo tempo. A princípio não há nenhum motivo em particular para excluir estes pontos, exceto possivelmente o fato de gerarem singularidades uma vez que pontos relacionados por permutações são identificados [2]. Outros argumentam que a remoção destes pontos é importante para a discussão de estatísticas não triviais [26].

De todo modo, o próximo passo é identificar os pontos que diferem apenas pela permutação segundo a relação \sim . Podemos definir a ação do grupo S_2 em $\mathcal{C}_2 = \mathbb{R}_2 \times r_2^2$ por:

$$\begin{aligned} \sigma \star (R, r) &= (\sigma \star R, \sigma \star r) \\ &= \left(\frac{1}{2}(r_{\sigma(1)} + r_{\sigma(2)}), r_{\sigma(1)} - r_{\sigma(2)} \right). \end{aligned}$$

Onde $\sigma \in S_2$, $R \in \mathbb{R}_2$ e $r \in r_2^2$. A partir desta ação do grupo, definimos a relação \sim como anteriormente. Note, no entanto que pela comutatividade da soma de \mathbb{R}^2 , tem-se que $\sigma \star R = R$ para todo $R \in \mathbb{R}_2$, ou seja, \mathbb{R}_2 não é afetado pela identificação segundo \sim . Por outro lado, dado $r \in r_2^2$, tem-se que $\sigma \star r = -r$ se σ for uma transposição e $\sigma \star r = r$ se $\sigma = \mathbf{1}$. Note que S_2 possui apenas dois elementos, uma transposição e a identidade. Assim, cada elemento r de r_2^2 deve ser identificado com $-r$. Note que esta relação é exatamente aquela induzida pela ação do grupo

$$Z_2 = \{ -1, +1 \}.$$

Deste modo, podemos usar Z_2 ao invés de S_2 . Note que esta “substituição” é delicada. Ambos os conjuntos \mathbb{R}_2 e r_2^2 são subconjuntos de \mathbb{R}^2 escritos de forma a enfatizar as coordenadas das partículas. A ação de S_2 que atua explicitamente sobre as coordenadas das partículas não é igual à ação de S_2 em \mathbb{R}^2 . Por exemplo, o espaço \mathbb{R}_2 é invariante sob a ação de S_2 , mas \mathbb{R}^2 não o é e nem \mathbb{R}^2 nem \mathbb{R}_2 são invariantes sob a ação de Z_2 . A ação de S_2 coincide com a de Z_2 apenas para r_2^2 , ou seja, apenas para r_2^2 podemos fazer a substituição de S_2 por Z_2 .

Note que a igualdade da ação de Z_2 e a de S_2 em r_2^N independe de N , a dimensionalidade das partículas. Ela sempre corresponde a multiplicar um ponto $r \in r_2^N$ por ± 1 . Do mesmo modo \mathbb{R}_2 é sempre invariante sob a ação de S_2 . Além disso, generalizando-se a definição anterior, ambos são subconjuntos de \mathbb{R}^N , $\mathbb{R}_2 = \mathbb{R}^N$ e $r_2^N = \mathbb{R}^N - \Delta$. Este conjunto de fatos nos permite escrever o espaço de configuração de duas partículas em

dimensão arbitrária como:

$$\frac{\mathcal{C}_2^N}{\mathcal{S}_2} = \mathbb{R}_2 \times \frac{\mathbb{r}_2^N}{\mathcal{S}_2} = \mathbb{R}_2 \times \frac{\mathbb{r}_2^N}{Z_2} = \mathbb{R}^N \times \frac{\mathbb{R}^N - \{0\}}{Z_2}.$$

Analisar o processo de identificação dos pontos de \mathbb{r}_2^2 é, inclusive, mais conveniente se o escrevermos explicitamente como $\mathbb{R}^2 - \{0\}$ isto é, como pares ordenados $r = (x, y)$, $x, y \in \mathbb{R}$. O espaço $\mathbb{R}^2 - \{0\}$ e o processo de identificação estão ilustrados na figura 2.5 a seguir. O processo de identificação dos pontos é similar a um processo de dobra de um origami ou colagem e o faremos em duas partes. Primeiro, consideramos os pontos $(x, y) \in \mathbb{R}^2 - \{0\}$ tais que $y < 0$. Como dito anteriormente, a condição para que dois pontos $(x, y), (z, w) \in \mathbb{R}^2 - \{0\}$ estejam relacionados é

$$(x, y) \sim (z, w) \iff (z, w) = \pm(x, y).$$

Escrito deste modo é simples de ver que pontos pertencentes ao terceiro quadrante são mapeados ao primeiro e pontos pertencentes ao quarto quadrante, ao segundo. O semieixo y negativo é mapeado no semieixo y positivo. No caso do eixo y em particular, é como se o dobrássemos em torno do eixo x , sobrepondo precisamente cada ponto $(0, -y)$ sobre o ponto $(0, y)$. Esta primeira etapa está ilustrada no diagrama a seguir.

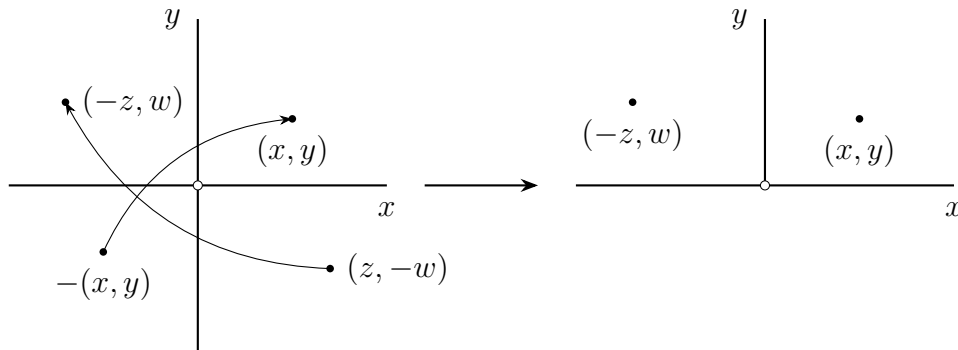


Figura 2.5: Identificação dos pontos no semiplano inferior com pontos do semiplano superior.

A segunda etapa da identificação dos pontos é identificar o semieixo x negativo com o positivo já que $(-x, 0) \sim (x, 0)$ para todo $x \in \mathbb{R}^+$. A primeira vista a “colagem” destes dois pontos pode não ser muito intuitiva. Note que estes são os únicos pontos restantes a serem identificados, pois nenhum ponto no segundo quadrante se relaciona com algum ponto do primeiro segundo \sim . Deste modo, podemos sobrepor as duas partes do eixo x como se o dobrássemos em torno de eixo y . A diferença agora é que apenas os pontos sobre o eixo “colam” uns nos outros. Os pontos do primeiro e segundo quadrantes não se relacionam. Isso torna possível “abrir” este espaço de modo similar a um filtro de café, por exemplo. A figura 2.6 a seguir ilustra esta segunda etapa.

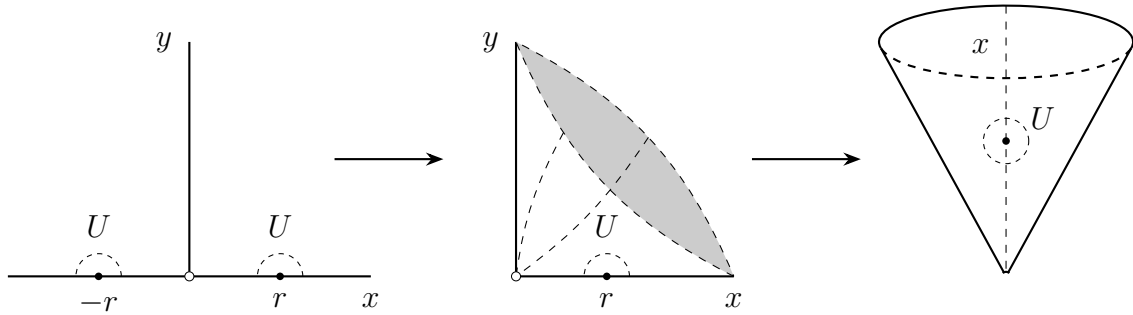


Figura 2.6: Processo de identificação do subconjunto U .

Note que o cone continua infinitamente, o corte mostrado é apenas para permitir uma melhor visualização. Note também que o cone não possui a ponta, já que a origem foi excluída. Note que o subconjunto U é dado por $\{(-x, 0), (x, 0)\}$. A rigor, o subconjunto U mostrado no cone é, na verdade a imagem de U sob o mapa quociente. No entanto, para facilitar o entendimento, preferimos manter o mesmo nome ao longo do processo.

Agora que sabemos um pouco sobre a topologia do espaço de configurações $\mathcal{C}_2/\mathbb{Z}_2$ estudemos as curvas correspondentes à permutação das duas partículas neste espaço. Como permutações das duas partículas foram identificadas, as curvas correspondentes a este processo correspondem a loops em $\mathcal{C}_2/\mathbb{Z}_2$, já que as configurações $(1, 2)$ e $(2, 1)$ correspondem ao mesmo ponto. A figura a seguir ilustra dois loops em $\mathcal{C}_2/\mathbb{Z}_2$.

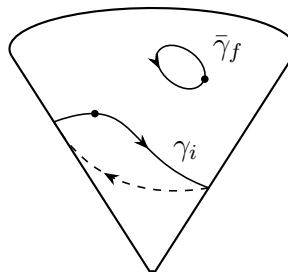


Figura 2.7: Dois loops distintos em $\mathcal{C}_2/\mathbb{Z}_2$.

Note que, como o cone não possui a ponta, os loops γ_i e $\bar{\gamma}_f$ são homotopicamente distintos. Se a ponta estivesse presente, poderíamos puxar o γ_i de modo que não percorresse toda a circunferência do cone e então deformá-lo em $\bar{\gamma}_f$. A ausência da ponta faz com que existam infinitos tipos de loops distintos. Isto é, existem infinitas classes de homotopia de loops em $\mathcal{C}_2/\mathbb{Z}_2$. Assim como γ_i não pode ser deformado em $\bar{\gamma}_f$, um loop que dá duas voltas ao redor do cone não pode ser deformado em γ_i . Assim, cada loop em $\mathcal{C}_2/\mathbb{Z}_2$ pode ser classificado por um número $n \in \mathbb{Z}$ correspondente ao número de voltas que dá em torno do cone. Um número de voltas, $n < 0$, indica que o loop percorre a circunferência no sentido contrário ao loops com $n > 0$.

É comum dizer que o espaço $\mathcal{C}_2/\mathbb{Z}_2$ é infinitamente conectado ou infinitamente conexo [2, 26, 27]. Nestes trabalhos, um espaço ser n -conexo significa que o grupo fundamental

$\pi_1(X, x_0)$ é de ordem n , ou seja, que existem n classes de homotopia de loops distintas. Formalmente falando, no entanto, um espaço é n -conexo quando todos os grupos de homotopia $\pi_i(X, x_0)$ são triviais para $i < n$. Estas duas definições são incompatíveis e mesmo que a primeira aparente ser usual em trabalhos de física, a evitaremos aqui para não causar confusão.

Neste caso foi possível determinar o espaço $\mathcal{C}_2/\mathcal{S}_2$ e a ordem do seu grupo fundamental por métodos diagramáticos — o processo de identificação de pontos de \mathcal{C}_2 . No entanto este método não é tão útil, se não completamente inviável, para espaços mais complicados, como é o caso do espaço de configurações de duas partículas idênticas em três dimensões ou mais de duas partículas em qualquer dimensão. Lembre-se que o grupo simétrico de ordem N , \mathcal{S}_N possui $N!$ elementos. Assim, mesmo para apenas três partículas, determinar diagramaticamente $\mathcal{C}^n/\mathcal{S}_3$ torna-se complicado. Enunciamos a seguir um resultado que é útil para determinar $\mathcal{C}^n/\mathcal{S}_2$.

Calcular o grupo fundamental de um espaço em geral não é uma tarefa fácil. Por este motivo, é interessante verificar se o espaço em questão não é homeomorfo a algum espaço cujo grupo fundamental já é conhecido, pois, neste caso, seus grupos fundamentais são isomorfos. No caso de duas partículas, felizmente, é possível escrever $\mathcal{C}^n/\mathcal{S}_2$ como um produto de espaços cujos grupos fundamentais são conhecidos. Usando, novamente que $\mathcal{C}_2/\mathcal{S}_2 = (\mathbb{R}^n - \Delta)/Z_2$:

$$\frac{\mathcal{C}^n}{\mathcal{S}_2} = \mathbb{R}^n \times \frac{\mathbb{R}^n - \Delta}{Z_2} \cong \mathbb{R}^n \times \mathbb{R}^+ \times \frac{\mathcal{S}^{n-1}}{Z_2}.$$

Onde

$$\mathcal{S}^{n-1} = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1 \right\}$$

é a $n - 1$ -esfera de raio unitário e $\mathbb{R}^+ = (0, \infty)$ é a semirreta real, excluindo-se o 0.

Para justificar a identificação

$$\frac{\mathbb{R}^n - \Delta}{Z_2} \cong \mathbb{R}^+ \times \frac{\mathcal{S}^{n-1}}{Z_2}$$

considere o seguinte argumento mais ou menos rigoroso. O espaço $\mathbb{R}^n - \Delta$ pode ser pensado como composto de inúmeras cascas de n -esferas concêntricas. Esta visão, por sua vez pode ser pensada como o produto de uma única n -esfera, de raio unitário com o intervalo $\mathbb{R}^+ = (0, \infty)$. Ou seja, em cada ponto, ou direção, da n -esfera, “colocamos” a semirreta $(0, \infty)$. Mais formalmente, podemos definir mapas

$$f : \mathbb{R}^n - \Delta \rightarrow \mathbb{R}^+ \times \mathcal{S}^{n-1}$$

$$f(x_1, \dots, x_n) \mapsto \left(r, \frac{x_1}{r}, \dots, \frac{x_n}{r} \right), \quad r = \left[\sum_{i=1}^n x_i^2 \right]^{\frac{1}{2}}$$

e

$$g : \mathbb{R}^+ \times \mathcal{S}^{n-1} \rightarrow \mathbb{R}^n - \Delta$$

$$g(r_+, s_1, \dots, s_n) \mapsto (r_+s_1, \dots, r_+s_n).$$

Note que r como definido acima pertence a \mathbb{R}^+ e

$$\left[\sum_{i=1}^n \left(\frac{x_i}{r} \right)^2 \right]^{\frac{1}{2}} = \frac{1}{r} \left[\sum_{i=1}^n x_i^2 \right]^{\frac{1}{2}} = \frac{r}{r} = 1.$$

De modo que $(x_1/r, \dots, x_n/r) \in \mathcal{S}^{n-1}$. Além disso, como $r_+ \in \mathbb{R}^+$, tem-se que (r_+s_1, \dots, r_+s_n) é sempre diferente de $\mathbf{0} \in \mathbb{R}^n$. Estes dois mapas são um par de bijeções inversas. De fato, para $(x_1, \dots, x_n) \in \mathbb{R}^n - \Delta$, tem-se que:

$$g(f(x_1, \dots, x_n)) = g\left(r, \frac{x_1}{r}, \dots, \frac{x_n}{r}\right)$$

$$= \left(\frac{rx_1}{r}, \dots, \frac{rx_n}{r}\right) = (x_1, \dots, x_n)$$

e, para $(r_+, s_1, \dots, s_n) \in \mathbb{R}^+ \times \mathcal{S}^{n-1}$,

$$f(g(r_+, s_1, \dots, s_n)) = f(r_+s_1, \dots, r_+s_n)$$

$$= \left(r_+, \frac{r_+s_1}{r_+}, \dots, \frac{r_+s_n}{r_+}\right) = (r_+, s_1, \dots, s_n).$$

Note que pela definição de \mathcal{S}^{n-1} :

$$\left[\sum_{i=1}^n (r_+s_i)^2 \right]^{\frac{1}{2}} = r_+ \left[\sum_{i=1}^n s_i^2 \right]^{\frac{1}{2}} = r_+.$$

Para que f e g sejam homeomorfismos entre estes dois espaços, é necessário que sejam contínuos. Não provaremos a continuidade formalmente aqui, mas pode-se pensar que, como $\mathbb{R}^n - \Delta$, \mathbb{R}^+ e \mathcal{S}^{n-1} são variedades suaves, e f e g são mapas suaves, a continuidade fica garantida. Isto é, $U \subset \mathbb{R}^n - \Delta$ é aberto se, e somente se, $f(U) \subset \mathbb{R}^+ \times \mathcal{S}^{n-1}$ o é. Por fim, como \mathbb{R}^+ é invariante sob a ação de Z_2 é possível escrever

$$\frac{\mathbb{R}^n - \Delta}{Z_2} \cong \frac{\mathbb{R}^+ \times \mathcal{S}^{n-1}}{Z_2} = \mathbb{R}^+ \times \frac{\mathcal{S}^{n-1}}{Z_2}.$$

Além disto, o mapa $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ é um homeomorfismo entre \mathbb{R}^+ e \mathbb{R} , pois para todo subconjunto $(a, b) \subseteq \mathbb{R}^+$, $\ln((a, b)) = (\ln(a), \ln(b)) \subseteq \mathbb{R}$ é aberto na topologia canônica. Similarmente, para todo subconjunto $(c, d) \subseteq \mathbb{R}$, $\exp((c, d)) = (e^c, e^d) \subseteq \mathbb{R}^+$ é aberto

na topologia do subespaço. Como ln e exp são inversas, tem-se que $\mathbb{R}^+ \cong \mathbb{R}$. Com isto, podemos escrever, de modo geral, o espaço $\mathcal{C}_n/\mathcal{S}_2$ como:

$$\mathcal{C}^n/\mathcal{S}_2 \cong \mathbb{R}^n \times \mathbb{R} \times \frac{\mathcal{S}^{n-1}}{Z_2}.$$

Esta decomposição é útil pois \mathbb{R}^n , $n \in \mathbb{N}$, são espaços conhecidos e relativamente simples, enquanto \mathcal{S}^n/Z_2 é a n -esfera com pontos antipodais — diametricamente opostos — identificados. Esta condição é precisamente um modo de definir o espaço projetivo real de dimensão n : \mathbb{RP}^n ou $\mathbb{P}^n(\mathbb{R})$. Estes espaços também são conhecidos e sabe-se que:

$$\pi_1(\mathbb{RP}^n, p_0) = \begin{cases} \mathbf{0}, & \text{se } n = 0; \\ \mathbb{Z}, & \text{se } n = 1; \\ \mathbb{Z}_2, & \text{se } n > 1; \end{cases}$$

onde $\mathbf{0}$ é o grupo trivial com um único elemento. Além disto, como descrito no apêndice B, o grupo fundamental de um produto de espaços topológicos é isomorfo ao produto dos grupos fundamentais de cada espaço. Usando os resultados anteriores, podemos escrever, então:

$$\pi_1(\mathbb{R}^n \times \mathbb{R} \times \mathbb{RP}^{n-1}, (r, x, p)) \cong \pi_1(\mathbb{R}^n, r) \times \pi_1(\mathbb{R}, x) \times \pi_1(\mathbb{RP}^{n-1}, p).$$

No caso em questão, onde $n = 2$, temos que:

$$\pi_1(\mathcal{C}^2/\mathcal{S}_2, c) \cong \mathbf{0} \times \mathbf{0} \times \mathbb{Z}.$$

Note que como o grupo $\mathbf{0}$ possui apenas um elemento, que chamaremos de e , a projeção $p_{\mathbb{Z}} : \mathbf{0} \times \mathbf{0} \times \mathbb{Z}$ dada por $p(e_1, e_2, z) = z$ é uma bijeção. Para cada elemento $z \in \mathbb{Z}$ existe um único elemento em $\mathbf{0} \times \mathbf{0} \times \mathbb{Z}$ cuja imagem é z . Este elemento é simplesmente (e_1, e_2, z) . Logo $\mathbf{0} \times \mathbf{0} \times \mathbb{Z} \cong \mathbb{Z}$ e tem-se, finalmente, que

$$\pi_1(\mathcal{C}^2/\mathcal{S}_2, c) \cong \mathbb{Z}.$$

Este resultado é exatamente o resultado esperado segundo a análise diagramática feita sobre o espaço $\mathcal{C}^2/\mathcal{S}_2$, identificando explicitamente os pontos de \mathcal{S}^2 segundo a ação de \mathcal{S}_2 . Isto é, que existem infinitos tipos de loops distintos, cada um classificado por um número inteiro, $n \in \mathbb{Z}$.

Retornando para a física, o fato de $\pi_1(\mathcal{C}^2/\mathcal{S}_2, c) = \mathbb{Z}$ terá consequências após a quantização — lembrando que esta discussão é para o espaço de configurações de partículas clássicas. Como existe um número infinito de classes de loops — permutações — é razoável supor que a condição de que a função ganhe uma fase igual a ± 1 deixe de ser

válida, podendo ganhar, no lugar, uma fase arbitrária $\phi(n)$ a depender de qual classe a permutação realizada pertence.

Duas partículas idênticas em \mathbb{R}^3

O caso de duas partículas em três dimensões não é muito diferente daquele em duas dimensões. A diferença principal é que o método que usamos para determinar a ordem do grupo fundamental do espaço de configurações $\mathcal{C}_2^3/\mathcal{S}_2$ já se torna mais complicado. Como argumentado a seção anterior, as coordenadas do centro de massa são invariantes sob ação de \mathcal{S}_2 de modo que precisamos determinar apenas r_2^3/\mathcal{S}_2 neste caso. Além disso também ficou estabelecido que a ação de \mathcal{S}_2 coincide com aquela de Z_2 .

O problema, no entanto é que ao invés de ter que identificar pontos pertencentes a apenas quatro quadrantes, os pontos de $\mathbb{R}^3 - \Delta$ estão divididos em oito octantes. O processo se torna significativamente mais difícil de visualizar e, como usamos um método visual para determinar os tipos de loops possíveis, é mais recomendado procurar outro método. Usaremos, portanto, os resultados já obtidos e escrevemos:

$$\frac{\mathcal{C}^3}{\mathcal{S}_2} \cong \mathbb{R}^3 \times \mathbb{R} \times \frac{\mathcal{S}^2}{Z_2} = \mathbb{R}^3 \times \mathbb{R} \times \mathbb{RP}^2$$

onde \mathcal{S}^2 é a 2-esfera, a esfera usual e \mathbb{RP}^2 o espaço projetivo de dimensão 2. Este espaço é obtido a partir da esfera pela identificação de pontos antipodais. O hemisfério superior de uma tal esfera está ilustrado na figura abaixo.

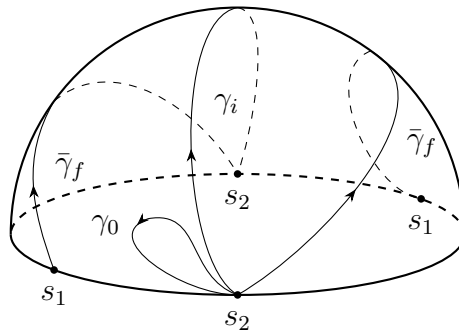


Figura 2.8: Loops no hemisfério superior de uma esfera com pontos antipodais identificados.

Note que todos os caminhos representados na figura, apesar de parecerem caminhos abertos, são loops, pois os pontos antipodais são identificados. Aqui surgiria uma dificuldade não antecipada para o método visual que usamos para partículas em duas dimensões: o espaço \mathbb{RP}^2 não pode ser imerso em \mathbb{R}^3 , o que significa que não podemos visualizá-lo completamente. Isto torna bastante difícil determinar quantas classes de loops distintas existem.

Usando a descrição analítica de $\mathcal{C}^3/\mathcal{S}_2$ e os resultados dispostos na seção anterior é direto verificar que

$$\pi_1(\mathcal{C}^3/\mathcal{S}_2, c) \cong \mathbf{0} \times \mathbf{0} \times Z_2 \cong Z_2.$$

Ou seja, existem exatamente duas classes distintas de loops em $\mathbb{R}\mathbb{P}^2$ e $\mathcal{C}^3/\mathcal{S}_2$. Da figura, é fácil ver que γ_0 pode ser contraído a um ponto enquanto γ_i e $\bar{\gamma}_f$ parecem não serem contráteis. No entanto, $\bar{\gamma}_f$ é contrátil enquanto γ_i não. Acontece que se um loop é não contrátil, como γ_i , tem-se que $\gamma_i^2 = \gamma_i * \gamma_i$ é e o loop $\bar{\gamma}_f$ pode ser considerado uma caminho intermediário na deformação de γ_i^2 a um ponto, isto é, é homotópico a γ_i^2 e, portanto, contrátil. Este resultado pode ser provado analiticamente mas é bastante contraintuitivo já que não podemos visualizar $\mathbb{R}\mathbb{P}^2$ completamente.

É interessante notar que apesar de $\mathcal{C}^3/\mathcal{S}_2$ ser topologicamente distinto de \mathbb{R}^3 , ele possui exatamente dois tipos de loops o que é consistente com a divisão de partículas em três dimensões em duas categorias, férmions e bósons. Isto indica que, mesmo que o argumento utilizado anteriormente e geralmente utilizado para justificar a divisão nessas duas classes de comportamento estatístico de partículas produza o resultado correto, o problema é mais profundo.

Isto, inclusive, pode ser lido como um comentário mais profundo sobre a metodologia científica. Usualmente diz-se que a natureza é a juiz final de todas as teorias. Argumento que muitas vezes é usado para retroativamente justificar algumas hipóteses feitas em um modelo, já que assim, o resultado final será condizente com o experimento. Essa linha de raciocínio já foi criticada por estudiosos de epistemologia da ciência, por exemplo, o livro *Contra o Método* de Michael Feyerabend [28]. Mas, além disso, serve de indicativo de que esta linha argumentativa pode levar a “problemas” como o discutido aqui. Mirman, Leinaas e Myrheim, dentre outros apontaram as falhas nas bases teóricas da mecânica quântica de partículas idênticas e até o momento da escrita deste trabalho, pouco mais de quarenta anos depois destas críticas terem sido feitas, esta abordagem ainda está presente nos livros textos da disciplina.

Reexaminações crítica dos fundamentos de uma teoria se fazem necessárias com o passar do tempo, até por se constatar que as primeiras tentativas de se explicar um novo fenômeno invariavelmente serão “incompletas” ou “fracas”. Um exemplo histórico é o próprio desenvolvimento da mecânica quântica. A formalização da teoria quântica em um conjunto de axiomas utilizando-se de conceitos como espaços de Hilbert e operadores neste espaço tornou a descrição anterior de fenômenos quânticos, alguns, inclusive, descorrelacionados, muito mais nítida e coesa.

Esta mesma formalização permitiu que inúmeros mais avanços fossem realizados. Neste sentido, pode-se entender também que argumentos não rigorosos, física ou matematicamente por exemplo, são uma espécie de mina, possivelmente rica de novos fenômenos. Por exemplo, a crítica de que os operadores de permutação, trocando apenas os índices

das partículas e não as trocando, fisicamente, as partículas em si foi o catalizador de toda a discussão até aqui e, como veremos a seguir, é suficiente para mostrar que o postulado de simetridade, justificado precisamente por necessidades experimentais, na verdade é um resultado direto do tratamento adequado de permutações.

n partículas idênticas em \mathbb{R}^N

Antes de tratar da quantização, gostaríamos de discutir brevemente o espaço de configurações do caso geral de n partículas em N dimensões. Definimos o espaço de configurações de modo geral por:

$$\mathcal{C}_n^N = \frac{X^n - \Delta}{\mathcal{S}_n},$$

como anteriormente. Para partículas em dimensão N , tem-se que $X = \mathbb{R}^N$. Além disso, a separação em termos do centro de massa e coordenadas relativas ainda é válida, a diferença é que o espaço relativo agora possui dimensão $L = nN - n$. O espaço correspondente ao centro de massa continua com dimensão igual à dimensão do espaço onde as partículas estão inseridas, no caso, N . Deste modo, escrevemos:

$$\mathcal{C}_n^N = \mathbb{R}^N \times \frac{\mathbb{R}^L - \Delta}{\mathcal{S}_n}.$$

O conjunto $\mathbb{R}^L - \Delta$, por sua vez é homeomorfo a $\mathbb{R} \times \mathcal{S}^{L-1}$, como estabelecido anteriormente. No entanto, este resultado já não é mais útil quanto antes pois $\mathcal{S}^{L-1}/\mathcal{S}_n$ para $n > 2$ não é espaço “canônico” como os espaços projetivos. No entanto, o grupo fundamental de \mathcal{C}_n^N já foi calculado. Tem-se:

$$\pi_1(\mathcal{C}_n^N, c) = \begin{cases} \mathcal{B}_n & \text{para } N = 2; \\ \mathcal{S}_n & \text{para } N > 2; \end{cases}$$

onde \mathcal{B}_n é o grupo tranças de n objetos e \mathcal{S}_n é simplesmente o grupo simétrico de ordem n . Trataremos mais do grupo \mathcal{B}_n e de sua significância para anyons na seção a seguir.

Por fim, alguns comentários sobre \mathcal{C}_n^N no geral. Primeiro, notamos que a ele pode ser dada a estrutura de uma variedade. Isto se dá porque o produto de variedades é uma variedade [21, 29]. Um detalhe é que o espaço quociente de um espaço topológico que também é uma variedade, como $\mathbb{R}^N - \Delta$ em geral não o é, porém, o grupo simétrico satisfaz propriedades o suficiente de modo que este a este espaço também possa ser dada a estrutura de variedade. Deste modo é possível definir coordenadas locais em \mathcal{C}_n^N para especificar uma configuração c . Isto será útil durante a quantização, assunto que começaremos a discutir a seguir.

2.3 - Anyons

Nesta seção trataremos de anyons propriamente. Anyons, introduzidos anteriormente, são partículas em duas dimensões e, devido a isto, são partículas que podem possuir qualquer valor de estatística. Isto é, podem ter um comportamento estatístico intermediário entre férmions e bósons. Como discutido anteriormente, esse comportamento surge devido às propriedades topológicas do espaço de configurações de partículas idênticas.

Até agora, tratamos do espaço de configurações clássicos, com alguns comentários sobre como ele influencia a descrição quântica do sistema. A partir de agora trataremos propriamente da quantização do sistema clássico para obter, formalmente, anyons e também, o postulado de simetridade para o caso de partículas em três dimensões. Na seção que se segue e na próxima serão apresentados dois processos para obter a descrição quântica adequada de partículas idênticas. Um usando a formulação de Feynman, segundo Laidlaw e DeWitt e Wu [1, 5] e o processo de quantização geométrica por Souriau e Leinaas e Myrheim [30, 31, 2].

Notamos que o processo de quantização geométrica apresentado por Souriau é mais rigoroso que aquele apresentado por Leinaas e Myrheim em seu artigo [2]. No entanto, nos ateremos aqui ao último por dois motivos. Primeiro, o artigo original de Souriau [30] está disponível apenas em francês, língua desconhecida para autor. Segundo, o livro, cujo título em tradução livre é Estrutura de Sistemas Dinâmicos, Uma Visão Simplética da Física [31], assim como o artigo original foram descobertos apenas na fase final de escrita deste trabalho por acidente e alguns dos pré-requisitos necessários estavam além do conhecimento do autor no momento da escrita. Notamos que os trabalhos de Souriau antecedem os de Leinaas e Myrheim e um provável motivo pelo qual não foram citados pela dupla tenha sido porque eram desconhecidos a eles, já que foram publicados originalmente em francês e ambos noroegueses.

2.3.1 - Quantização

Iniciaremos o processo de quantização de modo a obter uma descrição na formulação de Schrödinger da mecânica quântica, segundo o artigo de Leinaas e Myrheim [2], como mencionado. Este método de quantização, no entanto é melhor explicado em um conjunto de notas devido a Myrheim [32]. Estas notas contêm, também, uma quantização de modo a obter uma descrição na formulação de Heisenberg, no entanto, não seguiremos esta abordagem.

O processo de quantização consiste, essencialmente, em construir o espaço de Hilbert do sistema quântico, atribuindo a cada ponto c do espaço de configurações \mathcal{C} um espaço

de Hilbert unidimensional, \mathfrak{h}_c . Se

$$\mathbf{H} = \bigsqcup_{c \in \mathcal{C}} \mathfrak{h}_c = \bigcup_{c \in \mathcal{C}} \{c\} \times \mathfrak{h}_c$$

então podemos definir um mapa $\Pi : \mathbf{H} \rightarrow \mathcal{C}$, tal que para $|\psi\rangle \in \mathfrak{h}_c$, $\Pi(|\psi\rangle) = c$, isto é Π projeta um vetor do espaço quântico em sua correspondente configuração clássica. Notamos, primeiro, que representamos o espaço de configurações clássicas aqui por \mathcal{C} apenas para simplificar a notação, ainda se trata do espaço obtido a partir da identificação de pontos relacionados pela ação de grupo simétrico de ordem apropriada.

Em segundo lugar, note que a uma única configuração clássica corresponde um espaço unidimensional de configurações quânticas. Isto é de se esperar já que, segundo os postulados da mecânica quântica, vetores que diferem apenas por uma fase correspondem ao mesmo estado. Esta fase, a princípio é arbitrária, porém o requerimento adicional de que a norma seja igual a 1, para que seja uma probabilidade, impõem que a fase seja um elemento do grupo unitário $U(1)$. Por último, note que o espaço quântico obtido corresponde à representação no espaço real, já que as coordenadas no espaço de configurações são dadas em termos das coordenadas das partículas.

Note que sob o mapa Π , o espaço \mathfrak{h}_c corresponde a uma fibra, isto é: $\mathfrak{h}_c = \Pi^{-1}(\{c\})$, onde Π^{-1} aqui representa a imagem inversa ou pré-imagem e não um mapa inverso. Como todos os espaços \mathfrak{h}_c são, por exigência, espaços vetoriais de dimensão 1, dizemos, por abuso de linguagem, que \mathbf{H} é um fibrado de linhas sobre \mathcal{C} . O próximo passo consiste em escolher uma base para os espaços \mathfrak{h}_c . Como todos são unidimensionais, a base de cada espaço consiste em apenas um vetor que denotaremos por $|\chi_c\rangle$. Uma função de onda, Ψ é uma seção do fibrado de linhas, isto é, uma inversa direita do mapa Π . Assim, $\Psi : \mathcal{C} \rightarrow \mathbf{H}$ atribui a cada configuração c um vetor pertencente à fibra de c , isto é:

$$\Psi(c) = \psi(c)|\chi_c\rangle$$

e tal que

$$\Pi(\Psi(c)) = \Pi(\psi(c)|\chi_c\rangle) = c.$$

Note que $\psi(c)$ é um número complexo correspondente à, única, componente do vetor em \mathfrak{h}_c . A figura 2.9 a seguir é uma representação esquemática de um fibrado de linhas e uma seção. Na figura, cada linha vertical corresponde a um dos espaços \mathfrak{h}_c e a seção Ψ toma um valor em cada um destes espaços.

Uma vez definida as bases, impõem-se a condição que sejam normalizadas. Esta restrição ainda nos permite realizar mudanças de base da forma:

$$|\chi_c\rangle \mapsto |\tilde{\chi}_c\rangle = e^{i\alpha(c)}|\chi_c\rangle$$

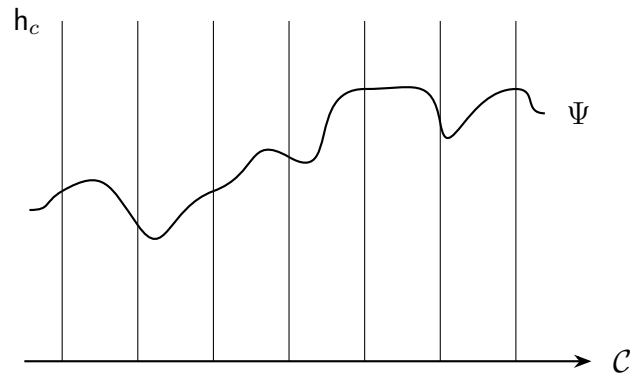


Figura 2.9: Representação esquemática do fibrado de linhas e de uma seção.

onde $\alpha : \mathcal{C} \rightarrow \mathbb{R}$ é uma fase. Esta transformação da base é chamada de transformação de calibre — *gauge* — local. Naturalmente, exigimos que a função de onda Ψ seja independente da escolha de base, significando, portanto, que deve ser invariante sob mudanças de calibre. Sob uma mudança de base, a função de onda se transforma segundo:

$$\psi|\chi_c\rangle \mapsto \tilde{\psi}|\tilde{\chi}_c\rangle = \tilde{\psi}e^{i\alpha(c)}|\chi_c\rangle$$

Para que seja invariante, portanto, devemos ter

$$\tilde{\psi} = e^{-i\alpha(c)}\psi.$$

Agora, o último passo é definir uma conexão em \mathbf{H} . Uma conexão é essencialmente uma regra para transporte paralelo de vetores, essencialmente, devemos responder à pergunta, como um vetor $\psi(c_1)|\chi_{c_1}\rangle$ é levado em $\psi(c_2)|\chi_{c_2}\rangle$ quando a função de onda Ψ evolui da configuração c_1 para a configuração c_2 . É importante recordar que o espaço de configurações \mathcal{C} em geral não é plano como seria se fosse apenas o produto de espaços de configuração de partículas únicas. A identificação segundo \mathcal{S}_N junto com a remoção de pontos coincidentes alterou significativamente a estrutura topológica de \mathcal{C} . Uma função de onda corresponde a um campo vetorial em \mathbf{H} , assim, dada uma evolução do sistema — uma curva, γ em \mathcal{C} — se desejamos saber como o vetor $\Psi(c)$ muda durante esta evolução, devemos transportá-lo paralelamente ao longo da curva γ . Isto é um transporte entre fibras.

Definimos, portanto, operadores lineares e unitários $T(c_2, c_1) : \mathbf{h}_{c_1} \rightarrow \mathbf{h}_{c_2}$ que transportam vetores de \mathbf{h}_{c_1} para \mathbf{h}_{c_2} paralelamente ao longo de alguma curva γ . Os operadores T , a princípio, dependem da curva γ , porém, segundo Leinaas e Myrheim, é suficiente definir um transporte paralelo infinitesimal, $T(c + dc, c)$ e postulamos a seguinte regra:

$$T(c + dc, c)|\chi_c\rangle = (1 + idc^i a_i(c))|\chi_{c+dc}\rangle$$

Onde c^i são as coordenadas locais de \mathcal{C} . Como foi discutido, podemos definir coordenadas locais em \mathcal{C} por se tratar de uma variedade. Por exemplo, no caso de $\mathcal{C} = \mathcal{C}_2^2/\mathcal{S}_2$, isto é, duas partículas em duas dimensões, as coordenadas locais são essencialmente $R = (x_{cm}, y_{cm})$ do centro de massa e $r = (x, y)$ de uma partícula relativa à outra. Assim, teríamos $c = (x_{cm}, y_{cm}, x, y)$ e os c^i são, simplesmente, as componentes de c na direção e_i .

A regra infinitesimal para transporte paralelo implica que os operadores $T(c_2, c_1)$ são escritos como

$$T(c_2, c_1)|\chi_{c_1}\rangle = \exp\left(i \int_{\gamma} dc^i a_i(c)\right)|\chi_{c_2}\rangle$$

onde γ é a curva ao longo da qual os vetores são transportados. Com os operadores T , podemos definir uma derivada em \mathbf{H} , chamada derivada covariante. A necessidade de usar T para definir esta derivada retorna à questão do transporte paralelo, o processo de derivação é essencialmente um processo de comparação entre os valores da função em um ponto e em outro. O transporte paralelo, faz-se necessário para que os vetores $\Psi(c_1)$ e $\Psi(c_2)$ possam ser comparados, por exemplo. A necessidade de uma estrutura diferencial em \mathbf{H} surge, pela necessidade de se definir o operador hamiltoniano do sistema, por exemplo, que geralmente inclui derivadas nas coordenadas devido ao operador momento.

Definimos, então D , a derivada covariante segundo a condição de que ela seja invariante de calibre. Assim:

$$\Psi(c + dc) = T(c + dc, c)(\Psi(c) + dc^i [D_i \Psi](c))$$

onde D_i representa a i -ésima componente de D e índices repetidos, uma soma. O operador de translação paralela é necessário pois $\Psi(c + dc) \in \mathfrak{h}_{c+dc}$ enquanto $\Psi(c)$ e $[D_i \Psi](c)$ pertencem a \mathfrak{h}_c . Como está, a definição é invariante de calibre pois não faz menção a uma base específica. Se escolhermos uma base, no entanto, podemos escrever:

$$\Psi(c + dc) = T(c + dc, c)(\psi(c) + dc^i [D_i \psi](c))|\chi_c\rangle$$

Note, se isto fosse uma derivada comum, obteríamos a derivada ao longo da direção dc , escrevendo $dc = tv$, onde $t \in \mathbb{R}$ e $v \in \mathcal{C}$ é a direção ao longo do qual queremos calcular a derivada, dividiríamos por t e tomaríamos o limite. Neste caso, no entanto, a curvatura do espaço gera termos extras que ficariam “sobrando” neste processo. A ideia de definir a derivada covariante é justamente defini-la de modo a anular o efeito destes termos extras. Para prosseguir com este processo, no entanto é mais conveniente determinar a forma de

D para cada componente separadamente. Assim, fazemos $dc = te_i$ e escrevemos:

$$\begin{aligned}\Psi(c + dc) &= T(c + te_i, c)(\psi(c) + t[D_i\psi](c))| \chi_c \rangle \\ &= (\psi(c) + t[D_i\psi](c))(1 + ta_i(c))| \chi_{c+te_i} \rangle \\ &= \left[\psi(c) + [ia_i(c)\psi(c) + [D_i\psi](c)]t + [ia_j(c)[D_i\psi](c)]t^2 \right] | \chi_{c+te_i} \rangle\end{aligned}$$

Além disto, tem-se que

$$\Psi(c + dc) = \psi(c + dc)| \chi_{c+te_i} \rangle = [\psi(c) + t[\partial_i\psi](c)]| \chi_{c+dc} \rangle.$$

Onde ∂_j é a derivada direcional usual, lembre-se que ψ é uma função de $\mathcal{C} \rightarrow \mathbb{C}$. Substituindo na equação acima:

$$0 = \left[[ia_i(c)\psi(c) - [\partial_i\psi](c) + [D_i\psi](c)]t + ia_j(c)[D_i\psi](c)t^2 \right] | \chi_{c+te_i} \rangle.$$

Por fim, dividindo por t e tomando o limite $t \rightarrow 0$, obtém-se:

$$0 = [ia_j(c)\psi(c) - [\partial_i\psi](c) + [D_i\psi](c)]| \chi_c \rangle$$

donde concluímos que, para garantir a invariância de calibre, devemos ter

$$D_i = \partial_i - ia_i.$$

Antes de prosseguirmos, chamamos atenção para o potencial de calibre, a e calibre α , notando que o potencial de calibre é definido a menos de uma escolha de calibre. Se começarmos com uma base $e^{i\alpha(c_1)}| \chi_{c_1} \rangle$, tem-se que:

$$T(c_2, c_1)| \chi_{c_1} \rangle = \exp \left(i\alpha(c_1) + i \int_{\gamma} dc^i a_i(c) \right) | \chi_{c_2} \rangle.$$

Como possuímos liberdade para determinar o calibre, podemos escolhê-lo de modo que $\alpha(c_2) = 0$ e escrever

$$\alpha(c_1) = - \int_{\gamma} dc^i [\partial_i\alpha](c)$$

de modo que

$$\tilde{a}_i = a_i - \partial_i\alpha.$$

Onde vemos \tilde{a} como o potencial de calibre correspondente a uma escolha de calibre distinta, $| \tilde{\chi}_{c_2} \rangle$.

Por fim, o último objeto do qual precisamos para poder utilizar este espaço quântico que construímos para partículas indistinguíveis é a curvatura da conexão, \mathcal{F}_{ij} , dada pelo

comutador das componentes de D :

$$\mathcal{F}_{ij} = i[D_i, D_j] = \partial_i a_j - \partial_j a_i.$$

É a partir deste objeto que é possível extrair o comportamento mais marcante de partículas quânticas indistinguíveis. Para tal, Leinaas e Myrheim fazem uma restrição no estudo, supondo que \mathcal{F}_{ij} seja identicamente nulo, exceto possivelmente nas singularidades, isto é, os pontos que foram removidos na construção do espaço de configurações. Esta hipótese, segundo eles, é mais um restrição no escopo do estudo do que uma hipótese física de fato. É importante salientar, inclusive, que em geral é possível tornar \mathcal{F}_{ij} nulo simplesmente por uma escolha adequada de calibre.

A condição para que \mathcal{F}_{ij} seja identicamente nula, escrevemos $\mathcal{F}_{ij} \equiv 0$, é que o potencial de calibre seja o gradiente de alguma função escalar α , isto é, que exista um escolha de calibre tal que $\tilde{a}_i = a_i - \partial_i \alpha \equiv 0$. A verificação de que $\tilde{a}_i \equiv 0 \implies \mathcal{F}_{ij} \equiv 0$ é direta. Por outro lado, temos que $\mathcal{F}_{ij} \equiv 0 \implies \tilde{a}_i \equiv 0$ somente se α não possuir ramificações, ou seja, se não for multivalorada. Esta condição garante que:

$$\oint_{\gamma} dc^i a_i = \oint_{\gamma} dc^i [\partial_i \alpha](c) = 0$$

desde que o caminho γ não circunde nenhuma singularidade. Deste modo, como estamos supondo que $\mathcal{F}_{ij} \equiv 0$, α só será multivalorada em espaços que não são simplesmente conexos e, como vimos, este não é o caso dos espaços de configurações \mathcal{C} de partículas indistinguíveis. Note que isto não significa que \tilde{a}_i não seja identicamente nula, mas sim que o potencial de calibre α_i passa a ser multivalorado, assim como o calibre α .

Por fim, as consequências do calibre ser multivalorado podem ser observadas considerando-se loops no espaço de configurações, correspondendo precisamente à permutação de partículas. Neste caso, tem-se:

$$T(c, c) | \chi_c \rangle = \exp \left(i \oint_{\gamma} dc^i a_i(c) \right) | \chi_c \rangle$$

onde cometemos um pequeno abuso de notação ao usar c também como variável de integração. Note que, como γ começa e termina em c o vetor $|\chi_c\rangle$ é transportado paralelamente por γ , adquirindo uma fase, $\phi(\gamma)$, dada por:

$$\phi(\gamma) = \oint_{\gamma} dc^i a_i(c).$$

O valor de $\phi(\gamma)$ depende da natureza de γ , mais especificamente, de sua classe de homotopia pois a ϕ é invariante sob deformações contínuas de γ e saltos de um ramo a outro são descontínuos. Além disso, tem-se que a fase adquirida ao se transportar o vetor

da base por um loop $\bar{\gamma}_f$ após outro, γ_i é produto das duas fases. Escrevendo $T(\gamma)$ ao invés de $T(c, c)$ para evidenciar o caminho, tem-se:

$$T(\bar{\gamma}_f \gamma_i) = T(\bar{\gamma}_f)T(\gamma_i).$$

Como $T(\gamma)$ depende apenas da classe de equivalência de γ , $[\gamma]$, temos que os operadores T definem, para cada ponto c , uma representação do grupo fundamental de \mathcal{C} . Deste modo, uma função de onda Ψ se transforma pela ação de uma representação unidimensional e unitária de $\pi_1(\mathcal{C}, c)$ quando partículas são permutadas. Escrevemos então:

$$\Psi(c) \mapsto T(\gamma)\Psi(c) \tag{2.5}$$

Como exemplo, considere o caso de duas partículas em três dimensões. O grupo fundamental de \mathcal{C}_2^3 é isomorfo a \mathcal{S}_2 e este grupo possui apenas duas representações unidimensionais, ambas unitárias, dadas por $T(\gamma) = 1$ e $T(\gamma) = \text{sgn}([\gamma])$. Aqui $[\gamma] \in \pi_1(\mathcal{C}_2^3, c)$ é a classe de equivalência do loop γ . Como $\pi_1(\mathcal{C}_2^3, c)$ é isomorfo a \mathcal{S}_2 , cada loop pode ser associado, ou à identidade ou à transposição de \mathcal{S}_2 . Por exemplo, podemos associar os loops contráteis à identidade e os não contráteis à transposição. Desde modo, loops não contráteis são permutações ímpares e os contráteis permutações, pares. Assim, se escolhermos a representação completamente simétrica, $T(\gamma) = 1$, o calibre será unicamente valorado e existirá apenas uma base: $\{|\chi_c\rangle\}$. Por outro lado, se escolhermos a representação completamente antissimétrica, $T(\gamma) = \text{sgn}([\gamma])$, o calibre será multivalorado e teremos duas bases distintas para cada ponto c : $\{|\chi_c\rangle, -|\chi_c\rangle\}$. Vemos então, que a representação simétrica está associada a bósons e a antissimétrica a férmions.

No caso geral, vimos que o grupo fundamental para N partículas em mais de duas dimensões é o grupo simétrico de ordem N . Para qualquer N , este grupo possui apenas duas representações unidimensionais, a completamente simétrica, correspondendo a bósons e a completamente antissimétrica correspondendo a férmions, definidas exatamente com antes. Note, portanto, que este processo de quantização produziu o postulado de simetridade como um resultado. Além disso, também descreve o que acontece em duas dimensões, onde o grupo fundamental é o grupo de tranças de n elementos. Este grupo possui mais de duas representações unidimensionais e fornece fatores estatísticos intermediários entre férmions e bósons. Voltaremos a tratar dele mais adiante.

Com isto, finalizamos esta seção. Antes de passarmos para a próxima, alguns comentários são importantes. Primeiro, vimos que, de fato, é possível obter o postulado de simetridade como um resultado da quantização adequada de sistemas de partículas idênticas. Note que este processo produz um espaço que não é simplesmente o produto de espaços de uma única partícula. O sistema de N partículas é quantizado como um todo.

Segundo, como mencionado antes, a abordagem de Leinaas e Myrheim e, conseqüentemente aquela reproduzida aqui carece de completude e formalidade. Por exemplo, em seu

artigo [2] e nas notas [32] é dito que os espaços h_x são fibras no entanto, assim como neste trabalho, nenhuma prova é fornecida para sustentar esta afirmação, assim como o espaço total não é construído. Além disso, nenhum argumento é fornecido sobre a exigência de que h_x deve ser unidimensional. Particularmente, também consideramos que o desenvolvimento envolvendo a derivada covariante e especialmente a curvatura da conexão são apresentadas com um nível de rigor abaixo do desejado neste trabalho. Quanto a isto, no entanto, existem limitações externas e tal tratamento fica reservado para trabalhos futuros. Passamos agora para outro método de descrever partículas quânticas idênticas, desta vez baseado na formulação de Feynman. Não serão obtidos resultados novos porém esta formulação é especialmente útil para visualizar o grupo de tranças.

2.3.2 - Formulação de Feynman

Um pouco antes de Leinaas e Myrheim, Laidlaw e DeWitt [1] obtiveram também o postulado de simetricidade não como uma hipótese mas como um resultado por meio da formulação de Feynman. Após isto, e após o artigo de Leinaas e Myrheim, Wu [5] obteve o mesmo resultado para partículas em duas dimensões — estatística arbitrária — explorando mais a fundo o resultado de Laidlaw e DeWitt para este caso. Incluímos o tratamento aqui porque a formulação de Feynman proporciona um modo de ver sistemas anyônicos muito útil e, apesar de não seguirmos por esta abordagem neste trabalho, uma ligação direta com teorias de campo topológicas.

Tratar de anyons é, de certo modo, mais simples quando feito pela formulação de Feynman. Isto se dá porque ela é construída explicitamente a partir de sistemas clássicos. A probabilidade de um sistema inicialmente no estado i ser encontrado no estado f é dado por uma soma sobre todos os caminhos conectando i e f no espaço de configurações do sistema. Neste sentido, a simplicidade de usar esta formulação é que tudo que é necessário para obter as propriedades anyônicas de partículas em duas dimensões é a descrição adequada do espaço de configurações.

O detalhe de usá-la, no entanto, é que, como apontado por Schulman [33], para espaços cujo grupo fundamental não é trivial, o propagador de Feynman é definido com uma combinação linear de amplitudes parciais, cada uma calculada sobre caminhos que pertencem à mesma classe de homotopia. Denotemos por $\Pi[a, b]$ o conjunto de todas as classes de homotopia de caminhos que vão de $a \in \mathcal{C}$ a $b \in \mathcal{C}$. Note que este conjunto possui a estrutura de um grupoide [teorema B.2]. Se $\alpha \in \Pi[a, b]$ e $f \in \alpha$ é um caminho em \mathcal{C} , escrevemos:

$$\mathcal{K}(a, t_0; b, t_1) = \sum_{\alpha \in \Pi[i, j]} \chi(\alpha) \mathcal{K}^\alpha$$

onde $\chi(\alpha)$ é um número complexo, um peso atribuído a α e

$$\mathcal{K}^\alpha = \int_{f \in \alpha} \mathcal{D}f \exp \left(\frac{i}{\hbar} \int_{t_0}^{t_1} dt \mathcal{L} \right)$$

é a amplitude de transição como definida usualmente, onde \mathcal{L} é a lagrangiana do sistema clássico.

O problema, então é calcular os pesos χ . Laidlaw e DeWitt [1] mostraram que χ é uma representação unidimensional do grupo fundamental do espaço de configurações. É interessante notar que isto vale mesmo quando não estamos tratando de loops em \mathcal{C} , afinal, como mencionado, o conjunto $\Pi[a, b]$ é um grupoide. A conexão com o grupo fundamental pode ser feita se escolhermos um ponto $c \in \mathcal{C}$ e escrevermos f como um produto $\gamma_a * \lambda * \bar{\gamma}_b$, onde γ_a é um caminho de a a c , λ um loop sitiado em c e $\bar{\gamma}_b$ um caminho de c a b . A definição da operação $*$ pode ser encontrada no apêndice B, definição B.3, mas trata-se essencialmente de uma composição de caminhos. Como $*$ é uma operação bem definida, também, no conjunto das classes de homotopias de caminhos, podemos então usar esta decomposição para estabelecer uma bijeção entre o grupo fundamental: $\pi_1(\mathcal{C}, c)$ e $\Pi[a, b]$.

Para formalizar esta ideia, denotemos por $\Omega(\mathcal{C}, a, b)$ o conjunto de todos os caminhos em \mathcal{C} de a a b . Fixamos um ponto $c_0 \in \mathcal{C}$ e definimos um mapa \mathcal{M} que associa a cada ponto $c \in \mathcal{C}$ um caminho que vai de c_0 a c . Tal mapa é chamado de *homotopy mesh* por Laidlaw e DeWitt. Traduziremos este termo aqui como malha homotópica, essencialmente porque para cada ponto $c \in \mathcal{C}$, \mathcal{M} seleciona um caminho de $\Omega(\mathcal{C}, c_0, c)$. Se considerarmos o conjunto de todos tais caminhos, podemos imaginar \mathcal{M} como recobrindo o espaço \mathcal{C} com uma malha de caminhos.

Note que existem infinitos mapas \mathcal{M} , um para cada seleção distinta de caminhos de $\Omega(\mathcal{C}, c_0, c)$. Veremos que isto não será um problema, no entanto. O mapa \mathcal{M} então, induz uma injeção Λ_{ab} de $\pi_1(\mathcal{C}, c_0)$ em $\Pi[a, b]$ dada por:

$$\begin{aligned} \Lambda_{ab} : \pi_1(\mathcal{C}, c) &\rightarrow \Pi[a, b] \\ [\lambda] &\mapsto [\mathcal{M}(a)] * [\lambda] * [\bar{\mathcal{M}}(b)] \end{aligned}$$

Onde $\bar{\mathcal{M}}(b)$ é o caminho que começa em b e termina em c_0 . É simples de verificar que Λ_{ab} é uma bijeção, basta lembrar das propriedades de $*$, teorema B.2, em especial, para qualquer caminho γ em \mathcal{C} , tem-se:

$$[\bar{\gamma}] * [\gamma] = [e_{c_1}] \quad \text{e} \quad [\gamma] * [\bar{\gamma}] = [e_{c_0}]$$

e

$$[\gamma] * [e_{x_1}] = [\gamma] \quad \text{e} \quad [e_{x_0}] * [\gamma] = [\gamma]$$

onde c_0 e c_1 são os pontos inicial e final de γ respectivamente e e_c é o caminho constante

sitiado em c . Assim, supondo

$$[\mathcal{M}(a)] * [\lambda] * [\overline{\mathcal{M}}(b)] = [\mathcal{M}(a)] * [\mu] * [\overline{\mathcal{M}}(b)],$$

para $\lambda, \mu \in \pi_1(\mathcal{C}, c)$, podemos multiplicar esta igualdade pela esquerda por $[\overline{\mathcal{M}}(i)]$ e à direita por $[\mathcal{M}(b)]$ de modo a obter $[\lambda] = [\mu]$, ou seja, Λ_{ab} é injetivo. Além disto, supondo que \mathcal{C} seja conexo por caminhos, então Λ_{ab} é sobrejetiva, pois para qualquer ponto c , existe um caminho de c_0 a c . Esta hipótese corresponde fisicamente a supor que seja possível levar o sistema a uma configuração c a partir de qualquer configuração c_0 , isto é, não existem configurações inacessíveis e é bastante razoável. Por fim, Λ_{ab} é compatível com a operação $*$ no sentido que

$$\begin{aligned} \Lambda_{ab} * \Lambda_{bc} &= [\mathcal{M}(a)] * [\lambda] * [\overline{\mathcal{M}}(b)] * [\mathcal{M}(b)] * [\mu] * [\overline{\mathcal{M}}(c)] \\ &= [\mathcal{M}(a)] * [\lambda * \mu] * [\overline{\mathcal{M}}(c)] = \Lambda_{ac}. \end{aligned}$$

Uma vez escolhida uma malha homotópica, portanto, podemos identificar cada amplitude parcial \mathcal{K}^α , onde $\alpha \in \Pi[a, b]$ com $\mathcal{K}^{\Lambda_{ab}(\lambda)}$ onde $\lambda \in \pi_1(\mathcal{C}, c_0)$ é tal que $\Lambda_{ab}(\lambda) = \alpha$. Assim, escrevemos o propagador de Feynman como

$$\mathcal{K}(a, t_0; b, t_1) = \sum_{\lambda \in \pi_1(\mathcal{C}, c)} \chi(\Lambda_{ab}(\lambda)) \mathcal{K}^{\Lambda_{ab}(\lambda)}.$$

Sobre o propagador impomos duas condições:

- i A amplitude de probabilidade obtida deve ser independente da malha homotópica escolhida.
- ii A regra usual de composição de amplitudes deve ser satisfeita. Isto é:

$$\mathcal{K}(a, t_0; c, t_2) = \int_{\mathcal{C}} db \mathcal{K}(a, t_0; b, t_1) \mathcal{K}(b, t_1; c, t_2).$$

Provaremos que, para as duas condições acima sejam satisfeitas, é suficiente que para todos os pesos tenhamos $\chi(\Lambda_{ab}(\lambda)) = \rho(\lambda)$ onde $\rho : \pi_1(\mathcal{C}, c_0) \rightarrow \mathbb{C}$ é uma representação unidimensional e unitária do grupo fundamental de (\mathcal{C}, c_0) . Para isto, consideremos primeiro o que acontece se fizermos uma mudança de malha homotópica. Se \mathcal{M} e \mathcal{N} são duas malhas distintas tem-se:

$$\begin{aligned} \Lambda_{ab}(\lambda) &= [\mathcal{M}(a)] * [\lambda] * [\overline{\mathcal{M}}(b)] \\ &= [\mathcal{N}(a)] * \left[[\overline{\mathcal{N}}(a)] * [\mathcal{M}(a)] \right] * [\lambda] * \left[[\overline{\mathcal{M}}(b)] * [\mathcal{N}(b)] \right] * [\overline{\mathcal{N}}(b)] \\ &= [\mathcal{N}(a)] * [\mu * \lambda * \nu] * [\overline{\mathcal{N}}(b)] = \Lambda_{ab}(\mu * \lambda * \nu). \end{aligned}$$

Assim, sob uma mudança de malha homotópica tem-se:

$$\mathcal{K}(a, t_0; b, t_1) = \sum_{\lambda} \rho(\lambda) \mathcal{K}^{\Lambda_{ab}(\lambda)} \longmapsto \sum_{\lambda} \rho(\lambda) \mathcal{K}^{\Lambda_{ab}(\mu * \lambda * \nu)}.$$

onde $\lambda, \mu, \nu \in \pi_1(\mathcal{C}, c_0)$ e μ, ν são fixos, determinados pela escolha de malha.

Note, portanto, que a mudança de malha consiste em uma possível reordenação das amplitudes parciais de cada classe de homotopia. De modo que cada peso ficaria associado à amplitude parcial de uma classe de homotopia distinta. Isto não é um problema. Denotando $\mu * \lambda * \nu = \beta$, tem-se

$$\lambda = \mu^{-1} * \beta * \nu^{-1}.$$

Se ρ é uma representação unidimensional, então podemos escrever:

$$\begin{aligned} \mathcal{K} &= \sum_{\lambda} \rho(\lambda) \mathcal{K}^{\Lambda_{ab}(\beta)} = \sum_{\lambda} \rho(\mu^{-1} * \beta * \nu^{-1}) \mathcal{K}^{\Lambda_{ab}(\beta)} \\ &= \rho(\mu^{-1} * \nu^{-1}) \sum_{\beta} \rho(\beta) \mathcal{K}^{\Lambda_{ab}(\beta)} = \rho(\mu^{-1} * \nu^{-1}) \mathcal{K} \end{aligned}$$

Donde obtém-se

$$|\rho(\mu^{-1} * \nu^{-1}) \mathcal{K}| = |\rho(\mu^{-1} * \nu^{-1})| |\mathcal{K}| = |\mathcal{K}|.$$

Para verificar a segunda propriedade, usamos o fato de Λ_{ab} ser compatível com a operação $*$. Tem-se:

$$\begin{aligned} \Lambda_{ac}(\gamma) &= [\mathcal{M}(a)] * [\gamma] * [\overline{\mathcal{M}}(c)] \\ &= [\mathcal{M}(a)] * [\alpha] * [\overline{\mathcal{M}}(b)] * [\mathcal{M}(b)] * [\beta] * [\overline{\mathcal{M}}(c)] \\ &= \Lambda_{ab}(\alpha) * \Lambda_{bc}(\beta) \end{aligned}$$

para $\alpha, \beta \in \pi_1(\mathcal{C}, c_0)$ tal que $\alpha * \beta = \gamma$. Com isto, tem-se:

$$\begin{aligned} \mathcal{K}(a, t_0; c, t_2) &= \sum_{\gamma} \rho(\gamma) \mathcal{K}^{\Lambda_{ac}(\gamma)} = \sum_{\gamma} \rho(\gamma) \sum_{\alpha * \beta = \gamma} \int_{\mathcal{C}} \mathcal{K}^{\Lambda_{ab}(\alpha)} \mathcal{K}^{\Lambda_{bc}(\beta)} db \\ &= \sum_{\gamma} \sum_{\alpha * \beta = \gamma} \rho(\alpha * \beta) \int_{\mathcal{C}} \mathcal{K}^{\Lambda_{ab}(\alpha)} \mathcal{K}^{\Lambda_{bc}(\beta)} db \\ &= \int_{\mathcal{C}} \left[\sum_{\alpha} \rho(\alpha) \mathcal{K}^{\Lambda_{ab}(\alpha)} \right] \left[\sum_{\beta} \rho(\beta) \mathcal{K}^{\Lambda_{bc}(\beta)} \right] db \\ &= \int_{\mathcal{C}} \mathcal{K}(a, t_0; b, t_1) \mathcal{K}(b, t_1; c, t_2) db. \end{aligned}$$

Onde na passagem do segunda para a terceira igualdade usamos que, pela definição do

propagador de Feynman — definido para caminhos pertencentes a uma única classe de homotopia —, o propagador de a a c pode ser dividido naquele de a a b multiplicado pelo de b a c e somando-se sobre todos os pontos b . O fato de Λ_{ab} ser compatível com $*$ garante que este produto continue na mesma classe de homotopia. Além disso, na passagem da segunda para a terceira linha usamos que a soma com a condição $\alpha * \beta = \gamma$ torna-se uma soma livre sobre α e β se somarmos sobre γ .

Fica estabelecido, portanto que é possível descrever sistemas de partículas indistinguíveis pela formulação de Feynman se escrevermos o propagador como:

$$\mathcal{K}(a, t_0; b, t_1) = \sum_{\lambda} \rho(\lambda) \mathcal{K}^{\Lambda_{ab}(\lambda)}. \quad (2.6)$$

onde $\lambda \in \pi_1(\mathcal{C}, c_0)$, Λ_{ab} é o mapa induzido por uma escolha arbitrária de malha homotópica \mathcal{M} e ρ é uma representação unidimensional e unitária de $\pi_1(\mathcal{C}, c_0)$. É interessante notar que o trabalho de Laidlaw e DeWitt [1] precede o de Leinaas e Myrheim [2], no entanto, ele foi utilizado para tratar especificamente de partículas em duas dimensões apenas por Wu [5] alguns anos depois da publicação do trabalho de Leinaas e Myrheim.

Com este resultado fica claro que a topologia do espaço de configurações tem um papel a desempenhar na descrição quântica do sistema. Para partículas em três dimensões, o espaço de configurações adequado não é simplesmente conexo. O grupo fundamental é isomorfo a \mathcal{S}_n , o grupo simétrico sobre n elementos. Em particular, este grupo possui apenas duas representações unidimensionais, dadas por:

$$\rho^S(\sigma) = 1 \quad \text{e} \quad \rho^A(\sigma) = \text{sgn } \sigma$$

para $\sigma \in \mathcal{S}_n$. Deste modo, existem duas descrições distintas de partículas em três dimensões, uma para cada representação:

$$\begin{aligned} \mathcal{K}(a, t_0; b, t_1) &= \sum_{\lambda} \rho^S(\lambda) \mathcal{K}^{\Lambda_{ab}(\lambda)}, \\ \mathcal{K}(a, t_0; b, t_1) &= \sum_{\lambda} \rho^A(\lambda) \mathcal{K}^{\Lambda_{ab}(\lambda)}. \end{aligned}$$

A representação ρ^S está associada à estatística de Bose-Einstein, pois independente de qualquer permutação, a fase relativa dos caminhos é sempre igual a 1. Por outro lado, ρ^A está associada à estatística de Fermi-Dirac, pois qualquer caminho que seja imagem, sob Λ_{ab} , de um loop associado a uma permutação ímpar, como a permutação de duas partículas adjacentes, por exemplo, adquire uma fase relativa igual a -1 . Para partículas em duas dimensões, o grupo fundamental é isomorfo ao grupo de tranças e existem mais representações possíveis. Já discutimos brevemente como ele afeta o comportamento das partículas na seção anterior e a seguir iremos estudá-lo mais a fundo.

Um último comentário antes de passarmos para a próxima seção. Foi dito que a formulação de Feynman fornece uma ligação direta com teorias de campos topológicas. Esta ligação se dá, essencialmente porque a representação do grupo fundamental que pesa todas as amplitudes parciais pode ser escrita de modo a modificar a lagrangiana do sistema com um termo de origem topológica, já que o grupo fundamental está diretamente ligado à topologia do sistema. Esta linha de raciocínio pode ser encontrada nas referências [26, 27], bem como no artigo de revisão [15].

2.3.3 - Representações e o grupo de tranças

O grupo de tranças de n elementos, note n é o número de elementos a serem trançados, não a ordem do grupo, \mathcal{B}_n pode ser obtido a partir de $n - 1$ geradores σ_i , $1 \leq i \leq n - 1$. Cada gerador corresponde a trançar um fio numerado por n com o fio adjacente, $n + 1$ no sentido anti-horário. Este processo pode ser representado graficamente de maneira bem simples e bastante muito convenientemente. A ação dos geradores esta ilustrada à esquerda na figura 2.10 a seguir.

Esta representação gráfica dos elementos do grupo é útil extremamente conveniente porque podemos interpretar os fios como as linhas de mundo de partículas, de modo que a ação dos elementos de \mathcal{B}_n é interpretada diretamente como sua ação na trajetória das partículas. Ou seja, são precisamente os caminhos, ou pelo menos as classes de equivalência de caminhos sobre as quais as amplitudes parciais do propagador de Feynman são calculadas. Neste sentido, a formulação de Feynman fornece, literalmente, uma visualização de sistemas de partículas bidimensionais. Neste esquema de representação, todos os elementos do grupo são convencionados como trançando no sentido anti-horário e suas inversas, no sentido horário. A ação de σ_i^{-1} está ilustrada à direita na figura a seguir.

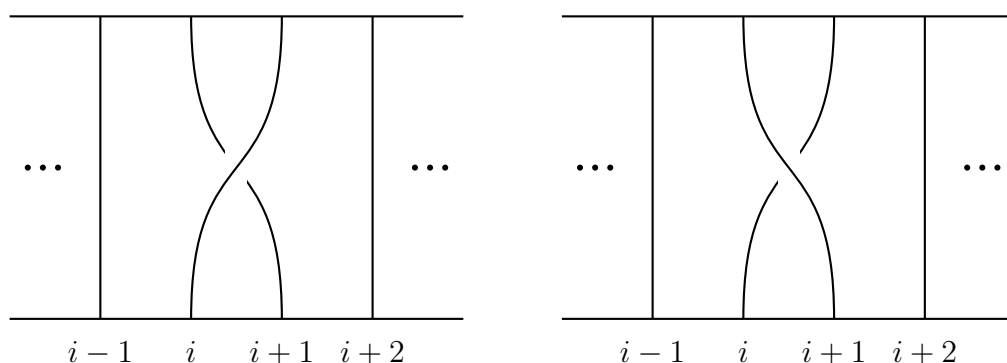


Figura 2.10: Ação do gerador σ_i do grupo de tranças à esquerda e seus inversos, à direita.

Os geradores satisfazem à duas condições chamadas condições de Artin, dadas por:

$$\begin{aligned}\sigma_i \sigma_{i+1} \sigma_i &= \sigma_{i+1} \sigma_i \sigma_{i+1}, \\ \sigma_i \sigma_j &= \sigma_j \sigma_i \text{ se } |i - j| \geq 2.\end{aligned}$$

A segunda condição meramente indica que trançar dois pares de fios disjuntos produz o mesmo resultado independentemente da ordem em que são trançados. A primeira condição é conhecida como relação de Yang-Baxter. Para visualizá-la, convencionamos que o produto de dois elementos do grupo de tranças é representado pela justaposição vertical de seus diagramas. Isto é o produto $\sigma_2 \sigma_1$ é representado pelo diagrama representando σ_2 acima daquele representando σ_1 . Note que esta convenção significa que os elementos do grupo atuam pela direita. A relação de Yang-Baxter está representada visualmente na figura abaixo.

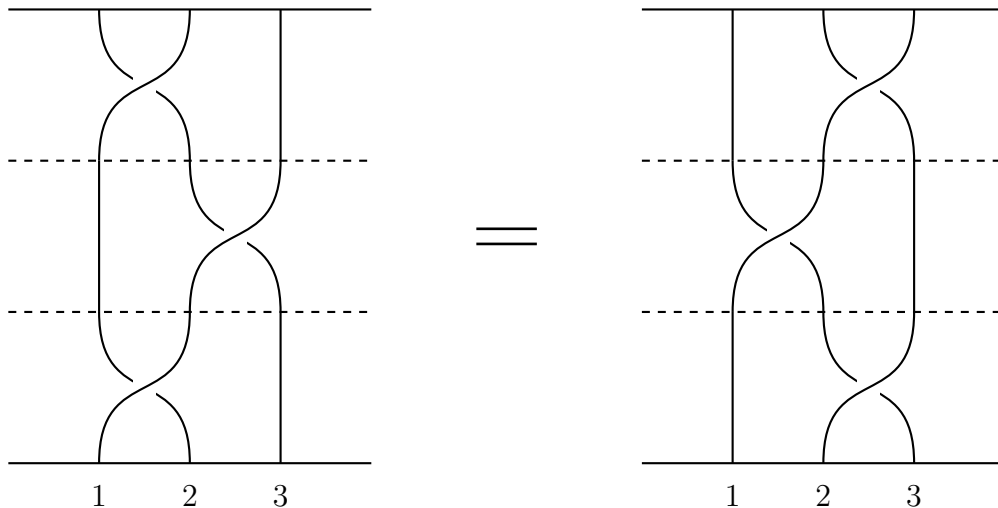


Figura 2.11: Representação visual da relação de Yang-Baxter.

A partir destas relações vê-se prontamente que o grupo de tranças é mais complexo que o grupo de permutações. Mais explicitamente, grupo simétrico de ordem n pode ser representado por um conjunto de $n - 1$ geradores, τ_i , $1 \leq i \leq n - 1$, com a condição de que todos satisfaçam $\tau_i^2 = 1$. Este conjunto em particular de geradores são as transposições de elementos adjacentes, isto é: $\tau_1 = (i \ i + 1)$, quando representada como um ciclo. Inclusive, o grupo simétrico de ordem n pode ser obtido a partir do grupo de tranças em n fios impondo-se a condição $\sigma_i^2 = 1$ para todo i sobre geradores do grupo de tranças. Isto evidencia a grande diferença entre estes dois grupos: o grupo de permutações é finito enquanto o de tranças não o é. Veremos as consequências disso mais a frente.

Agora que o \mathcal{B}_n está definido, vejamos quais são suas representações unidimensionais e unitárias. Pela definição, o mapa $\rho : \mathcal{B}_n \rightarrow \mathbf{U}(1)$ será uma tal representação de \mathcal{B}_n se for bijetivo e:

$$\rho(\sigma_i \sigma_j) = \rho(\sigma_i) \rho(\sigma_j).$$

Então, aplicando-se ρ à relação de Yang-Baxter, obtém-se:

$$\rho(\sigma_i \sigma_{i+1} \sigma_i) = \rho(\sigma_i) \rho(\sigma_{i+1}) \rho(\sigma_i).$$

Mas

$$\rho(\sigma_i \sigma_{i+1} \sigma_i) = \rho(\sigma_{i+1} \sigma_i \sigma_{i+1}) = \rho(\sigma_{i+1}) \rho(\sigma_i) \rho(\sigma_{i+1}).$$

Note que como $U(1)$ é abeliano, obtemos:

$$\begin{aligned} \rho(\sigma_i) \rho(\sigma_{i+1}) \rho(\sigma_i) &= \rho(\sigma_{i+1}) \rho(\sigma_i) \rho(\sigma_{i+1}) \\ \implies [\rho(\sigma_i)]^2 \rho(\sigma_{i+1}) &= [\rho(\sigma_{i+1})]^2 \rho(\sigma_i) \\ \implies \rho(\sigma_i) &= \rho(\sigma_{i+1}). \end{aligned}$$

O fato de $U(1)$ ser abeliano automaticamente garante a segunda condição. O resultado acima implica que $\rho(\sigma_i) = \rho(\sigma_j)$ para todo $1 \leq i, j \leq n$. Representaremos o valor comum de ρ para todos os geradores por:

$$\rho(\sigma_i) = e^{i\theta}$$

onde $\theta \in [0, 2\pi]$ é arbitrário. Cada escolha de θ fornece uma representação de \mathcal{B}_n . O parâmetro θ é comumente chamado de **parâmetro estatístico**, **ângulo estatístico** ou simplesmente **estatística**. Relembre que já nos deparamos com este parâmetro anteriormente, denotado na ocasião por ν . Note a semelhança com o resultado obtido no processo de quantização na formulação de Schrödinger, equação 2.5. O fato do parâmetro estatístico poder tomar seu valor de um intervalo contínuo indica que em duas dimensões existe um contínuo entre bósons, $\theta = 0$, e férmions, $\theta = \pi$.

Agora, existem duas coisas interessantes a serem notadas. Primeiro, a diferença entre a atuação do grupo de tranças e do grupo de permutações sobre a trajetória das partículas. Em particular, a diferença devido à condição $\tau_i^2 = 1$ no grupo de permutações cujo efeito é ilustrado na figura abaixo. Lembre-se que configurações que diferem apenas por permutações de partículas estão identificadas, ou seja, representam o mesmo ponto.

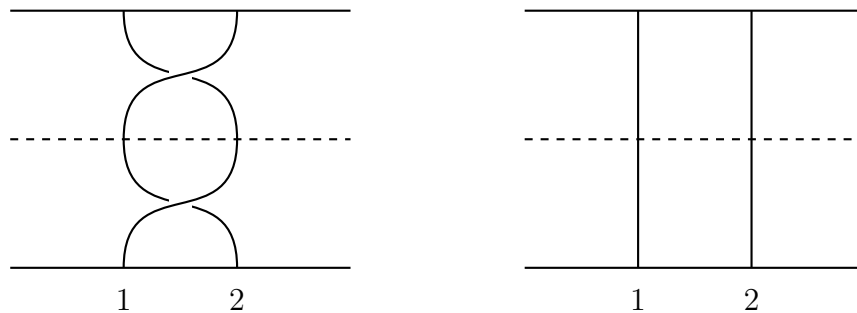


Figura 2.12: Comparação entre a aplicação de $\sigma_1^2 \neq 1$ e $\tau_1^2 = 1$.

A ação de ambos o grupo de tranças quando o grupo de permutações corresponde

ao processo físico de trocar duas partículas de lugar. A diferença, como pode-se ver na imagem é que, do ponto de vista de permutações simples, isto é, do grupo simétrico, permutar as mesmas duas partículas duas vezes é equivalente a não fazer nada, enquanto do ponto de vista do grupo de tranças, permutar as partículas duas vezes, trança as linhas de mundo das partículas e não equivale a não fazer nada. Deste modo, mesmo que o estado final após as duas operações sejam o mesmo, o ato de trançar partículas possui um efeito dinâmico que não está presente na permutação simples.

Mais formalmente, considere um loop γ situado em uma dada configuração c_0 . Se evoluirmos o sistema segundo $\gamma^2 = \gamma * \gamma$, segundo a equação 2.5 obteremos o seguinte:

$$\Psi(c_0) \mapsto \rho([\gamma * \gamma])\Psi(c_0).$$

onde ρ é uma representação do grupo fundamental do espaço de configurações do sistema em questão, $\pi_1(\mathcal{C}, c_0)$.

Agora, antes de investigarmos mais a fundo, vejamos o que significa ρ ser uma representação unitária. Uma representação é, essencialmente, um homomorfismo de grupos. Um homomorfismo de grupos é um mapa, R , entre um grupo $(\mathcal{G}, *)$ e (\mathcal{H}, \star) tal que:

$$R(g_1 * g_2) = R(g_1) \star R(g_2)$$

onde $*$ e \star são as operações de \mathcal{G} e \mathcal{H} , respectivamente. Para os espaços de Hilbert usuais da mecânica quântica, uma representação unitária e unidimensional de um grupo $(\mathcal{G}, *)$ é um homomorfismo entre \mathcal{G} e $U(1)$, onde $U(1)$ é o grupo de todos os números complexos de módulo igual a 1. Assim, podemos escrever ρ como:

$$\begin{aligned} \rho : \pi_1(\mathcal{C}, c_0) &\rightarrow U(1) \\ [\gamma] &\mapsto e^{i\theta([\gamma])} \end{aligned}$$

onde $\theta([\gamma]) \in \mathbb{R}$ é a menos de múltiplos de 2π , isto é $\theta([\gamma]) \in [2\pi n, 2\pi(n+1)]$ para algum $n \in \mathbb{Z}$. Alternativamente, poderíamos evidenciar o fator de 2π e escrever ρ como:

$$\rho([\gamma]) = e^{2\pi i\nu([\gamma])}$$

onde $\nu([\gamma]) \in [n, n+1]$ para algum $n \in \mathbb{Z}$. Do fato de uma representação ser um homomorfismo decorre que:

$$\rho([\gamma * \gamma]) = [\rho([\gamma])]^2 = e^{i2\theta([\gamma])}.$$

Agora, consideremos o caso de partículas em três dimensões. Neste caso, o grupo fundamental é isomorfo ao grupo simétrico de ordem n , onde n é o número de partículas

do sistema. Como discutido anteriormente, para qualquer n , o grupo simétrico possui duas representações unidimensionais, a completamente simétrica dada por

$$\rho^S([\gamma]) = 1$$

e a completamente antissimétrica, dada por

$$\rho^A([\gamma]) = \text{sgn}([\gamma]).$$

Podemos escrever a representação completamente simétrica na forma exponencial simplesmente definindo $\theta([\gamma]) = \nu([\gamma]) = 0$ para todo loop γ . Por outro lado, o sinal de uma permutação define um mapa para Z_2 , de modo que $\text{sgn}([\gamma]) \in \{-1, +1\}$. Por sua vez, o grupo Z_2 é homomorfo ao grupo $\mathbb{Z}/2\mathbb{Z}$. Este grupo definido identificando-se todos os números pares e todos os ímpares respectivamente. Escrevemos

$$\mathbb{Z}/2\mathbb{Z} = \{n \bmod 2, n \in \mathbb{Z}\} = \{0, 1\}.$$

A operação em $\mathbb{Z}/2\mathbb{Z}$ é a operação de soma usual. Assim, podemos escrever a representação completamente antissimétrica em forma exponencial definindo:

$$\theta([\gamma]) = \begin{cases} 0 & \text{se } \text{sgn}([\gamma]) = 1, \\ \pi & \text{se } \text{sgn}([\gamma]) = -1. \end{cases}$$

Com estas definições, temos que

$$[\rho^S([\gamma])]^2 = e^{2\pi i 0} e^{2\pi i 0} = 1$$

e

$$[\rho^A([\gamma])]^2 = e^{i\pi} e^{i\pi} = e^{2\pi i} = 1.$$

Repare a semelhança deste resultado com a descrição feita na seção 2.2.1, onde associamos uma fase proporcional ao ângulo no qual uma partícula foi rotacionada ao redor da outra. Uma permutação simples corresponde a um ângulo de π e uma fase $e^{i\pi} = -1$. Permutarmos as mesmas partículas duas vezes corresponde a um ângulo de 2π , três vezes, 3π e assim por diante. Repare como esta descrição é reproduzida exatamente pela representação completamente antissimétrica. Ela é reproduzida também pela representação completamente simétrica se ao invés se escrevermos

$$\rho([\gamma]) = e^{\pi i \nu([\gamma])}.$$

Temos liberdade para fazer isto pois $\theta([\gamma])$ é definido a menos de múltiplos de 2π como

dito. Se definirmos um mapa $m : \pi_1(\mathcal{C}^3, c_0) \rightarrow \mathbb{R}$ como

$$m([\gamma]) = \begin{cases} 0 & \text{se } \text{sgn}([\gamma]) = 1, \\ \pi & \text{se } \text{sgn}([\gamma]) = -1. \end{cases}$$

podemos unificar as duas representações como

$$\rho^\nu([\gamma]) = e^{im([\gamma])\nu} \quad (2.7)$$

e $\nu = 0$ para a representação completamente simétrica e $\nu = 1$ para representação completamente simétrica. Escrito desta maneira, ν é o parâmetro estatístico ou a estatística das partículas. Vê-se claramente que $\nu = 0$ corresponde a bósons e $\nu = 1$ a férmions. E m classifica a trajetória das partículas segundo o grupo fundamental. Permutar as mesmas partículas duas vezes é uma permutação par, três, é uma permutação ímpar. Permutar as mesmas partículas três vezes corresponderia a uma fase de $e^{3\pi i} = e^{\pi i} = -1$. Note que a vantagem de usar o mapa m é que basta determinar a paridade da permutação ao invés de determinar todos os ângulos de rotação, o que pode ser complicado para muitas partículas. Outra vantagem do mapa m é que podemos defini-lo de modo análogo para sistemas de partículas em duas dimensões e assim, obter uma descrição essencialmente unificada.

Apesar da construção do mapa m para partículas em duas dimensões ser mais complicada, a ideia subjacente é a mesma. Primeiro, note que para partículas com estatística θ , permutar duas partículas uma ao longo de um loop γ tal que $[\gamma] = \sigma_i$, um vez transforma a função de onda da seguinte forma:

$$\Psi(c_0) \mapsto \rho^\theta([\gamma])\Psi(c_0) = \rho^\theta(\sigma_i)\Psi(c_0) = e^{i\theta}\Psi(c_0).$$

Similarmente:

$$\Psi(c_0) \mapsto \rho^\theta([\gamma * \gamma])\Psi(c_0) = [\rho^\theta(\sigma_i)]^2\Psi(c_0) = e^{2\theta i}\Psi(c_0).$$

Agora, se considerarmos um caminho $\gamma \in \sigma_i\sigma_j$, temos:

$$\Psi(c_0) \mapsto \rho^\theta([\gamma])\Psi(c_0) = \rho^\theta(\sigma_i)\rho^\theta(\sigma_j)\Psi(c_0) = e^{2\theta i}\Psi(c_0).$$

Por fim, se $\gamma \in \sigma_i^{k_i}\sigma_j^{k_j}$, $k_i, k_j \in \mathbb{Z}$, obtém-se:

$$\Psi(c_0) \mapsto \rho^\theta([\gamma])\Psi(c_0) = [\rho^\theta(\sigma_i)]^{k_i}[\rho^\theta(\sigma_j)]^{k_j}\Psi(c_0) = e^{(k_i+k_j)\theta i}\Psi(c_0).$$

A partir destes resultados constata-se que a fase ganha na permutação de anyons com estatística θ depende apenas do número total de permutações realizadas. Como σ_i são

geradores do grupo de tranças, qualquer elemento do grupo pode ser escrito como uma palavra formada por estes geradores. Isto é, se \mathbf{G} é o conjunto de todos os geradores de \mathcal{B}_n e \mathbf{G}^{-1} o conjunto de todos os inversos dos geradores, então qualquer trançamento $\tau \in \mathcal{B}_n$ pode ser escrito como:

$$\tau = \prod_{i=1}^s g_i, \quad g \in \mathbf{G} \cup \mathbf{G}^{-1}.$$

para algum $s \in \mathbb{N}^* = \mathbb{N} - \{0\}$. Note que podemos pensar uma trança τ como uma sequência (g_1, \dots, g_s) , possivelmente infinita de elementos de $\mathbf{G} \cup \mathbf{G}^{-1}$. Assim, definimos um mapa auxiliar $p : \mathbf{G} \cup \mathbf{G}^{-1} \rightarrow \mathbb{Z}$ por:

$$p(g) = \begin{cases} +1 & \text{se } g \in \mathbf{G}, \\ -1 & \text{se } g \in \mathbf{G}^{-1} \end{cases}$$

e com este mapa definimos $m : \mathcal{B}_n \rightarrow \mathbb{Z}$ por

$$m(\tau) = \sum_{g \in \tau} p(g).$$

Note que por definição, o produto de duas tranças $\tau_1, \tau_2 \in \mathcal{B}_n$ pode ser representando simplesmente como a concatenação das sequências, isto é:

$$\tau_1 \tau_2 = (g_1, \dots, g_{s_1})(h_1, \dots, h_{s_2}) = (g_1, \dots, g_{s_1}, h_1, \dots, h_{s_2}).$$

Deste modo

$$m(\tau_1 \tau_2) = \sum_{g \in \tau_1} p(g) + \sum_{h \in \tau_2} p(h) = m(\tau_1) + m(\tau_2).$$

Ou seja, m define um homomorfismo entre \mathcal{B}_n e \mathbb{Z} . Por fim, usamos o mapa m para escrever a representação ρ como:

$$\rho^\theta([\gamma]) = e^{im([\gamma])\theta}. \quad (2.8)$$

Lembrando que $\pi_1(\mathcal{C}_n^2, c_0) \cong \mathcal{B}_n$ de modo que cada classe de equivalência de $[\gamma]$ é associada a uma única trança de \mathcal{B}_n . Além disso, como $\theta \in [0, 2\pi]$ e $m([\gamma]) \in \mathbb{Z}$, o produto $m([\gamma])\theta$ pertence ao intervalo $[2\pi n, 2\pi(n+1)]$ para algum $n \in \mathbb{Z}$ como requerido por ρ ser uma representação unidimensional e unitária. Apesar de não termos o feito, podemos evidenciar o fator de π contribuído por cada permutação na expressão acima, basta redefinir m para que o resultado seja multiplicado por π e θ como pertencente ao intervalo $[0, 2]$. Deste modo a representação do grupo fundamental para sistemas em duas dimensões é escrita de modo completamente análoga àquela em três dimensões.

Escrito desta forma, fica nítido como anyons são uma generalização do comportamento estatístico de partículas em três dimensões. Desconsiderando os detalhes dos mapas m

em duas e três dimensões, a imagem de ambos está contida no conjunto dos inteiros e são usados essencialmente para contar quantas vezes cada partícula foi permutada ao redor de todas as outras. Além disso, tem-se que em três dimensões, o parâmetro estatístico é restrito a $\theta = 0$ e $\theta = 1$, enquanto que em duas $\theta \in [0, 2\pi]$. Fica estabelecido, portanto, que existe um contínuo de estatísticas entre as estatísticas bosônica e fermiônica para partículas em sistemas bidimensionais.

Investigamos, agora, a propriedade mais exótica de anyons, chamada de **fusão**. Começamos com um sistema de três partículas em posições 1, 2 e 3 e queremos trançar permutar a partícula na posição 3 com as partículas nas posições 1, mantendo aquela na posição 2 “fixa”. O processo é tal que trocamos as partículas originalmente nas posições 1 e 3, passando por 2 e depois, retornamos a partícula originalmente em 2, agora em 3, à sua posição original, permutando-a com aquela agora em 2. O resultado é que “giramos” as partículas originalmente em 1 e 2 uma em relação à outra, pois aquela que estava originalmente em 1 agora está à esquerda.

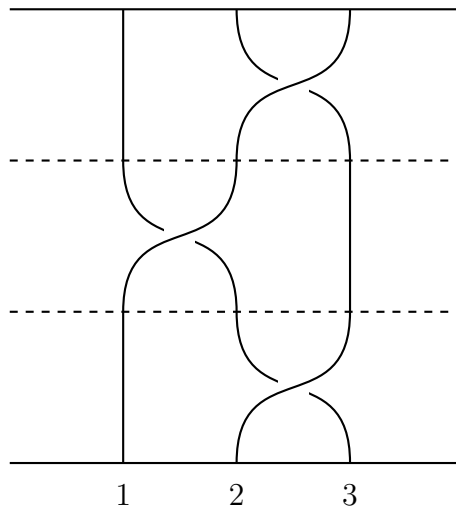


Figura 2.13: Trançamento de três partículas segundo $\tau = \sigma_2\sigma_1\sigma_2$.

Este procedimento corresponde ao trançamento $\tau = \sigma_2\sigma_1\sigma_2$ mostrado na figura acima. Supondo que estas partículas possuem estatística θ , usando a representação 2.8 vê-se que este processo altera a função de onda segundo:

$$\Psi(c_0) \mapsto e^{im(\tau)\theta}\Psi(c_0) = e^{3i\theta}.$$

Este resultado é curioso pois, se realizássemos exatamente o mesmo processo em três dimensões, trocando os σ_i pelas transposições τ_i , segundo a representação 2.7, teríamos:

$$\Psi(c_0) \mapsto e^{im(\tau)\nu}\Psi(c_0) = e^{\pi i\nu}.$$

Que é o mesmo resultado que obteríamos se trocássemos duas partículas adjacentes uma

vez. Note, também que, caso permutássemos partículas adjacentes em duas dimensões, a fase adquirida pela função de onda seria apenas $e^{i\theta}$. Existe uma diferença entre as duas situações.

Para entender a origem desta diferença, suponha que as duas partículas, originalmente nas posições 1 e 2 formassem um estado ligado, ou, simplesmente que a distância entre as posições 1 e 3 e 2 e 3 são muito maiores que a distância entre 1 e 2. Neste caso, o processo corresponderia a permutar a partícula em 3 com a partícula composta 1 – 2 e então permutar as partículas originalmente em 1 e 2, efetivamente rotacionando a partícula composta 1 – 2 por um ângulo de π . Em particular, no processo realizado, esta rotação é feita no sentido horário. Agora, note que podemos escrever

$$e^{3i\theta} = e^{(4-1)i\theta} = e^{i4\theta} e^{-i\theta}.$$

A última expressão por sua vez pode ser interpretada como uma partícula de estatística θ sendo permutada com outra no sentido horário. E, então, uma partícula de estatística 4θ sendo permutada com outra. Assim, poderíamos entender que, permutar uma partícula de estatística θ com uma partícula composta por duas, também de estatística θ acrescenta uma fase de 4θ . Neste sentido, mesmo que as partículas não sejam idênticas, permutá-las ainda acrescenta uma fase na função de onda determinada por um tipo de fusão das estatísticas de cada partícula.

A rigor, para obter uma descrição adequada disto, deveríamos ter partido de várias partículas, separadas em tipos, com partículas do mesmo tipo sendo indistinguíveis e partículas de tipos distintos, distinguíveis e então realizar o processo de quantização. No entanto esta análise é suficiente para ilustrar o fenômeno de fusão de anyons. Usando θ para nomear um tipo de anyon, poderíamos escrever, heurísticamente:

$$\theta \times \theta = 4\theta$$

Onde a operação \times consiste na fusão destes anyons, isto é, em aproximá-los de modo que a distância relativa entre eles seja muito menor que todas as outras. Ou formar um estado ligado, no entanto a primeira exigência é suficiente e mais geral. Assim, esta partícula composta seria um anyon de outra espécie, com estatística igual a 4θ .

Se quiséssemos determinar o resultado de $\theta \times \theta \times \theta$, poderíamos utilizar o mesmo tipo de análise, permutando um anyon de tipo θ com uma partícula composta de três anyons deste mesmo tipo. Resultados sobre isto estão dispostos em [15] e notamos que a análise realizada aqui é essencialmente heurística e não rigorosa. O último assunto a tratamos neste capítulo, também está disposto em [15] assim como em [26] e trata do conceito de anyons não abelianos.

Este tipo de anyons, as vezes chamado de “nonabelion” surge quando consideramos partículas se movendo em espaços mais complicados que o espaço euclidiano \mathbb{R}^n . Lembrem-

se que desde o início, supusemos que as partículas se moviam em \mathbb{R}^n , no entanto, a primeira definição do espaço de configurações foi:

$$\mathcal{C}_n^d = \frac{X^n - \Delta}{\mathcal{S}_n}$$

onde d era a dimensionalidade do espaço X e n o número de partículas. Poderíamos muito bem ter escolhido partículas restritas a uma esfera, ou um toro, por exemplo. Como discutido em [26], se escolhermos $X = \mathcal{S}^2$, a esfera, então surgem condições extras sobre os geradores do grupo de tranças e os anyons só podem ter estatísticas racionais.

Além disso, argumentamos, na seção sobre quantização na formulação de Schrödinger, seção 2.3.1, que não havia, no artigo de Leinaas e Myrheim [2] uma justificativa sobre porque escolher apenas espaços de Hilbert unidimensionais. Segundo Lerda [26], se permitirmos funções de onda multidimensionais, obtemos o que ele se referiu como anyons generalizados. Estes tipos de anyons são conhecidos como anyons não abelianos na literatura contemporânea. O diferencial deste tipo de partículas é que o sistema evolui com representações multidimensionais do grupo fundamental, o que pode levar a comportamentos muito mais ricos, especialmente no que diz respeito a fusão.

É possível descrever as regras de fusão de anyons não abelianos de maneira muito similar a como é descrito o processo de “soma” de spin, quando se considera uma partícula composta por outras partículas com spin [15]. Neste caso, podemos escrever, para anyons não abelianos de tipos ψ e ϕ :

$$\psi \times \phi = \sum_{\rho} N_{\rho}^{\psi\phi} \rho$$

onde ρ é outro tipo de anyon não abeliano e o somatório é feito sobre todos os tipos possíveis de serem obtidos a partir da fusão de anyons dos tipos ψ e ϕ . A riqueza de fenômenos para anyons não abelianos advém do fato de que representações unitárias de dimensões superiores não são, necessariamente, abelianas. De modo que permutar dois anyons em uma dada ordem pode gerar uma fase diferente daquela gerada pela permutação na ordem contrária.

A descrição de anyons não abelianos é mais comumente feita usando-se ferramentas de teoria de campos, em particular teoria de campos topológica ou conformes. Não abordaremos estes assuntos aqui, o próximo capítulo é dedicado à descrição de anyons por meio de categorias. Notamos, no entanto, que por descreverem o mesmo fenômeno, todas estas descrições são equivalentes, ao menos aproximadamente. Discutiremos mais sobre isto adiante. Ambas as descrições via teoria de campos topológica e teoria de campos conformes aparecem principalmente em trabalhos tratando de modelos específicos para a realização experimental ou utilização de anyons [15, 34, 35].

Por fim, um comentário sobre o porquê de se pesquisar anyons. O principal motivo presente na literatura é a possível utilização de anyons para construir um computador quântico.

tico [15]. Como anyons não abelianos evoluem segundo representações multi-dimensionais do grupo fundamental, pode-se usar a permutação destes anyons para realizar operações no sistema quântico, como rotações, por exemplo. Estas operações por sua vez podem ser usadas para realizar computações com qubits, a versão quântica do bit de computador. A vantagem de usar anyons para isto, no entanto é que, como as operações são de natureza topológica, elas são robustas. Por exemplo, permutar duas partículas por caminhos que pertençam à mesma classe de homotopia irá implementar a mesma operação, eliminando possíveis imprecisões no processo. Do mesmo modo, um dado estado em um sistema de anyons só pode ser alterado se houver uma permutação de anyons, um processo que dificilmente será realizado por flutuações ambientais. Sistemas quânticos de anyons não são completamente imunes a erros, no entanto. Apesar disso, um computador quântico feito utilizando-se estas partículas seria mais robusto que computadores quânticos de outros tipos e um grande avanço.

Capítulo 3

Sobre categorias

Neste capítulo passaremos a tratar de um modo particular de descrever anyons matematicamente por meio de **teoria de categorias**. Começamos com uma breve história sobre o surgimento de teoria de categorias tratando, em particular, a ideia central que deu origem ao ramo. Em seguida, usamos esta breve história para justificar o porquê de usar tal teoria para descrever anyons e quais são as vantagens e desvantagens desta descrição.

Teoria de categorias pode ser descrita como uma vista panorâmica da matemática [36]. De longe, os detalhes tornam-se invisíveis, porém, padrões que antes não eram possíveis de se observar tornam-se nítidos. Alternativamente, poderíamos dizer, em primeira aproximação, que teoria de categorias é o estudo de álgebras abstratas de funções [37]. Do mesmo modo que teoria de grupos é a abstração da ideia de um sistema de permutações de um conjunto ou das simetrias de um objeto geométrico, teoria de categorias surge a partir da ideia de um sistema de funções entre objetos.

Lane [38], um dos fundadores da teoria escreve que: teoria de categorias surge da constatação que muitas propriedades de sistemas matemáticos podem ser unificadas e simplificadas em uma apresentação usando diagramas de flechas, como exemplificado abaixo. Cada flecha corresponde a uma função $f : X \rightarrow Y$ no sentido usual, isto é, X e Y são conjuntos e f atribui a cada $x \in X$ um valor $f(x) \in Y$.

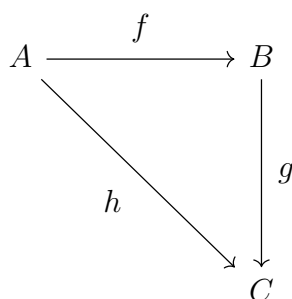


Figura 3.1: Exemplo de um diagrama de flechas.

Dizemos que o diagrama **comuta** se $h = g \circ f$. Pensamos então a operação de composição como um tipo de produto entre funções e a teoria de categoria, então, considera as “álgebras abstratas” que surgem deste produto. Como veremos, esta teoria e estes diagramas se aplicam a construções mais gerais do que funções entre conjuntos, de modo que

podemos pensar estas álgebras como sendo compostas por objetos A, B, C, \dots e mapas, $f : A \rightarrow B, g : B \rightarrow C, \dots$ entre estes objetos tais que são fechados sob a operação de composição. Veremos a definição precisa adiante.

Neste sentido, ao invés de simples funções e conjuntos, podemos tomar os objetos A, B e C como sendo, por exemplo, espaços topológicos e f, g e h como sendo mapas contínuos entre estes espaços. Isto corresponderia à categoria de todos os espaços topológicos. Do mesmo modo, poderíamos tomar os objetos como sendo grupos e as funções como homomorfismos, formando a categoria de todos os grupos. Neste sentido, também, podemos dizer [37] que teoria de categorias foi inventada na tradição do *Erlanger Programm* de Felix Klein. Isto é como um modo de estudar sistemas matemáticos em termos de suas “transformações admissíveis” ou “mapas que preservam estrutura”.

É interessante notar que este ramo, que surgiu em 1945 com o artigo de Eilenberg e MacLane [16] intitulado, em tradução livre, “Teoria Geral de Equivalências Naturais” posteriormente revelou possuir conexões profundas com lógica e por volta de 1970 começou a ser aplicado em ciência da computação, linguística e filosofia. É surpreendente que existam aplicações em ramos tão diversos. Como se revelou depois, teoria de categorias serve como uma espécie de linguagem matemática universal, tal qual teoria de conjuntos e isto nos trás de volta à questão de “fornecer uma visão panorâmica da matemática”. Vejamos o que isto quer dizer com alguns exemplos.

Produto cartesiano

Primeiro, tratemos do produto cartesiano usual entre dois conjuntos X e Y , isto é, o conjunto $X \times Y$ que consiste de pares ordenados (x, y) onde $x \in X$ e $y \in Y$. As projeções $p : X \times Y \rightarrow X$ e $q : X \times Y \rightarrow Y$, dadas por $(x, y) \mapsto x$ e $(x, y) \mapsto y$ respectivamente, são tais que qualquer função $h : W \rightarrow X \times Y$ é determinada unicamente pelas compostas $p \circ h$ e $q \circ h$. Por outro lado, dadas funções $f : W \rightarrow X$ e $g : W \rightarrow Y$ existe uma única função $h : W \rightarrow X \times Y$, dada por $h(w) \mapsto (f(w), g(w))$ que faz o diagrama abaixo comutar.

$$\begin{array}{ccccc}
 & & W & & \\
 & g \swarrow & \downarrow h & \searrow f & \\
 X & \xleftarrow{q} & X \times Y & \xrightarrow{p} & Y
 \end{array}$$

Representando por $\langle f, g \rangle$ os pares de funções de algum conjunto qualquer W para X e Y , isto é, pares de funções $f : W \rightarrow X$ e $g : W \rightarrow Y$ podemos dizer que o par $\langle p, q \rangle$ é “universal” entre estes pares. Isto porque qualquer um destes pares é fatorado unicamente, através de h , pelo par $\langle p, q \rangle$. Esta propriedade, então, descreve o produto cartesiano unicamente a menos de bijeções. O diagrama acima, descreve igualmente bem

o produto cartesiano de conjuntos, de espaços topológicos e o produto direto de grupos. Basta trocar as funções pelas flechas adequadas para cada categoria: funções contínuas para espaços topológicos e homomorfismos para grupos.

Se representarmos por $[X, Y]$ o conjunto de todas as funções de $X \rightarrow Y$ e por $[\langle U, V \rangle, \langle X, Y \rangle]$ o conjunto de todos os pares de funções, $\langle f, g \rangle$, $f : U \rightarrow X$ e $g : V \rightarrow Y$, então a correspondência $h \mapsto p \circ h \times q \circ h = f \times g$ estabelece uma bijeção:

$$[W, X \times Y] \cong [\langle W, W \rangle, \langle X, Y \rangle].$$

A prova de que esta correspondência é simples. Já argumentamos sobre a sobrejeção: para qualquer par $\langle f, g \rangle$ definimos $h = f \times g$. A injetividade também é direta, dadas $h_1 = h_2$ isto implica que $p \circ h_1 = p \circ h_2$ e $q \circ h_1 = q \circ h_2$ logo os pares de funções, $\langle f_1, g_1 \rangle$ e $\langle f_2, g_2 \rangle$, fatorados por h_1 e h_2 são iguais. Esta bijeção, inclusive, é **natural** pois é definida “do mesmo modo” para todos os conjuntos W e todos os pares de conjuntos $\langle X, Y \rangle$. A ideia de naturalidade é central em teoria de categorias e a formalizaremos mais adiante.

Por último, a construção “produto cartesiano” é chamada de **funtor**, outro conceito central em teoria de categorias que também formalizaremos adiante. Ela é um funtor pois se aplica adequadamente tanto para conjuntos quanto para funções entre conjuntos. Duas funções $k : X \rightarrow X'$ e $l : Y \rightarrow Y'$ possuem, como produto cartesiano, a função

$$\begin{aligned} k \times l : X \times Y &\rightarrow X' \times Y' \\ (x, y) &\mapsto (k(x), l(y)). \end{aligned}$$

Note que o conjunto de um elemento, $\mathbf{1} = \{0\}$ serve como a identidade sob a operação de produto cartesiano, dado que existem bijeções

$$1 \times X \xrightarrow{\lambda} X \xleftarrow{\rho} X \times 1$$

dadas por $\lambda(0, x) = x$ e $\rho(x, 0) = x$. A verificação de que λ e ρ são, de fato, bijeções é direta.

Antes de passarmos para o próximo exemplo, então, reiteramos que o produto cartesiano como definido pela propriedade de que o par $\langle p, q \rangle$ seja “universal” e pelo diagrama mostrado, vale se os objetos são conjuntos, espaços topológicos, grupos ou outras estruturas. Neste sentido, esta definição é insensível aos detalhes das estruturas em consideração. Note que existe, uma diferenciação no sentido de que, se X, Y e W forem grupos, por exemplo, as funções são substituídas por homomorfismos. No entanto, uma vez que definamos categorias, esta diferenciação será absorvida na definição dos objetos de uma categoria e de suas flechas.

Monoides e grupos

Um **monoide** é uma estrutura

$$(M, \star, e)$$

onde M é um conjunto, \star é uma operação binária em R e e é um elemento selecionado de M , tais que:

$$(r \star s) \star t = r \star (s \star t) \quad \text{e} \quad e \star r = r = r \star e$$

para todo $r, s, t \in R$. Deste modo, a operação é associativa e o elemento nomeado, e , é a unidade desta operação. Usualmente omite-se o símbolo da operação e escreve-se simplesmente rs para o produto $r \star s$. Note que um monoide seria um grupo se cada elemento possuísse um inverso sob \star .

A definição acima é a definição algébrica usual de um monoide, presente em livros de álgebra. Podemos, porém, “categorificar” esta definição, isto é, expressá-la em termos de categorias. Para tal, considere a operação \star explicitamente como um mapa $\star : M \times M \rightarrow M$ e definimos o mapa $\eta : \mathbf{1} \rightarrow M$ que seleciona um elemento de M como a identidade. Tem-se: $\eta(0) = e$. Se $\text{id} : M \rightarrow M$ é o mapa identidade em M , então, usando estes mapas, os axiomas de um monoide equivalem a exigir que os seguintes diagramas comutem.

$$\begin{array}{ccc}
 M \times M \times M & \xrightarrow{\text{id} \times \star} & M \times M \\
 \star \times \text{id} \downarrow & & \downarrow \star \\
 M \times M & \xrightarrow{\star} & M
 \end{array}
 \qquad
 \begin{array}{ccccc}
 \mathbf{1} \times M & \xrightarrow{\eta \times \text{id}} & M \times M & \xleftarrow{\text{id} \times \eta} & M \times \mathbf{1} \\
 & \searrow \lambda & \downarrow \star & \swarrow \rho & \\
 & & M & &
 \end{array}$$

Note, $\mathbf{1} = \{0\}$ é o conjunto de um único elemento e λ e ρ são as bijeções do exemplo anterior. Dizer que estes diagramas comutam, corresponde exatamente a dizer que

$$\star \circ (\text{id} \times \star) = \star \circ (\star \times \text{id})$$

e

$$\star \circ (\eta \times \text{id}) = \lambda \quad \text{e} \quad \star \circ (\text{id} \times \eta) = \rho.$$

A primeira equação corresponde à associatividade de \star pois,

$$\begin{aligned}
 [\star \circ (\text{id} \times \star)](r, s, t) &= \star(r, s \star t) = r \star (s \star t), \\
 [\star \circ (\star \times \text{id})](r, s, t) &= \star(r \star s, t) = (r \star s) \star t
 \end{aligned}$$

para $s, r, t \in M$. A igualdade, então, informa que \star é associativa em M . As duas outras

equações correspondem às propriedades do elemento neutro pois:

$$\begin{aligned} [\star \circ (\eta \times \text{id})](0, s) &= \star(e, s) = e \star s, \\ [\star \circ (\text{id} \times \eta)](s, 0) &= \star(s, e) = s \star e. \end{aligned}$$

Como $\lambda(0, s) = s$ e $\rho(s, 0) = s$, tem-se justamente que $e \star s = s = s \star e$. Os diagramas acima poderiam ser reescritos com os elementos em vez dos conjuntos. Os mostraremos aqui apenas para ajudar a ilustrar como diagramas em teoria de categorias devem ser lidos.

$$\begin{array}{ccc} (s, r, t) & \xrightarrow{\text{id} \times \star} & (s, r \star t) \\ \star \times \text{id} \downarrow & & \downarrow \star \\ (s \star r, t) & \xrightarrow{\star} & (s \star r \star t); \end{array} \quad \begin{array}{ccc} (0, s) & \xrightarrow{\eta \times \text{id}} & (e \star s) \\ & \searrow \lambda & \downarrow \star \\ & & s; \end{array} \quad \begin{array}{ccc} (s \star e) & \xleftarrow{\text{id} \times \eta} & (s, 0) \\ \star \downarrow & & \nearrow \rho \\ s. & & \end{array}$$

Como dito anteriormente, um grupo é um monoide onde cada elemento possui uma inversa, deste modo, podemos apresentar uma versão categorificada dos axiomas de grupo se impusermos mais condições além da comutatividade dos diagramas acima. Seja $\delta : G \rightarrow G \times G$ o mapa diagonal dado por $g \mapsto (g, g)$. A estrutura (G, \star, e) será um grupo se, além de satisfazer as propriedades anteriores, existir um mapa $\zeta : G \rightarrow G$, tal que o seguinte diagrama comute.

$$\begin{array}{ccccc} G & \xrightarrow{\delta} & G \times G & \xrightarrow{\text{id} \times \zeta} & G \times G \\ \downarrow & & & & \downarrow \star \\ \mathbf{1} & \xrightarrow{\eta} & & & G. \end{array}$$

A flecha não nomeada é a única flecha possível entre G e $\mathbf{1}$, isto é, a que envia todos os elementos de G no único elemento de $\mathbf{1}$. O mapa ζ é o mapa que atribui a cada elemento do grupo seu elemento inverso. A comutatividade deste diagrama, então, implica que

$$g \star g^{-1} = e.$$

Note que este diagrama estabelece apenas a existência de uma inversa esquerda em G . Isto é suficiente, no entanto pois é possível provar que se G é um monoide e a condição para ζ é satisfeita, então toda inversa esquerda é também uma inversa direita. A saber, pela associatividade tem-se

$$(g \star h) \star j = g \star (h \star j)$$

Tomando $h = g^{-1}$ e $j = [g^{-1}]^{-1}$, isto é, j é a inversa da inversa de g . Então

$$\begin{aligned}(g \star g^{-1}) \star [g^{-1}]^{-1} &= g \star (g^{-1} \star [g^{-1}]^{-1}) \implies e \star [g^{-1}]^{-1} = g \star e \\ &\implies g^{-1} \star [g^{-1}]^{-1} = e = g^{-1} \star g.\end{aligned}$$

Ou seja, g^{-1} é uma inversa esquerda e direita de g . Grupos são um raro caso onde os axiomas usualmente usados na definição, aqueles que incluem inversas à esquerda e à direita, são mais fortes do que o necessário.

Agora, note que todo este tempo trabalhamos em uma categoria onde os objetos são conjuntos e as flechas são funções. Se mudarmos para uma categoria onde os objetos são espaços topológicos e as flechas são funções contínuas, obtemos, imediatamente e sem qualquer alteração, a definição de grupo topológico. Pois estes grupos são precisamente aqueles para os quais a operação de multiplicação e atribuição de inversa são contínuas. Do mesmo modo, se considerarmos uma categoria onde os objetos são variedades suaves e as flechas são mapas suaves, obtemos a definição de grupos de Lie. Assim, tanto grupos comuns, grupos topológicos e grupos de Lie podem ser descritos como grupos diagramáticos em uma categoria adequada. Do mesmo modo definem-se monoides topológicos e monoides suaves.

Finalizamos os exemplos aqui. A generalidade de categorias leva, naturalmente, a resultados muito mais profundos do que aqueles apresentados nestes exemplos porém, é necessário mais conhecimento sobre matemática, em particular álgebra abstrata e topologia do que está presente na formação básica de estudantes de física e do que está contido no escopo deste trabalho. Assim, passaremos agora à questão de porque usar tais categorias para tratar de anyons.

No capítulo anterior ficou estabelecido que as propriedades mais marcantes de anyons são sua estatística e a capacidade de se fundirem, gerando anyons de um novo tipo estatístico. Estas propriedades essencialmente tratam de como anyons se relacionam com outros. Fusão e estatística anyônica não significam nada se houver apenas um anyon no sistema. Do mesmo modo, teoria de categorias, como dito, é um ramo que foca em estudar como objetos se relacionam com outros, ou seja, em mapas entre objetos. Segundo Awodey [37] um *slogan* de teoria de categorias é:

São as flechas que realmente importam!

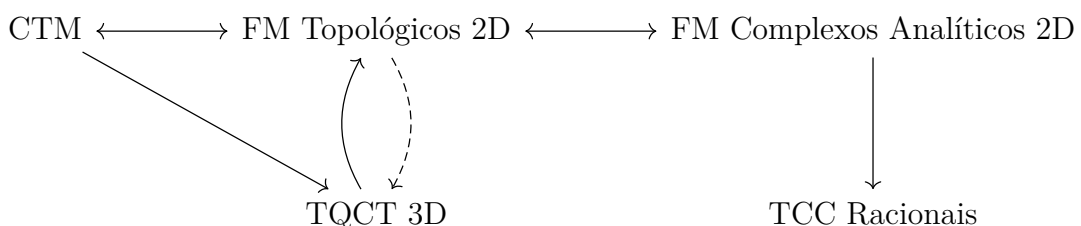
Sob está ótica, entendemos que o uso de categorias para descrever anyons é útil precisamente por capturar a estatística e as regras de fusão, que são as propriedades essenciais destas partículas. Em troca, no entanto, perde-se de vista a física por trás. Lembre-se que anyons não abelianos surgem quando consideramos que o espaço de configuração das

partículas é alguma variedade distinta do espaço euclidiano. A principal razão pela qual o espaço de configurações tomaria a forma de uma tal variedade é devido à interação entre as partículas. Deste modo, poderíamos tratar de aspectos relacionados à estatística e fusão de anyons não abelianos de modo relativamente simples usando a descrição categorial, no entanto, obter informações sobre os termos de interação à partir desta descrição é um problema bastante mais complicado. Tratamos desta questão por dois motivos.

O primeiro é que, como veremos, na descrição por categorias um sistema de anyons é especificado completamente ao se especificar os tipos de anyons envolvidos e suas regras de fusão. Em particular, isto significa que podemos criar regras de fusão de maneira arbitrária exceto por algumas limitações relativas ao tipo de categoria que descreve estas partículas. Esta limitação é similar ao requerimento de que o espaço de configurações de um sistema quântico seja um espaço de Hilbert; a menos desta limitação ele é essencialmente arbitrário. Então, se é possível criar a vontade regras de fusão, seria interessante que houvesse um método de traduzir estas regras e o correspondente sistema anyônico para uma outra formulação, uma que dá ênfase maior para a física que gera este sistema.

Segundo, anyons no mundo real são mais complicados do que discutido no capítulo anterior. Em particular, pelo simples fato do mundo em que vivemos possuir três dimensões espaciais. Isto significa que não existem anyons “puros”, qualquer anyon existente deve viver dentro de algum material onde as interações são efetivamente bidimensionais. Isto implica, em particular que todo sistema de anyons surge de um sistema de muitas partículas interagentes e, portanto, é bastante complicado. O efeito Hall quântico teve um enorme peso no desenvolvimento histórico do assunto e ainda é um dos principais candidatos para detecção experimental de anyons. As descrições que surgiram a partir do estudo deste efeito tratam anyons em termos de teoria de campos topológica ou teoria de campos conformes. Mencionamos brevemente como a descrição via teoria de campos topológica pode surgir na seção 2.3.2.

Como dito na introdução, não tratamos do efeito Hall quântico justamente pois não é objetivo deste trabalho investigar as formulações via teorias de campo. No entanto, as mencionamos aqui para contextualizar o uso de categorias. Além do argumento acima, existe um outro motivo para usar categorias: certas categorias, as que descrevem anyons, são aproximadamente equivalente a certos tipos de teorias de campos topológica e certos tipos de teorias de campos conformes. Esta equivalência é esperada, já que todas as descrições tratam do mesmo fenômeno, no entanto também é de interesse para a física-matemática. Ela está ilustrada no diagrama a seguir [39, 40].



No diagrama, CTM significa categorias tensoriais modulares, FM, funtores modulares, TQCT, teoria quântica de campos topológica e TCC, teoria de campos conformes. Exceto por categorias tensoriais modulares, os tipos de categorias que descrevem anyons, não exploraremos nenhum dos outros conceitos além da, extremamente breve, apresentação feita acima. As relações entre estes objetos é discutida por Bakalov e Kirillov em seu livro, Lições sobre Categorias Tensoriais e Funtores Modulares [39] em tradução livre, que reúne os principais resultados sobre o assunto.

Como dito anteriormente, a vantagem de se usar categorias para descrever anyons é que foco é dado às regras de fusão e à estatística e apenas estas propriedades descrevem completamente o modelo. Veremos como isto se dá ao longo deste capítulo e porque esta descrição é interessante, em particular para computação quântica com ânyons. Começaremos então, como uma exposição sobre os conceitos básicos de categorias, definindo precisamente o que são categorias, funtores e o conceito de naturalidade e outros conceitos correlatos.

3.1 - Categorias

Assim como o conceito de espaços topológicos surgiu para sustentar o estudo de funções contínuas, diz-se que o conceito de categoria surgiu para sustentar o estudo de funtores e estes, por sua vez, surgiram para sustentar o conceito de transformações naturais. Naturalmente, agora que estes conceitos estão definidos, são apresentados na ordem inversa. É isto que faremos aqui, começando por definir categorias.

Definição 3.1: Uma **categoria**, \mathbf{C} , consiste em

- Uma coleção de **objetos**, $\text{Ob}(\mathbf{C})$;
- Para objetos $A, B \in \text{Ob}(\mathbf{C})$, uma coleção $\mathbf{C}[A, B]$ de **flechas** de A a B ;
- Para cada objeto $A \in \text{Ob}(\mathbf{C})$, uma flecha $\text{id}_A \in \mathbf{C}[A, A]$, chamada **identidade** em A ;
- Para objetos $A, B, C \in \text{Ob}(\mathbf{C})$, uma operação

$$\begin{aligned} \circ : \mathbf{C}[B, C] \times \mathbf{C}[A, B] &\rightarrow \mathbf{C}[A, C] \\ (g, f) &\mapsto g \circ f, \end{aligned}$$

chamado **composição**,

que satisfazem aos seguintes axiomas:

- **Associatividade:** para cada $f \in \mathbf{C}[A, B]$, $g \in \mathbf{C}[B, C]$ e $h \in \mathbf{C}[C, D]$ tem-se

$$(h \circ g) \circ f = h \circ (g \circ f);$$

- **Identidade:** para cada $f \in \mathbf{C}[A, B]$, tem-se

$$f \circ \text{id}_A = f = \text{id}_B \circ f.$$

Primeiro, notamos que as flechas de uma categoria também são referidas como **setas** ou **morfismos**. Evitaremos usar terminologia associada com morfismos aqui para distanciar o conceito de flechas em categorias daquele de funções e morfismos usuais. Os motivos para tal ficarão mais evidentes conforme avançamos. Além disso, frequentemente escreve-se $A \in \mathbf{C}$ para indicar que $A \in \text{Ob}(\mathbf{C})$, $f : A \rightarrow B$ ou $A \xrightarrow{f} B$ para indicar $f \in \mathbf{C}[A, B]$ e gf para indicar $g \circ f$. Se não houver ambiguidade quanto à categoria em questão, escreveremos apenas $[A, B]$ para indicar o conjunto de flechas de A a B . Além disto, a definição de categoria é tal que dada uma sequência

$$A_0 \xrightarrow{f_0} A_1 \xrightarrow{f_1} \dots \xrightarrow{f_n} A_n$$

de flechas de \mathbf{C} é possível construir exatamente um mapa

$$A_0 \longrightarrow A_n.$$

Tal mapa é dado por $f_1 f_2 \dots f_n$. Existem, evidentemente, diversos modos de inserir parênteses entre as operações de composição destes n mapas, no entanto a associatividade implica que todos são iguais. Para cada flecha $f \in \mathbf{C}[A, B]$, chamamos A de **fonte** ou **domínio** de f e B , **alvo** ou **contradomínio** de f . Dados objetos e flechas de uma categoria, frequentemente fala-se de **diagramas comutativos**, por exemplo

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ h \downarrow & & \downarrow g \\ C & \xrightarrow{i} D \xrightarrow{j} & E \end{array}$$

Diz-se que este diagrama **comuta** se $gf = jih$. De modo mais geral, um diagrama comuta se quaisquer dois caminhos entre um objeto X e um objeto Y , as flechas de X a Y obtidas por composição das flechas ao longo de um caminho é igual à flecha obtida pela composição daquelas do outro caminho.

Por último, chamamos atenção para a palavra **coleção** usada na definição. Os objetos

de uma categoria podem ou não formar um conjunto, assim como as flechas entre objetos de modo que a palavra “coleção” aqui não necessariamente possui o mesmo significado de conjunto. Basicamente isto se dá porque existem coleções que são “grandes demais” para serem conjuntos. Isto se relaciona, por exemplo, com o paradoxo de Russel e com o fato de que o conjunto de todos os conjuntos não é um conjunto e, de modo mais geral, com os axiomas da teoria de conjuntos.

O paradoxo de Russel afirma que o conjunto de todos os conjuntos que não pertencem a si mesmos não é um conjunto. Primeiro é importante notar que conjuntos de conjuntos são perfeitamente válidos. A topologia de um espaço, por exemplo, é um exemplo. Segundo, é simples de pensar em conjuntos que não são elementos de si mesmos, por exemplo \mathbb{R} . O conjunto dos números reais não é um número real. O paradoxo é uma crítica à chamada teoria de conjuntos ingênua — *naive set theory* — que existia antes que teoria de conjuntos fosse formalmente desenvolvida. Seu propósito é mostrar que só porque é possível delimitar um grupo de coisas com base em uma dada propriedade destas coisas, isto não torna este agrupamento de coisas um conjunto.

Similarmente, o resultado de que o conjunto de todos os conjuntos não é um conjunto decorre do paradoxo de Cantor. Este resultado essencialmente trata da questão de que a cardinalidade de um conjunto é estritamente menor que a cardinalidade do conjunto de suas partes. A cardinalidade de um conjunto é essencialmente uma medida do número de elementos que possui. Dois conjuntos possuem a mesma cardinalidade se existe uma bijeção entre eles.

Quando a teoria de conjuntos foi formalizada com um conjunto de axiomas, estes paradoxos foram eliminados. Geralmente fala-se então de **classes**. A definição de classe depende do conjunto de axiomas adotados para se “definir” conjuntos. O conjunto de axiomas mais comum é o ZFC, Zermelo-Fraenkel, com o axioma da escolha. Com estes axiomas, classe é um conceito informal que corresponde a qualquer agrupamento de coisas segundo uma propriedade em comum. Assim, a coleção de todos os conjuntos que não são elementos de si mesmos, e a coleção de todos os conjuntos formam classes. Note que todo conjunto é uma classe.

Mais detalhes sobre o papel de classes na teoria de categorias podem ser encontradas nas referências [36, 38]. A distinção entre conjuntos e classes, no entanto, fornece a seguinte definição sobre categorias retirada de [36]:

Definição 3.2: Uma categoria, \mathbf{C} é dita **pequena** quando a classe ou coleção de todas as flechas de \mathbf{C} é um conjunto. Ela é dita **grande** caso contrário. Uma categoria é dita **localmente pequena** se para todo par de objetos $A, B \in \mathbf{C}$, $\mathbf{C}[A, B]$ é um conjunto.

Notamos, por fim que esta discussão é muito mais profunda pois acontece que a lin-

guagem categorial pode ser usada para fornecer outra axiomatização de conjuntos, como descrito no capítulo três de [36] e no artigo Repensando Teoria de Conjuntos, de Leinster [41]. Ele argumenta que esta axiomatização captura as propriedades que se exigem de conjuntos de modo muito mais intuitivo, além de estar mais alinhada com o uso cotidiano de conjuntos, por matemáticos. Chamamos atenção para o fato de que linguagem categorial não quer dizer que categorias são definidas anteriormente a esta axiomatização, apesar de isto poder ser feito [38]. Quer dizer apenas que usamos ideias oriundas de teoria de categorias para formular estes axiomas, de modo similar a como a definição de continuidade em espaços topológicos é fundamentada na definição anterior de continuidade para funções reais de modo a abrangê-la. Findado este último comentário, não discutiremos mais sobre estas questões de teoria de conjuntos, axiomatizações e usaremos simplesmente o termo coleções e a definição 3.2 quando for relevante. Vejamos agora alguns exemplos de categorias.

Mon

A definição de monoide já foi apresentada anteriormente, mas a repetimos aqui por conveniência. Um **monoide** é uma estrutura

$$(M, \star, e)$$

onde M é um conjunto, \star é uma operação binária em M e e é um elemento selecionado de M , tais que:

$$(r \star s) \star t = r \star (s \star t) \quad \text{e} \quad 1 \star r = r = r \star 1$$

para todo $r, s, t \in M$. Deste modo, a operação é associativa e o elemento nomeado é a unidade desta operação. Usualmente omite-se o símbolo da operação e escreve-se, simplesmente, rs para $r \star s$.

Dados dois monoides, (M, \star, e_M) e $(N, *, e_N)$ um mapa

$$\phi : M \rightarrow N$$

é um **morfismo de monoides** se respeita a estrutura de cada monoide, isto é, se é um mapa tal que:

$$\phi(r \star s) = \phi(r) * \phi(s) \quad \text{e} \quad \phi(e_M) = e_N$$

para todo $r, s \in R$. Dados dois morfismos de monoides $\phi : M \rightarrow N$ e $\psi : N \rightarrow S$ entre monoides (M, \star, e_M) , $(N, *, e_N)$ e (L, \bullet, e_L) , a composta $\psi \circ \phi : M \rightarrow S$ é, também um

morfismo de monoides. De fato, tem-se

$$\begin{aligned} (\psi \circ \phi)(r \star s) &= \psi(\phi(r \star s)) = \psi(\phi(r) * \phi(s)) \\ &= \psi(\phi(r)) \bullet \psi(\phi(s)) = (\psi \circ \phi)(r) \bullet (\psi \circ \phi)(s) \end{aligned}$$

e

$$(\psi \circ \phi)(e_M) = \psi(\phi(e_M)) = \psi(e_N) = e_L.$$

Deste modo, se considerarmos a coleção de todos os monoides e a coleção de todos os morfismos de monoides, obtemos a categoria **Mon**. Cada morfismo de monoide $\phi : M \rightarrow N$ dá origem a uma flecha entre os objetos M e N :

$$M \xrightarrow{\phi} N.$$

A composição de morfismos de monoides dá origem à composição de flechas, que representamos no diagrama abaixo.

$$\begin{array}{ccc} & N & \\ \phi \nearrow & & \searrow \psi \\ M & \xrightarrow{\psi \circ \phi} & L. \end{array}$$

A associatividade da composição equivale à comutatividade do seguinte diagrama:

$$\begin{array}{ccccc} M & \xrightarrow{\phi} & N & & \\ & \searrow \psi \circ \phi & \downarrow \psi & \searrow \chi \circ \psi & \\ & & L & \xrightarrow{\chi} & O. \end{array}$$

A verificação que é imediata, pois pela definição de morfismos de monoide, para todo $s \in M$, tem-se:

$$((\chi \circ \psi) \circ \phi)(s) = (\chi \circ \psi \circ \phi)(s) = (\chi \circ (\psi \circ \phi))(s).$$

Por fim, para cada monoide M , a identidade em M dá origem à flecha

$$M \xrightarrow{\text{id}_M} M$$

tal que para toda flecha $\phi : M \rightarrow N$, tem-se:

$$\phi \circ \text{id}_M = \phi = \text{id}_N \circ \phi.$$

Deste modo Mon satisfaz à todos os axiomas de uma categoria.

Pos

Um **conjunto parcialmente ordenado**, ou *poset*, é uma estrutura

$$(A, \preceq)$$

onde A é um conjunto e \preceq é uma **ordem parcial**, isto é, uma relação binária que satisfaz às seguintes condições:

- **Reflexividade:** $a \preceq a$;
- **Transitividade:** Se $a \preceq b$ e $b \preceq c$, então $a \preceq c$;
- **Antissimetricidade:** Se $a \preceq b$ e $b \preceq a$, então $a = b$.

A ordem parcial \preceq lê-se “precede ou é igual a”, assim como qualquer ordem parcial que surgir a partir de agora e que não seja conhecida. Os números reais, \mathbb{R} , com o ordenamento usual, \leq , formam um *poset*. Tal *poset* é também linearmente ordenado pois todos os elementos são comparáveis, isto é, para todo $x, y \in \mathbb{R}$, ou $x \leq y$ ou $y \leq x$. O conjunto das partes de um conjunto A , $\mathcal{P}(A)$ com a relação \subseteq também forma um conjunto parcialmente ordenado.

Dados dois *posets* (A, \preceq) e (B, \leq) , um **morfismo de posets** é um mapa $m : A \rightarrow B$ tal que

$$a \preceq a' \implies m(a) \leq m(a').$$

Este tipo de mapa é chamado de mapa monotônico. Funções $f : \mathbb{R} \rightarrow \mathbb{R}$ que são monotonicamente crescentes são morfismos de *posets* de \mathbb{R} em \mathbb{R} , por exemplo. Agora note que para todo *poset*, (A, \preceq) a função identidade, id_A é um mapa monotônico de $A \rightarrow A$. Além disto, dados três conjuntos parcialmente ordenados (A, \preceq) , (B, \leq) e (C, \dashv) e mapas monotônicos $f : A \rightarrow B$ e $g : B \rightarrow C$, tem-se que a composta $g \circ f : A \rightarrow C$ também é monotônica.

$$a \preceq a' \implies f(a) \leq f(a') \implies g(f(a)) \dashv g(f(a'))$$

logo

$$a \preceq a' \implies (g \circ f)(a) \dashv (g \circ f)(a').$$

A verificação de que a composição de mapas monotônicos é associativa e que a função identidade é a unidade para a operação de composição são diretas e análogas ao exemplo

anterior.

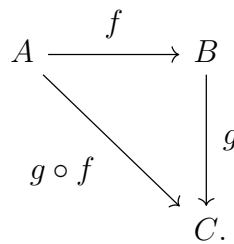
Deste modo se considerarmos a coleção de todos os conjuntos parcialmente ordenados e a coleção de todas as funções monotônicas entre estes conjuntos, obtemos uma categoria, chamada **Pos**. Os objetos de **Pos** são os *posets* e as flechas, as funções monotônicas. Em vista dos resultados acima, **Pos** de fato forma uma categoria.

Set

A categoria **Set** é a categoria cujos objetos são conjuntos – todos os conjuntos – e as flechas são as funções entre estes conjuntos. A composição de funções fornece a composição de flechas, isto é, dadas duas funções $f : A \rightarrow B$ e $g : B \rightarrow C$, a função composta é dada por $g \circ f : A \rightarrow C$, definida por:

$$(g \circ f)(a) = g(f(a))$$

para todo $a \in A$. A operação de composição está representada no diagrama a seguir.



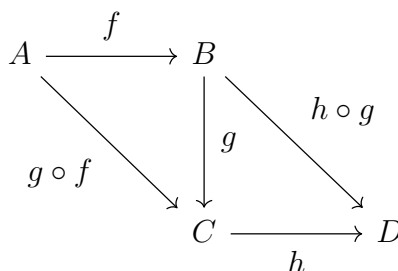
Como se sabe, a composição usual de funções, \circ , é associativa. Assim, considerando mais uma função $h : C \rightarrow D$:

$$(h \circ g) \circ f = h \circ (g \circ f)$$

já que para qualquer $a \in A$, tem-se:

$$((h \circ g) \circ f)(a) = h(g(f(a))) = (h \circ (g \circ f))(a)$$

segundo a definição de função composta. A associatividade é representada pelo seguinte diagrama de flechas:



Por fim, como todo conjunto possui uma função identidade:

$$id_A : A \rightarrow A,$$

dada por

$$id_A(a) = a,$$

cada objeto de **Set** possui uma flecha identidade tal que

$$f \circ id_A = f = id_B \circ f.$$

Portanto, funções usuais entre conjuntos satisfazem os axiomas e **Set** de fato é uma categoria. Note que é necessário apenas a operação de composição e unidades para entender funções como flechas de uma categoria e o domínio e codomínio destas funções como objetos desta categoria. Enfatizamos que não é necessário nenhum conhecimento adicional sobre quais conjuntos são esses ou informações extras sobre as funções.

Além da categoria **Set**, existe a categoria **Set_{fin}** cujos objetos são todos os conjuntos finitos e as flechas são as funções entre estes conjuntos. Inclusive existem muitas categorias obtidas a partir de restrições dos conjuntos que serão os objetos e das funções que serão as flechas. Por exemplo, tomando os conjuntos finitos como objetos e funções injetivas como flechas, obtemos uma nova categoria, já que a composição de injeções é também uma injeção e já que a identidade é uma injeção. Apesar da simplicidade aparente, **Set** é um categoria bastante importante, não estudaremos suas propriedades mais interessantes mas voltaremos a nos encontrar com ela adiante.

Categorias de conjuntos estruturados

Todos os exemplos anteriores apresentaram um tipo particular de categorias, as vezes referidas como **categorias concretas**. Informalmente, categorias concretas são aquelas cujos objetos são conjuntos, possivelmente munidos de estruturas extras e cujas flechas são funções entre estes conjuntos que preservam estas estruturas. Mais exemplos de categorias deste tipo incluem:

- A categoria **Grp** de grupos e homomorfismos de grupos;
- A categoria **Vect_K** de espaços vetoriais sobre um corpo K e mapas lineares;
- A categoria **Grph** de grafos e homomorfismos de grafos;
- A categoria **Rng** de todos os anéis e homomorfismos de anéis;
- A categoria **Top** de espaços topológicos e mapas contínuos.

Discutimos anteriormente que a coleção de objetos e de flechas de uma categoria não necessariamente formam conjuntos. Este é um ponto de divergência, por assim dizer, de categorias em relação a outros conceitos como grupos, espaços topológicos e afins. O outro ponto de divergência é que os objetos de uma categoria não necessariamente se assemelham a objetos de um conjunto usual e, mais importante, as flechas de uma categoria podem ser bem distintas de funções. Vejamos agora exemplos de categorias deste tipo.

Rel

Considere a coleção de todos os conjuntos e a coleção de todas as relações nestes conjuntos. Definimos flechas

$$A \xrightarrow{R} B$$

entre conjuntos como uma relação entre eles, isto é, como um subconjunto, $R \subseteq B \times A$. Cada conjunto, A , possui uma flecha identidade, id_A obtida a partir da relação identidade. Isto é:

$$\text{id}_A = \{ (a, a) \in A \times A \mid a \in A \} \subseteq A \times A.$$

Além disso, dadas relações $R \subseteq B \times A$ e $S \subseteq C \times B$ define-se a relação composta $S \circ R \subseteq C \times A$ por:

$$(c, a) \in S \circ R \iff \exists b. (c, b) \in S \wedge (b, a) \in R.$$

Esta operação de composição é associativa pois, dadas três relações, $S \subseteq B \times A$, $R \subseteq C \times B$ e $T \subseteq D \times C$, tem-se:

$$\begin{aligned} (d, a) \in T \circ (R \circ S) &\iff \exists c. (d, c) \in T \wedge (c, a) \in R \circ S \\ &\iff \exists c. (d, c) \in T \wedge [\exists b. (c, b) \in R \wedge (b, a) \in S] \\ &\iff \exists b. (b, a) \in S \wedge [\exists c. (d, c) \in T \wedge (c, b) \in R] \\ &\iff \exists b. (b, a) \in S \wedge (d, b) \in T \circ R \\ &\iff (d, a) \in (T \circ R) \circ S. \end{aligned}$$

Além disso, pela definição de composição, tem-se $R \circ \text{id}_A = R = \text{id}_B \circ R$, pois

$$(b, a) \in R \circ \text{id}_A \iff \exists a. (b, a) \in R \wedge (a, a) \in \text{id}_A.$$

Como para todo $a \in A$, $(a, a) \in \text{id}_A$, tem-se que a condição acima é simplesmente a definição da relação R . Um argumento análogo mostra que $\text{id}_B \circ R = R$ de modo que a igualdade $R \circ \text{id}_A = R = \text{id}_B \circ R$ é garantida.

Deste modo, fica estabelecido que a coleção de todos os conjuntos e a coleção de todas as relações nestes conjuntos de fato forma uma categoria. Tal categoria é chamada **Rel** e é um exemplo de categoria onde as flechas não são funções no sentido usual. Notamos, inclusive, que **Rel** é uma “expansão” da categoria **Set** pois cada função dá origem a uma relação, seu gráfico. No entanto, nem toda relação pode ser associada a uma função.

Grupos e monoídes

O exemplo anterior ilustra que as flechas de uma categoria não necessariamente são funções. Neste, trataremos de categorias cujos objetos não são conjuntos. Considere a categoria com um único objeto, que denotaremos por $*$. Equipada com uma única flecha identidade

$$* \xrightarrow{\text{id}_*} *$$

esta construção forma uma categoria. Note que não importa qual a natureza do objeto e nem da flecha identidade. Como não existe nenhuma outra flecha, todos os outros axiomas são automaticamente satisfeitos. Esta categoria é frequentemente referida como **1**. Existem, no entanto, outras categorias com apenas um objeto.

Um monoíde forma uma categoria com apenas um objeto. A natureza do objeto não importa e a estrutura do monoíde é carregada pelas flechas da categoria. Assim, se (M, \star, e) é um monoíde, M pode ser visto como uma categoria, \mathbf{M} , da seguinte forma: para cada elemento $r \in M$, associe uma flecha

$$* \xrightarrow{r} *$$

em \mathbf{M} . A composição de flechas r e s é a flecha associada ao produto $s \star r$ em M :

$$\begin{array}{ccc}
 & * & \\
 r \nearrow & & \searrow s \\
 * & \xrightarrow{s \circ r = s \star r} & *
 \end{array}$$

Como \star é associativa e e a unidade em M gera uma flecha identidade para \mathbf{M} , esta construção de fato gera uma categoria. Tal categoria evidencia a característica fundamental de um monoíde. Como o objeto $*$ não tem nenhuma estrutura, deste ponto de vista, toda a informação sobre um monoíde é carregada pelas flechas. Essencialmente, então, um monoíde trata de mapas de um objeto em si mesmo que não são necessariamente reversíveis. Se, para cada flecha r de \mathbf{M} , atribuirmos uma flecha inversa r^{-1} tal que $r \star r^{-1} = e$, então obtemos um grupo, que trata de todos os modos de mapear um objeto

em si mesmo de maneira reversível, isto é, das simetrias de um objeto.

Elaborando um pouco mais, considere um conjunto qualquer X . O conjunto de todas as função $f : X \rightarrow X$ forma um monoide. Cada função f mapeia X em X mas não o faz necessariamente de modo reversível. Por exemplo, se f não for injetiva, dois elementos originalmente distintos podem ser mapeados em um mesmo elemento. Esta função não inversível. Por outro lado, o conjunto de todas as bijeções de $X \rightarrow X$ formam o bem conhecido grupo de permutações. Cada bijeção leva X em X de tal modo que é possível desfazer o processo.

Por fim, note que grupos, vistos como categorias apontam para uma propriedade interessante de flechas. Um grupo pode ser descrito como uma categoria de um elemento onde todas as flechas são inversíveis, isto é, onde existe, para toda flecha g uma flecha g^{-1} tal que

$$g \circ g^{-1} = \text{id}_* = g^{-1} \circ g.$$

onde id_* é a flecha identidade da categoria. Pares de flechas para os quais estas propriedade vale são chamadas de **isomorfismos**, g e g^{-1} formam um par de isomorfismos inversos. Frequentemente, isomorfismos são morfismos entre estruturas algébricas que também são bijeções. Um homomorfismo de grupos que é bijetivo é um isomorfismo de grupos, por exemplo. Esta propriedade sugere que existam tipos especiais de flechas, em particular, deve existir um análogo para flechas dos conceitos de injetividade e sobrejetividade. Trataremos disto a partir de agora.

3.1.1 - Flechas especiais

Um dos lemas estabelecidos da teoria de categorias é que “as flechas são o que realmente importam”. No espírito deste lema, a descrição categoria visa descrever propriedades, não em termos de objetos, mas em termos de flechas. Por exemplo, usando-se esta filosofia, pode-se descrever um conjunto de um elemento como o conjunto para o qual existe exatamente uma função $f : Y \rightarrow \mathbf{1}$, para qualquer conjunto Y . Esta função é aquela que envia todos os elementos de Y no único elemento de $\mathbf{1}$. Note o caso particular de $Y = \emptyset$, para qualquer conjunto X existe uma única função $f : \emptyset \rightarrow X$, a função que não faz nada, já que \emptyset não possui nenhum elemento.

Usaremos esta abordagem para definir algumas propriedades de interesse para flechas nesta seção e para objetos na seção seguinte. A maioria das definições a seguir serão duais, isto é, definiremos dois conceitos simultaneamente, um à esquerda e outra à direita. Eventualmente trataremos sobre como estes conceitos “duais” estão relacionados. Começamos com a seguinte definição:

Definição 3.3: Em uma categoria \mathbf{C} , uma flecha

$$B \xrightarrow{m} A; \quad A \xrightarrow{e} B$$

é dita **mônica** ou **épica**, respectivamente, se, para cada par de flechas paralelas

$$X \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} B; \quad B \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} X$$

tem-se:

$$m \circ f = m \circ g \implies f = g; \quad f \circ e = g \circ e \implies f = g$$

Uma observação: flechas mônicas e épicas são mais comumente conhecidas como monomorfismos e epimorfismos respectivamente. Mônica e épica são neologismo que usamos neste trabalho para traduzir os termos *monic* e *epic* que são os adjetivos, ou as vezes substantivos ingleses usualmente utilizados para se referir a estes conceitos. Escolhemos esta terminologia em adequação com a terminologia que adotamos para tratar teoria de categorias usando o termo “flecha” ao invés de “morfismo” e por conveniência. Ao ler outros textos, tenha em mente que os conceitos aqui referidos como flechas mônicas ou épicas são usualmente chamados, respectivamente, de **monomorfismos** e **epimorfismos** ou em inglês, *mono* e *epi*, ou ainda, *monic arrow* e *epic arrow*. *Mono* e *epi* as vezes são encontrados em português.

Frequentemente, denotam-se flechas mônicas e épicas por

$$B \succ \longrightarrow A \quad \text{e} \quad A \longrightarrow \succ B$$

respectivamente. Usaremos esta notação quando for conveniente. Pela definição, então, vê-se que uma flecha é mônica se pode ser cancelada à esquerda e épica se pode ser cancelada à direita. Estas propriedades são generalizações dos conceitos de injetividade de sobrejetividade para categorias. Para ver isto, considere a seguinte situação: suponha que $m \circ f = m \circ g$ para funções $m : B \rightarrow X$ e $f, g : A \rightarrow B$. Então, para todo $a \in A$

$$m(f(a)) = m(g(a)) \implies f(a) = g(a) \iff f = g$$

se, somente se, m for injetiva. Ou seja, uma flecha é mônica em \mathbf{Set} se, somente se, for uma função injetiva. Do mesmo modo, suponha $e : X \rightarrow A$ tal que $f \circ e = g \circ e$. Tem-se:

$$f(e(x)) = g(e(x)) \implies f = g$$

se somente se a igualdade vale para todo $a = e(x) \in A$, ou seja, se, somente se, e for sobrejetiva.

Note, também, que a propriedade de ser mônica ou épica diz respeito à capacidade de ser cancelada pela esquerda ou direita, respectivamente. No caso de **Set**, esta possibilidade é equivalente à existência de uma inversa esquerda e direita, já que uma função injetiva possui, possivelmente múltiplas, inversas esquerdas e uma função sobrejetiva possui, possivelmente múltiplas, inversas direitas. O conceito de inversas também possui uma descrição categorial:

Definição 3.4: Dado uma flecha

$$A \xrightarrow{f} B,$$

e um par de flechas

$$A \xrightarrow{s} B; \quad B \xrightarrow{r} A,$$

se

$$f \circ s = \text{id}_B \quad \text{e} \quad r \circ f = \text{id}_A$$

então, s é uma **seção** de r e r é uma **retração** de s .

Note, então, que uma seção, s , de uma flecha f é uma inversa direita de f e uma retração, r , de uma flecha g , uma inversa esquerda de g . Antes de prosseguirmos, para elucidar o porquê destes nomes, consideremos estes conceitos na categoria **Set** onde as flechas são funções comuns. Tem-se:

- Uma retração, r de uma função $f : X \rightarrow Y$, isto é, sua inversa esquerda, trás de volta todos os pontos mapeados por f ao longo de suas trajetórias originais, possivelmente mapeando pontos que não pertencem à imagem de f em pontos arbitrários de X . Existe uma retração para cada escolha distinta de como mapear os pontos não pertencentes à $f(X)$ em pontos de X .
- Por outro lado, uma seção, s de uma função $g : X \rightarrow Y$ é uma inversa direita. Assim, a seção escolhe, para cada ponto y , um ponto $x \in g^{-1}(\{y\})$, que possivelmente possui mais de um elemento, e mapeia y em x . Deste modo, tem-se que $g(s(y)) = y$. Para cada escolha de x existe uma seção s de g . Podemos pensar s então como escolhendo uma seção de $g^{-1}(\{y\})$ ou uma seção do domínio de g .

A discussão acima, também, nos indica uma relação interessante entre seções e flechas mônicas e retrações e flechas épicas. Uma função possui uma inversa esquerda se é uma injeção e uma inversa direita se é uma sobrejeção. Além disto, um inversa esquerda necessariamente é um sobrejeção e uma inversa direita, uma injeção. Se o fato de uma flecha ser mônica está relacionado com injetividade e o fato de ser épica, com sobrejetividade,

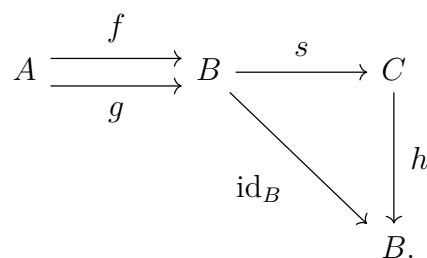
então uma seção daria origem a uma flecha mônica, e uma retração, a uma flecha épica. De fato, este é o caso.

Proposição 3.1: Toda seção é uma flecha mônica e toda retração, uma flecha épica.

Prova: Sejam f, g duas flechas de A em B e s a seção de uma flecha h de C em B . Suponha, então que tenhamos $s \circ f = s \circ g$. Como s é uma seção de h , $h \circ s = \text{id}_B$. Deste modo

$$f = \text{id}_B \circ f = h \circ (s \circ f) = h \circ (s \circ g) = \text{id}_B \circ g = g.$$

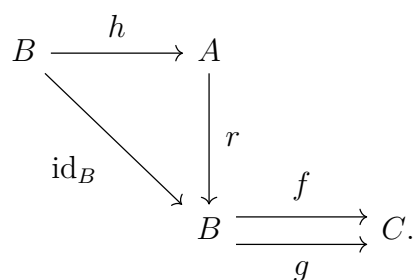
Ou seja, s é mônica. O argumento está ilustrado no diagrama abaixo.



Do mesmo modo, se agora f e g são flechas de B em C e se r for uma retração de uma flecha $h : A \rightarrow B$, então, supondo $f \circ r = g \circ r$ tem-se

$$f = f \circ \text{id}_B = (f \circ r) \circ h = (g \circ r) \circ h = g \circ \text{id}_B = g.$$

Ou seja, r é épica. O argumento está ilustrado no diagrama a seguir.



□

Em vista deste resultado, uma seção e uma retração frequentemente são chamadas de *split monic* e *split epic*, respectivamente, apesar dos termos *section* e *retraction* também existirem em inglês. A tradução de *split* para o português é usualmente feita como “cinde”, do verbo cindir. Ou seja, uma seção seria uma monomorfismo que cinde, ou, na terminologia que adotamos, uma flecha mônica que cinde. Do mesmo modo, uma retração seria uma flecha épica que cinde. Não encontramos explicações para o uso de *split* na notação, porém, ela é conveniente quando desejamos olhar seções e retrações como flechas

mônicas e épicas, por isto, nos referiremos a estes conceitos como **flechas mônicas que cindem** e **flechas épicas que cindem** quando for conveniente.

Antes de passarmos para o próximo conceito, notamos que toda seção é uma flecha mônica e toda retração é uma flecha épica mas o contrário nem sempre é verdade [38]. Por exemplo, em **Set** é fácil de ver que o contrário vale. Toda seção é uma injeção e, portanto, possui uma inversa direita, de modo que pode ser cancelada à direita. Analogamente, toda retração é uma sobrejeção e, portanto, possui uma inversa esquerda, podendo ser cancelada à esquerda. Isto não vale na categoria **Grp** de todos os grupos ou **Top** de todos os espaços topológicos, por exemplo.

O motivo para tal é essencialmente porque, nestas categorias, uma flecha ser cancelável de um dos lados não equivale a dizer que esta flecha possui uma inversa daquele lado. Como funções, esta equivalência vale, mas, as flechas em **Grp** e **Top** são, além de funções, homomorfismos de grupo ou funções contínuas. Cancelabilidade à esquerda — ser mônica — só seria equivalente à existência de uma flecha inversa à esquerda — ser uma seção — em **Top** se pudéssemos garantir que para qualquer injeção contínua pelo menos uma de suas inversas esquerdas é, também, contínua. Experiência comum com cálculo básico indica que esta é uma exigência que não pode ser cumprida.

Esta discussão deixa nítida a diferença entre os conceitos de flechas mônicas e épicas e suas contrapartes que cindem, seções e retrações, respectivamente. Ser mônica ou ser épica trata de ser cancelável à esquerda e à direita enquanto ser uma seção ou uma retração trata de ser uma inversa direita ou esquerda, respectivamente para alguma flecha. É neste ponto que encarar flechas como funções pode ser prejudicial, pois a categoria **Set** é especial sob muitos aspectos e suas propriedades em geral não são carregadas para outras. Além disso, lembre-se da categoria **Rel**, onde as flechas são relações, ou seja, subconjuntos. Tanto ser cancelável quando possuir inversa não são conceitos que se traduzem bem em **Rel** se pensarmos nas flechas como subconjuntos explicitamente. Não há motivo, então, para crer que resultados que valham em **Set**, que tratam de funções comuns, valham em **Rel**.

Apesar de funções serem um tipo particular de relações, o modo usual de se pensar estas duas coisas é diferente. Funções usualmente são pensadas como regras aplicadas aos elementos de um conjunto de modo a produzir os elementos de outro. Por outro lado, relações são comparações entre dois elementos, por exemplo, $x > y$, $A \subset B$. No primeiro caso, um elemento é obtido a partir de outro enquanto no segundo ambos os elementos são dados. Ressaltamos isto porque, como dissemos, flechas não necessariamente são funções e o motivo pelo qual escolhemos usar o termo “flechas” ao invés de “morfismos”, que é mais comum, é para distanciar o primeiro da ideia de funções.

Notado isto, passamos ao próximo conceito, bastante familiar, o de uma flecha que possui tanto uma inversa direita quando uma inversa esquerda. Usualmente tais flechas são chamadas de isomorfismos, iremos nos ater à terminologia adotada, no entanto, e nos

referir a elas como flechas inversíveis, como feito em [38]. Vejamos a definição formal:

Definição 3.5: Dado uma flecha

$$A \xrightarrow{f} B$$

Se existir uma flecha

$$B \xrightarrow{f'} A$$

tal que

$$f' \circ f = id_A \quad \text{e} \quad f \circ f' = id_B$$

então dizemos que f e f' formam um **par de flechas inversas**. Cada flecha é dita **inversível**. Se existe um par de flechas inversas entre objetos A e B , estes objetos são ditos **isomorfos**.

Proposição 3.2: Dado uma flecha f , a flecha f' , que forma com f um par de flechas inversas, quando existe, é única. Chamamos f' de **inversa** de f e escrevemos f^{-1} .

Prova: Dada uma flecha $f : A \rightarrow B$, suponha que existam duas flechas $f'_1, f'_2 : B \rightarrow A$ tais que

$$f'_1 \circ f = id_A \quad \text{e} \quad f \circ f'_1 = id_B;$$

$$f'_2 \circ f = id_A \quad \text{e} \quad f \circ f'_2 = id_B.$$

Estas condições indicam que os seguintes diagramas comutam:

$$A \xrightarrow{f} B \begin{array}{c} \xrightarrow{f'_1} \\ \xrightarrow{f'_2} \end{array} A \quad \text{e} \quad B \begin{array}{c} \xrightarrow{f'_1} \\ \xrightarrow{f'_2} \end{array} A \xrightarrow{f} B$$

A partir destes diagramas, obtemos:

$$\begin{array}{ccc} B & \xrightarrow{f'_1} & A \\ & \searrow id_B & \downarrow f \\ & & B \end{array} \begin{array}{c} \xrightarrow{f'_1} \\ \xrightarrow{f'_2} \end{array} A;$$

$$\begin{array}{ccc} B & \begin{array}{c} \xrightarrow{f'_1} \\ \xrightarrow{f'_2} \end{array} & A \xrightarrow{f} B \\ & \searrow id_A & \downarrow f'_1 \\ & & A. \end{array}$$

Estes diagramas comutam, pela definição de f , f'_1 e f'_2 . Assim escolhendo, por exemplo, o primeiro diagrama, tem-se:

$$f'_1 \circ f \circ f'_1 = f'_2 \circ \text{id}_B \implies f'_1 \circ \text{id}_B = f'_2 \circ \text{id}_B \implies f'_1 = f'_2.$$

O mesmo resultado pode ser obtido do segundo. Note, no entanto que além disto, o primeiro diagrama implica que f é mônica e o segundo, que f é épica. Assim, não só a inversa de uma flecha, quando existe é única, uma flecha que é inversível é simultaneamente mônica e épica. Isto é esperado já que a própria definição implica que toda flecha inversível é simultaneamente uma seção e uma retração.

□

Uma categoria onde todas as flechas são inversíveis é chamada de um **grupoide**. Um exemplo de grupoide é o grupoide fundamental de um espaço topológico, X , definido no apêndice B. Ele consiste de todas as classes de equivalência de caminhos em X sob a relação de homotopia de caminhos. Os objetos do grupoide fundamental de X , $\pi(X)$ são pontos de X e as flechas são classes de equivalência de caminhos entre dois pontos x e y . A composição de flechas é o produto de caminhos, $*$, como definido no apêndice. Todos os caminhos possuem um caminho inverso, mas só se pode formar o produto de caminhos cujos pontos final e inicial coincidem. No apêndice definimos grupoides em termos algébricos, como grupos onde a operação binária é definida apenas para um subconjunto de pontos. No entanto grupoides parecem ser de maior interesse sob a perspectiva categorial.

Prosseguindo, é evidente que nem toda seção e nem toda retração é uma flecha inversível. Mesmo em **Set**, uma flecha só é inversível, correspondendo a ser uma função bijetiva, se é simultaneamente injetiva e sobrejetiva. Uma seção, como vimos, necessariamente é injetiva, pois é a inversa esquerda de alguma outra função. Não há nenhuma outra condição que imponha que ela seja, também, sobrejetiva. A situação para retrações é análoga. Isto significa que a coleção de todas as flechas inversíveis está contida naquelas de seções e retrações, que por sua vez estão contidas nas coleções de flechas mônicas e épicas respectivamente. As relações estão representadas no diagrama 3.2 adiante.

Como notado na legenda, o gancho nas flechas indica que são mapas de inclusão, ou seja, incluem a coleção de seções, por exemplo, na coleção de flechas mônicas. São análogas ao mapa de inclusão para conjuntos. Agora, lembre-se que nem toda flecha mônica é uma seção e nem toda flecha épica é uma retração. Isto sugere perguntar se existem flechas que são simultaneamente mônicas e épicas, mas que não são simultaneamente seções ou retrações e, conseqüentemente, não são flechas inversíveis. A resposta para esta pergunta é sim. Como notamos, os conceitos de “monicidade” e “epicidade” tratam da propriedade de uma flecha ser cancelável. Esta propriedade, por sua vez pode não ser equivalente à propriedade de possuir inversas. Então, nestes termos, é perfeitamente possível que

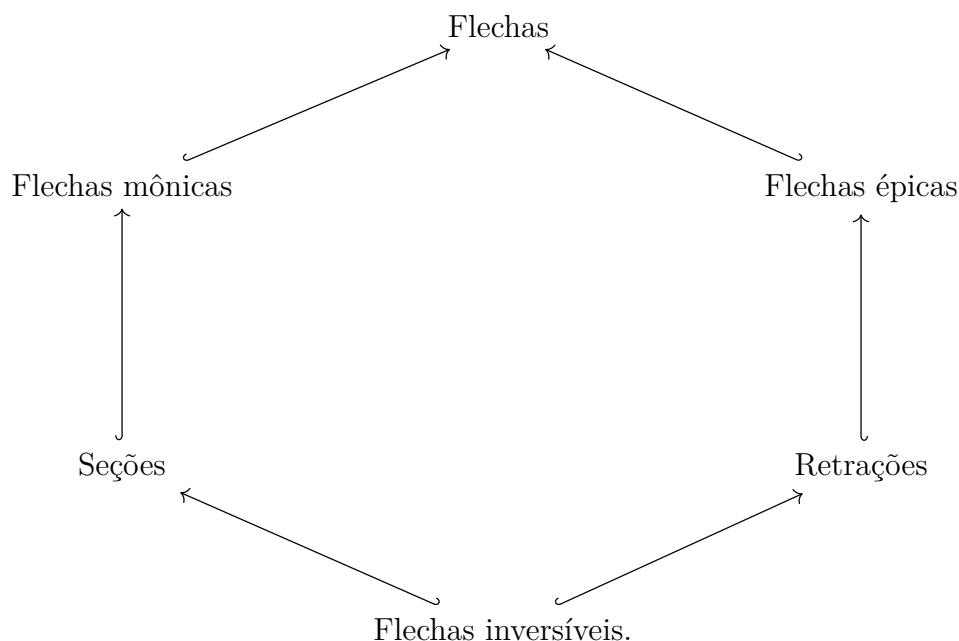


Figura 3.2: Relação entre tipos de flechas em uma categoria. A flecha com o gancho indica inclusão.

existam flechas que sejam canceláveis de ambos os lados mas que não possuam um ou outra, ou mesmo nenhuma, inversa.

Definição 3.6: Uma flecha que é simultaneamente épica e mônica é chamada de **bimorfismo**.

Notamos que no caso de bimorfismos não há escapatória do “-morfismo”. Alguns sugerem que este conceito não é importante o suficiente para ter uma palavra própria e conseqüentemente deveria ser referidos como *monic epics*, que traduziríamos como “flechas mônicas-épicas”. Uma outra terminologia sugerida é *bimonomorphism* ou *bimonic arrow*, que traduziríamos como “flecha bimônica”. Esta última é inspirada pela princípio da dualidade que veremos mais adiante. Como não encontraremos este conceito novamente além desta seção, manteremos o termo “bimorfismo”. Este conceito, então, nos permite completar o diagrama anterior, ilustrando uma espécie de hierarquia entre os conceitos básicos que tratam de flechas em uma determinada categoria. Notamos que, o fato de nem todo biformismo ser uma flecha inversível, nos leva à seguinte definição:

Definição 3.7: Uma categoria, C é dita **balanceada** quando todo bimorfismo de C é também inversível em C .

A propriedade de uma categoria ser balanceada, fornece informações úteis sobre a categoria mas, assim como bimorfismos não são de muita importância, o fato de uma categoria ser balanceada também não recebe tanta atenção. Assim como bimorfismos,

esta é a única vez que falaremos de categorias balanceadas neste trabalho. Notamos, por exemplo que **Set** é balanceada. Finalizamos então esta seção com o esquema completo mostrando as relações entre flechas de uma dada categoria.

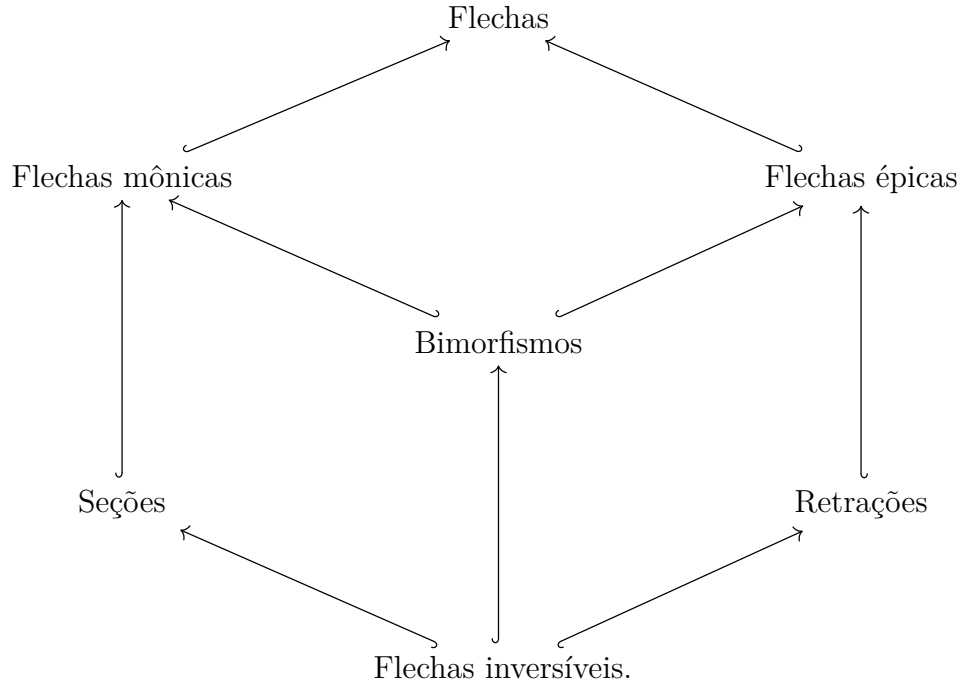


Figura 3.3: Relação completa entre os tipos de flechas em uma categoria.

3.1.2 - Objetos especiais

Assim como existem flechas distintas em uma categoria, também existem objetos distintos. Em particular, tais objetos são distintos por ocuparem posições “extremas”. Assim como a definição de flechas mônicas e épicas, a definição a seguir é “dual”.

Definição 3.8: Considere dois objetos I e T de uma categoria C . Se, para todo objeto A de C existem flechas únicas

$$I \longrightarrow A; \quad A \longrightarrow T.$$

então o objeto I é dito **inicial** e o objeto T é dito **final** ou **terminal**.

Objetos iniciais e finais são familiares de matemática elementar mas geralmente não é lhes dado esta ênfase. O motivo disto é justamente a distinção na abordagem de teoria de categorias em estudar objetos através de como estes se relacionam com outros. Para ilustrar objetos finais que de fato são conhecidos, vejamos um exemplo em **Set**.

Considere o conjunto $\mathbf{1} = \{*\}$, o conjunto de um elemento e o conjunto vazio, \emptyset .

Afirmamos que $\mathbf{1}$ é terminal em \mathbf{Set} e \emptyset é inicial em \mathbf{Set} .

$$A \longrightarrow \mathbf{1}.$$

Esta flecha é dada pela única função $f : A \rightarrow \mathbf{1}$ que mapeia todos os elementos de A no único elemento de $\mathbf{1}$. Assim, $\mathbf{1}$ é terminal em \mathbf{Set} . Por outro lado, para todo conjunto A existe uma única flecha,

$$\emptyset \longrightarrow A,$$

dada pela função $f : \emptyset \rightarrow A$ que não faz nada. A existência desta função pode parecer problemática da primeira vez que encontrada, um argumento para sua existência é o seguinte: uma função é definida de tal forma que para cada elemento, x , de seu domínio existe um elemento, y , no contradomínio tal que $f(x) = y$. Logicamente, afirmar que “para todo elemento no domínio, existe um elemento no contradomínio que é sua imagem sob f ” é verdade por vacuidade se o domínio é o conjunto vazio, pois ele não possui nenhum elemento. Mais formalmente, uma função, $f : A \rightarrow B$ é uma relação $f \subseteq A \times B$ que satisfaz à esta propriedade. Como $\emptyset \times B = \emptyset$ e o conjunto vazio é único, para cada conjunto A existe uma única função $f : \emptyset \rightarrow A$ determinada pela relação $\emptyset \subset \emptyset \times A$. Deste modo \emptyset é inicial em \mathbf{Set} .

Este exemplo também ilustra um fato interessante sobre objetos iniciais e finais. Note que o conjunto vazio é único, no entanto o conjunto de um elemento não o é. Existem infinitos tais conjuntos, no entanto todos eles são isomorfos entre si. Quaisquer dois conjuntos de um único elemento diferem apenas em como chamamos seu elemento. Enunciamos este resultado em geral para categorias agora.

Proposição 3.3: Quaisquer dois objetos iniciais ou finais em uma categoria \mathbf{C} são isomorfos.

Proposição 3.4: Considere dois objetos iniciais I_1 e I_2 , por definição, então, tem-se que para todo objeto A da categoria existem flechas

$$I_1 \longrightarrow A \quad \text{e} \quad I_2 \longrightarrow A.$$

Em particular, então, existem flechas únicas

$$I_1 \xrightarrow{f} I_2 \quad \text{e} \quad I_2 \xrightarrow{g} I_1.$$

Note, portanto que as compostas $g \circ f$ e $f \circ g$ são flechas de I_1 para I_1 e I_2 para I_2 , respectivamente. Mas, como I_1 e I_2 são iniciais, estas flechas devem ser únicas. Por outro lado, pelos axiomas de uma categoria, a cada elemento está associado uma flecha identidade, em particular, a I_1 está associada $\text{id}_{I_1} : I_1 \rightarrow I_1$ e a I_2 ,

$\text{id}_{I_2} : I_2 \rightarrow I_2$. Novamente, estas flechas devem ser únicas. Deste modo, concluímos que $g \circ f = \text{id}_{I_1}$ e $f \circ g = \text{id}_{I_2}$. Logo f e g são um par de flechas inversas e I_1 e I_2 são isomorfos. O argumento para mostrar que objetos terminais também são isomorfos é análogo, basta inverter a direção das flechas acima.

Objetos iniciais e finais são definidos em termos das flechas que saem e chegam, respectivamente a eles, mas isto não significa que não existem flechas que cujo alvo é um objeto inicial e cuja fonte é um objeto terminal. Estas flechas, em geral, representam conceitos importantes. Por exemplo, a ideia de uma flecha saindo de $\mathbf{1}$ a um conjunto A em **Set** pode ser usada para obter a noção de elemento de A [41] ao usar linguagem categórica para se definir conjuntos. Isto porque uma função de $\mathbf{1}$ em qualquer conjunto, A , corresponde simplesmente a uma escolha de um elemento de A , assim, denotados o elemento $a \in A$ pela flecha $a : \mathbf{1} \rightarrow A$.

De modo geral flechas cuja fonte são objetos terminais são chamadas de **elementos globais** e são uma espécie de generalização do conceito de elemento de um conjunto. Podemos ver isto se considerarmos duas flechas $f, g : A \rightarrow B$. Se existe um elemento global, $a : T \rightarrow A$, então tem-se

$$T \xrightarrow{a} A \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} B.$$

Como T é terminal, obtém-se que $f \circ a = g \circ a$, ou seja, quaisquer duas flechas f, g coincidem no elemento global, a . Neste sentido, então, elementos globais são uma espécie de “elemento” especial de um objeto A . Note que em **Set**, duas flechas que coincidem em todos os elementos globais de um conjunto A são iguais, já que isto equivale a dizer que $f(a) = g(a)$ para todo $a \in A$. Isto não necessariamente vale em categorias mais complicadas. O caso de flechas cujo alvo é um objeto inicial é mais complicado e não o trataremos aqui.

Por fim, assim como uma flecha pode ser simultaneamente uma seção e uma retração, ou mônica e épica, também podemos ter objetos que são simultaneamente iniciais e finais.

Definição 3.9: Um objeto que é simultaneamente inicial e final é um chamado **objeto nulo** ou **objeto zero**.

Objetos nulos de uma categoria são usualmente representados por $\mathbf{0}$ e possuem a seguinte propriedade: para quaisquer elementos A, B existe uma única flecha

$$A \longrightarrow \mathbf{0} \longrightarrow B$$

isto é, a flecha de A a B obtida pela composição através de $\mathbf{0}$. Tal flecha é chamada **flecha**

zero. Voltaremos a encontrar este tipo de flecha no futuro. Finalizamos esta seção com um exemplo de objeto zero.

Considere a categoria \mathbf{Grp} cujos objetos são grupos e as flechas, homomorfismos de grupo. O grupo trivial $\mathbf{0}$ é um objeto zero em \mathbf{Grp} . Primeiro tem-se que para qualquer grupo, (G, \star, e) , o mapa $m : \mathbf{0} \rightarrow G$ dado por $m(0) = e$, onde 0 é o único elemento de $\mathbf{0}$, é um homomorfismo. De fato, denotado a operação de grupo de $\mathbf{0}$ aditivamente, tem-se

$$m(0 + 0) = m(0) = e = e \star e = m(0) \star m(0).$$

Este mapa é o único homomorfismo de $\mathbf{0}$ em G , pois, se escolhermos qualquer outro elemento $g \in G$ e definirmos $m(0) = g$, como $g = h \star i$ para algum $h, i \in G$, tem-se:

$$m(0 + 0) = m(0) = g = h \star i \neq m(0) \star m(0).$$

Mesmo se h ou i sejam iguais a g , os axiomas de grupo garantem que o outro elemento será e e e não pertence à imagem de m . Logo a única escolha que torna m um homomorfismo é $m(0) = e$.

Por outro lado, para qualquer grupo (G, \star, e) , o mapa $n : G \rightarrow \mathbf{0}$ dado por $n(g) = 0$ para todo $g \in G$ é a única função possível de G em $\mathbf{0}$. Além disso, para quaisquer $g, h \in G$

$$n(g \star h) = 0 = 0 + 0 = n(g) + n(h).$$

Logo n é o único homomorfismo de G em $\mathbf{0}$. Assim, fica estabelecido que $\mathbf{0}$ é um objeto zero de \mathbf{Grp} . Note que assim como objetos iniciais e finais, objetos zero são definidos a menos de um isomorfismo.

3.1.3 - Princípio de dualidade

Nas seções anteriores encontramos vários conceitos que são duais em certo sentido: flechas mônicas e épicas, seções e retrações e objetos iniciais e finais. Esta dualidade é originada no **princípio da dualidade** que existe em teoria de categorias. Este princípio é extremamente útil e essencialmente faz com que conceitos em teoria de categorias sejam definidos aos pares. Informalmente, para qualquer conceito em teoria de categorias, pode-se obter seu dual invertendo-se todas as flechas envolvidas na definição. Por exemplo, se um objeto inicial, I , é definido como um objeto tal que, para qualquer objeto A , exista uma única flecha de $I \rightarrow A$. Seu dual é aquele objeto T , definido como o objeto para o qual, dado um objeto A , existe uma única flecha $A \rightarrow T$. É neste sentido que os conceitos de objeto inicial e final são duais. Formalizaremos este princípio agora.

Em termos lógicos, teoria de categorias consiste em um certo conjunto de afirmações Σ que envolvem objetos e flechas. Estas afirmações são formadas a partir de afirmações

atômicas correspondendo aos axiomas de teoria de categorias, isto é sentenças como “ A é o alvo de f ”, “ id_A é a identidade de A ”, “ h é a composta de f e g ” e assim por diante. Em outras palavras os axiomas servem como uma espécie de lista de propriedades às quais as coleções devem satisfazer para formar uma categoria. Note, então, que objetos e flechas não são definidos, quaisquer duas coleções de coisas podem ser objetos e flechas de uma categoria se satisfizerem aos axiomas, a natureza destas coisas não importa.

Uma **afirmação** Σ então é definida como qualquer frase ou fórmula bem formada construída a partir das afirmações atômicas mencionadas e por meio dos conectivos lógicos usuais, “e”, “ou” e “não”, e quantificadores usuais, “existe”, “para todo”. Assim, $f : A \rightarrow B$, por exemplo, é uma abreviação para a afirmação: “ A é a fonte de f e B é o alvo de f ”. Uma **sentença** é uma afirmação onde todas as variáveis estão quantificadas, isto é, não existe nenhuma variável “livre”. Por exemplo, a sentença “para toda flecha f , existem objetos A e B e A é a fonte de f e B é o alvo de f ” é uma sentença. Esta sentença é um dos axiomas de uma categoria, toda flecha possui uma fonte e um alvo. A afirmação anterior, $f : A \rightarrow B$, por outro lado, não é uma sentença, pois f, A e B estão “livres”.

O **dual**, Σ^* de um afirmação, Σ em teoria de categorias é obtida ao realizar as seguintes substituições em Σ : alvo é trocado por fonte e fonte é trocado por alvo e h é a composição de g e f é trocado por h é a composição de f e g . Em outras palavras, flechas e compostas são invertidas. A lógica, os conectivos e quantificadores, permanece inalterada. Note que o dual do dual de qualquer afirmação é igual à afirmação original, isto é $\Sigma^{**} = \Sigma$. Além disso, o dual de uma afirmação envolvendo diagramas é a afirmação envolvendo os diagramas com suas flechas invertidas.

O potencial de se considerar o dual de uma afirmação em teoria de categorias é que os axiomas são autoduais, isto é, se Σ é um axioma então $\Sigma^* = \Sigma$. Isto significa que dada qualquer prova válida de um resultado a partir dos axiomas, o dual desta prova é uma prova válida para o resultado dual. Um exemplo disso é a prova da proposição 3.3, por exemplo, onde para se obter a prova de que todos os objetos finais de uma categoria são isomorfos basta inverter as flechas na prova para objetos iniciais. Inclusive, a própria definição de objetos inicial e final é dual.

Provemos então que os axiomas, de fato, são autoduais. Faremos isto usando a construção de categoria oposta. Dada uma categoria C , a categoria oposta a C , é a categoria C^{op} cujos objetos são os objetos de C e as flechas são as flechas de C porém com a fonte e alvo trocadas. Provar que os axiomas são autoduais equivale a provar que C^{op} de fato é uma categoria. Considere então a correspondência:

$$A \xrightarrow{f} B \quad \longmapsto \quad A \xleftarrow{f^*} B$$

que associa uma flecha f^* em C^{op} para cada flecha f de C . Note que f^* é a mesma flecha que f , nós apenas decidimos que em C^{op} o significado de fonte e alvo são trocados. Se f fosse uma função, por exemplo, f^* seria a mesma regra, porém com domínio e

contradomínio trocados. A composição de em \mathbf{C}^{op} , é a composição de \mathbf{C} porém com a ordem trocada. Como todas as flechas são invertidas, tem-se:

$$\begin{array}{ccc}
 & B & \\
 f \nearrow & & \searrow g \\
 A & \xrightarrow{h = g \circ f} & C
 \end{array}
 \quad \longmapsto \quad
 \begin{array}{ccc}
 & B & \\
 f^* \nwarrow & & \nearrow g^* \\
 A & \xleftarrow{h^* = f^* \circ g^*} & C
 \end{array}$$

Com isto, então temos o suficiente para provar que \mathbf{C}^{op} é uma categoria. Primeiro, para todo objeto A , a flecha identidade de A , $\text{id}_A : A \rightarrow A$ é levada em $\text{id}_A^* : A \rightarrow A$. Note que a flecha identidade permanece inalterada já que seu alvo e fonte são iguais. Assim, tem-se que, para qualquer flecha $f : A \rightarrow B$,

$$\text{id}_B \circ f = f = f \circ \text{id}_A \implies f^* \circ \text{id}_B^* = f^* = \text{id}_A^* \circ f^*$$

Ou seja, todo objeto em \mathbf{C}^{op} possui uma flecha identidade. Além disso, dadas flechas $f : A \rightarrow B, g : B \rightarrow C$ e $h : C \rightarrow D$ em \mathbf{C} , tem-se:

$$(h \circ g) \circ f = h \circ (g \circ f) \implies f^* \circ (h \circ g)^* = (g \circ f)^* \circ h^* = f^* \circ (g^* \circ h^*) = (f^* \circ g^*) \circ h^*.$$

Ou seja, a composição em \mathbf{C}^{op} é associativa. Deste modo, \mathbf{C}^{op} de fato é uma categoria o que implica que os axiomas de categoria são autoduais. Trocar a fonte e o alvo de todas as flechas mantém inalteradas as propriedades das flechas identidades e da operação de composição. Todos os outros axiomas são autoduais, por exemplo, o dual da afirmação de que toda flecha possui uma fonte e um alvo é que toda flecha possui um alvo e uma fonte. Para nitidez, repetimos os axiomas de uma categoria de uma forma que deixa evidente que são autoduais:

Definição 3.10: Uma categoria, \mathbf{C} , consiste em

- Uma coleção de objetos, $\text{Ob}(\mathbf{C})$;
- Uma coleção de flechas, $\text{Flc}(\mathbf{C})$;
- Duas atribuições

$$\text{Fon} : \text{Flc}(\mathbf{C}) \rightarrow \text{Ob}(\mathbf{C}),$$

$$\text{Alv} : \text{Flc}(\mathbf{C}) \rightarrow \text{Ob}(\mathbf{C})$$

que associam a cada flecha f , sua fonte e seu alvo. Escrevemos, $f : A \rightarrow B$ para a flecha cuja fonte é objeto A e o alvo, o objeto B ;

- Uma operação, \circ , chamada composição, que atribui a cada par de flechas (f, g) , tais que $\text{Alv}(f) = \text{Fon}(g)$, uma flecha $g \circ f$ onde $\text{Fon}(g \circ f) = \text{Fon}(f)$ e $\text{Alv}(g \circ f) = \text{Alv}(g)$;

que satisfazem aos seguintes axiomas:

- **Identidade:** para cada objeto A existe uma flecha $\text{id}_A : A \rightarrow A$, chamada identidade em A , tal que para toda flecha $f : A \rightarrow B$, tem-se

$$f \circ \text{id}_A = f = \text{id}_B \circ f.$$

- **Associatividade:** para todas as flechas $f : A \rightarrow B, g : B \rightarrow C$ e $h : C \rightarrow D$ tem-se

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Escrita deste modo, é evidente que os axiomas são autoduais. Assim, enunciamos o princípio da dualidade como

Proposição 3.5: Se uma afirmação Σ vale para todas as categorias, então a afirmação dual, Σ^* também vale.

Note o requerimento que Σ seja uma afirmação sobre categorias que vale para **todas** as categorias, isto significa que Σ é construída apenas com os axiomas de categoria e não utiliza nenhuma propriedade que pode valer para uma dada categoria mas não para outra. Por exemplo, o princípio da dualidade não se aplica a uma sentença que usa do fato de toda flecha mônica ser uma seção, pois, como vimos, isto não vale em certas categorias.

Agora, então, deve ser simples de perceber que os conceitos de seção e retração, flechas mônicas e épicas e objetos finais e iniciais são duais. Qualquer um pode ser obtido de sua contraparte invertendo-se todas as flechas na definição e teoremas que valem para um produzem um teorema dual que vale para o outro. O princípio da dualidade, então, fornece dois conceitos pelo preço de um e dois teoremas pelo preço de um, pois, basta provar o teorema uma vez que o princípio garante que seu dual é válido. Usaremos muito do princípio da dualidade ao tratar de limites adiante.

3.2 – Funtores

O próximo conceito fundamental em teoria de categorias é aquele de funtor ou functor. Funtores desempenham para categorias um papel análogo a homomorfismos de grupo para grupos ou funções contínuas para espaços topológicos, isto é, são mapas entre categorias. Vejamos a definição formal:

Definição 3.11: Sejam \mathbf{A} e \mathbf{B} categorias. Um **funtor** de \mathbf{A} em \mathbf{B} , representado por $F : \mathbf{A} \rightarrow \mathbf{B}$, consiste em

- Uma função

$$\text{Ob}(\mathbf{A}) \rightarrow \text{Ob}(\mathbf{B})$$

escrita como $A \mapsto F(A)$;

- Para cada par $A, A' \in \mathbf{A}$, uma função

$$\mathbf{A}[A, A'] \rightarrow \mathbf{B}[F(A), F(A')]$$

escrita como $f \mapsto F(f)$,

que satisfazem

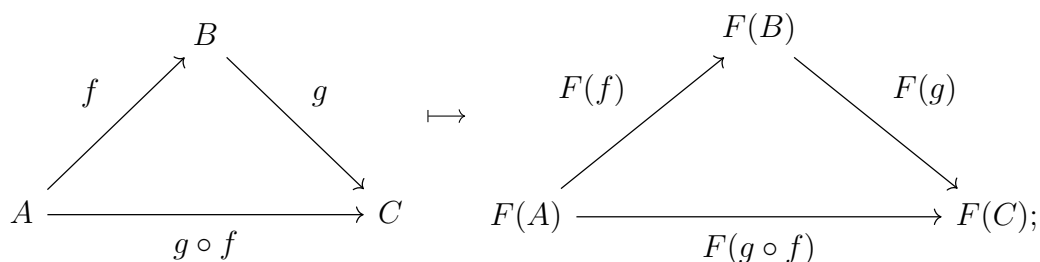
- $F(f' \circ f) = F(f') \circ F(f)$ sempre que $\text{Alv}(f) = \text{Fon}(f')$;
- $F(\text{id}_A) = \text{id}_{F(A)}$, para todo $A \in \mathbf{A}$.

Pela definição é simples de constatar a afirmação anterior, de que funtores são para categorias o que homomorfismos de grupo são para grupos. Funtores são mapas que preservam a “estrutura” de uma categoria, isto é, a composição de flechas e a existência de flechas identidades para cada objeto. Inclusive, pode-se definir a composição de funtores de modo análogo à composição de flechas e obter a categoria \mathbf{Cat} , cujos objetos são categorias pequenas e cujas flechas são funtores entre estas categorias. A restrição a categorias pequenas é para evitar problemas similares ao paradoxo de Cantor, isto é, àqueles encontrados quando se tenta definir o conjunto de todos os conjuntos.

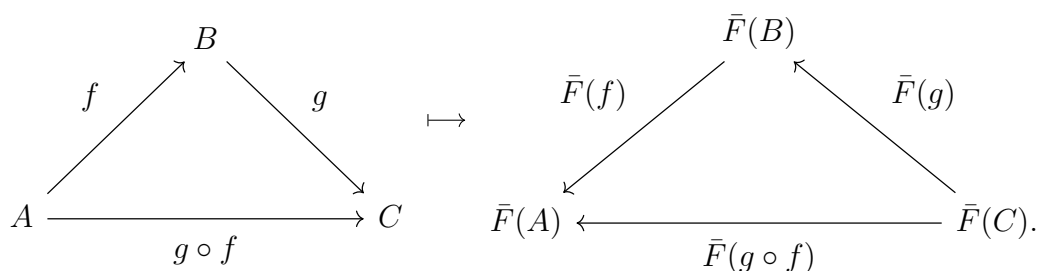
Note que, apesar da definição acima, existem duas possibilidades para um funtor. A definição garante que objetos que estão relacionados em \mathbf{A} , isto é, objetos tais que exista uma flecha entre eles, continuam relacionados em \mathbf{B} . No entanto, a direção da flecha não é essencial para manter esta estrutura. Assim, um funtor pode preservar a direção das flechas de modo que uma flecha $f : A \rightarrow B$ é levada em $F(f) : F(A) \rightarrow F(B)$, ou pode inverter a direção de modo que f é levada em $F(f) : F(B) \rightarrow F(A)$. Deste modo, definimos dois tipos de funtores:

Definição 3.12:

- Um **funtor covariante** é um funtor $F : \mathbf{A} \rightarrow \mathbf{B}$ tal que



- Um **funtor contravariante** é um funtor $\bar{F} : \mathbf{A} \rightarrow \mathbf{B}$ tal que



Repare a semelhança de um funtor contravariante com a ideia de categoria oposta. De fato, um funtor contravariante, $\bar{F} : \mathbf{A} \rightarrow \mathbf{B}$ pode ser escrito como um funtor covariante: $F : \mathbf{A}^{op} \rightarrow \mathbf{B}$, pois as flechas de \mathbf{A}^{op} são exatamente as flechas $f^* : B \rightarrow A$, para toda flecha $f : A \rightarrow B$ em \mathbf{A} . Deste modo, $f^* \mapsto F(f^*) : F(B) \rightarrow F(A) = \bar{F}(f)$. Em vista deste resultado, todo funtor doravante será covariante a menos que especificado o contrário. Assim como homomorfismos de grupos ou funções contínuas entre espaços topológicos, a ideia de funtor é relativamente simples, no entanto são ferramentas essenciais para se estudar categorias. Vejamos então alguns exemplos:

Grupo fundamental

Começamos com um exemplo relevante para este trabalho. Como discutido no B, para cada ponto x de um espaço topológico X , pode-se atribuir um grupo, $\pi_1(X, x)$, o grupo fundamental sitiado em x . Além disso, cada função contínua que preserva o ponto base, isto é, funções $f : (X, x) \rightarrow (Y, y)$, onde Y é outro espaço topológico e $y \in Y$, induz um homomorfismo entre $\pi_1(X, x)$ e $\pi_1(Y, y)$. Como demonstrado no teorema B.5, este homomorfismo satisfaz certas propriedades que chamamos de functoriais. Para entender o motivo deste nome, considere a categoria \mathbf{Top}_* de todos os espaços topológicos pontuados, (X, x) . Existe um funtor

$$\pi_1 : \mathbf{Top}_* \rightarrow \mathbf{Grp}$$

dado por

$$(X, x) \mapsto \pi_1(X, x)$$

$$(f : (X, x) \rightarrow (Y, y)) \mapsto (f_* : \pi_1(X, x) \rightarrow \pi_1(Y, y))$$

O teorema B.5 garante, exatamente que estes dois mapas formem um funtor pois garante que o homomorfismo induzido preserve a composição de mapas entre espaços pontuados, ou seja, preserve a composição de flechas em \mathbf{Top}_* e que preserve as flechas identidade de \mathbf{Top}_* . Ou seja, garante que os dois mapas acima formem um funtor.

Historicamente, os primeiros exemplos de funtores surgiram em estudos de topologia algébrica, pois lá, a estratégia para estudar um dado espaço é reunir dados sobre este espaço de maneira inteligente, de modo a formar alguma estrutura algébrica, como o grupo fundamental por exemplo e então estudar esta estrutura algébrica ao invés do espaço em si. Assim, este ramo está repleto de funtores entre categorias de espaços e categorias de estruturas algébricas.

Funtores esquecimento

Funtor esquecimento ou funtor de esquecimento são talvez os tipos mais simples de funtores. Este é um termo informal que engloba vários funtores que possuem uma propriedade em comum: a de “esquecer” estruturas. Por exemplo, existe um funtor $U : \mathbf{Grp} \rightarrow \mathbf{Set}$ definido do seguinte modo: se G é um grupo, $U(G)$ é o conjunto dos elementos de G e se $f : G \rightarrow H$ é um homomorfismo de grupos, $U(f)$ é a própria função f . Ou seja, U essencialmente esquece as propriedades de grupo de um grupo, ele esquece que existe uma operação de grupo em $U(G)$ e esquece que f é um homomorfismo de grupos. Similarmente, existem funtores $\mathbf{Rng} \rightarrow \mathbf{Set}$ e $\mathbf{Vect}_K \rightarrow \mathbf{Set}$ que esquecem as propriedades de anel e de espaço vetorial sobre um corpo K , respectivamente.

Funtores esquecimento não necessariamente esquecem toda a estrutura dos objetos de uma categoria. Por exemplo, existe um funtor $U : \mathbf{Rng} \rightarrow \mathbf{Ab}$, onde \mathbf{Ab} é a categoria cujos objetos são grupos abelianos. Este funtor esquece multiplicação de um anel, deixando apenas o grupo abeliano subjacente. Do mesmo modo, existe um funtor $U : \mathbf{Rng} \rightarrow \mathbf{Mon}$ que esquece a adição no anel, deixando apenas o monoide multiplicativo.

Por fim, também existem funtores esquecimento que se esquecem de propriedades ao invés de estrutura. Por exemplo, o funtor $U : \mathbf{Ab} \rightarrow \mathbf{Grp}$, dado por $U(A) = A$ e $U(f) = f$ onde A é um grupo abeliano e f é um homomorfismo entre grupos abelinos, se esquece que grupos abelianos são abelianos. A operação de esquecer parece trivial mas por vezes pode ser bastante útil. Na física, por exemplo, frequentemente passa-se da interpretação de um espaço como simplesmente um conjunto de pontos para um espaço vetorial e vice-versa. O mesmo ocorre quando se trata de estruturas algébricas que são “combinações”

de outras, por exemplo, um anel, que é um grupo e um monoide.

Funtores Hom

Existem dois funtores relacionados que são bastante importantes, um é covariante e o outro, contravariante. Para defini-los, considere uma categoria localmente pequena \mathcal{C} , isto é, uma categoria onde a coleção de flechas entre quaisquer dois objetos formam um conjunto, e seja K um objeto de \mathcal{C} . Para cada objeto $A \in \mathcal{C}$, tem-se os conjuntos de flechas

$$L(A) = \mathcal{C}[K, A] \quad \text{e} \quad R(A) = \mathcal{C}[A, K].$$

Assim, definimos os pares de atribuições:

$$\begin{array}{ccc} \mathcal{C} & \longrightarrow & \text{Set} & & \mathcal{C} & \longrightarrow & \text{Set} \\ A & \longmapsto & L(A); & & A & \longmapsto & R(A). \end{array}$$

Estas são as atribuições para elementos de \mathcal{C} . Elas mapeiam cada elemento A de \mathcal{C} no conjunto de flechas cujo alvo é A , $L(A)$ e no conjunto de flechas cuja fonte é A , $R(A)$. Além disso, cada flecha

$$A \xrightarrow{f} B$$

de \mathcal{C} é mapeada em

$$\begin{array}{ccc} \mathcal{C}[K, A] & \xrightarrow{L(f)} & \mathcal{C}[K, B] & & \mathcal{C}[B, K] & \xrightarrow{R(f)} & \mathcal{C}[A, K] \\ r & \longmapsto & f \circ r; & & l & \longmapsto & l \circ f; \end{array}$$

Escrevemos $L(f)(r) = f \circ r$ e $R(f)(l) = l \circ f$. Note que L é covariante e R é contravariante.

Provemos que estes conjuntos de atribuições realmente constituem funtores. Considere flechas

$$A \xrightarrow{f} B \xrightarrow{g} C$$

de \mathcal{C} e flechas

$$K \xrightarrow{r} A \quad \text{e} \quad A \xrightarrow{l} K$$

de $L(A)$ e $R(A)$ respectivamente. Então

$$\begin{array}{ll} (L(g) \circ L(f))(r) = L(g)(L(f)(r)) & (R(f) \circ R(g))(l) = R(f)(R(g)(l)) \\ = L(g)(f \circ r) & = R(f)(l \circ g) \\ = g \circ (f \circ r) & = (l \circ g) \circ f \\ = (g \circ f) \circ r & = l \circ (g \circ f) \\ = L(g \circ f)(r); & = R(g \circ f)(l). \end{array}$$

A prova de que $L(\text{id}_A) = \text{id}_{L(A)}$ e $R(\text{id}_A) = \text{id}_{R(A)}$ é direta, pois para todas as flechas $r \in L(A)$ e $l \in R(A)$, tem-se $L(\text{id}_A)(r) = r$ e $R(\text{id}_A)(l) = l$, ou seja, as flechas identidades de \mathbf{C} são mapeadas nos mapas identidade em $L(A)$ e $R(A)$. Assim, os dois pares de atribuições dados definem funtores.

Por fim, notamos que o nome funtores hom é uma tradução de *Hom-Functors*. Este nome, por sua vez está relacionado ao nome dos conjuntos $\mathbf{C}[A, B]$ de flechas, que também são chamados de *Hom-sets*. Esta terminologia é originária da ideia de que flechas são como homomorfismos entre estruturas algébricas no entanto não é mais muito usado justamente porque $\mathbf{C}[A, B]$ não necessariamente são conjuntos, isto só vale para categorias pequenas ou localmente pequenas. A importância de funtores hom está em sua relação com o lema de Yoneda, um dos resultados fundamentais de teoria de categorias. Nos depararemos com este lema mais a frente.

Categorias produto

O último exemplo que trataremos nesta seção é o de produto de categorias, uma construção muito similar ao produto cartesiano de conjuntos e que pode ser usada para definir “funtores de duas variáveis”, chamados **bifuntores**. Como dito, a construção do produto de categorias é similar ao produto cartesiano, dadas duas categorias \mathbf{A} e \mathbf{B} , podemos formar uma nova categoria $\mathbf{A} \times \mathbf{B}$, chamada **categoria produto** de \mathbf{A} e \mathbf{B} . Um objeto de $\mathbf{A} \times \mathbf{B}$ é um par $\langle A, B \rangle$ de objetos com $A \in \mathbf{A}$ e $B \in \mathbf{B}$ e uma flecha $\langle A, B \rangle \rightarrow \langle A', B' \rangle$, um par $\langle f, g \rangle$ de flechas $f : A \rightarrow A'$ e $g : B \rightarrow B'$. A composição de flechas

$$\langle A, B \rangle \xrightarrow{\langle f, g \rangle} \langle A', B' \rangle \xrightarrow{\langle f', g' \rangle} \langle A'', B'' \rangle$$

é definida em termos das composições em \mathbf{A} e \mathbf{B} por

$$\langle f', g' \rangle \circ \langle f, g \rangle = \langle f' \circ f, g' \circ g \rangle.$$

Assim como existem funções projeção no produto cartesiano, que são funtores, pois se aplicam tanto aos objetos, conjuntos, quanto às flechas, funções entre estes conjuntos, aqui também definimos funtores projeção

$$\mathbf{A} \xleftarrow{P} \mathbf{A} \times \mathbf{B} \xrightarrow{Q} \mathbf{B}.$$

A ação dos funtores P e Q é dada por

$$\begin{array}{ccc} \langle A, B \rangle & \xrightarrow{P} & A \\ \langle f, g \rangle & \xrightarrow{\quad} & f; \end{array} \quad \text{e} \quad \begin{array}{ccc} \langle A, B \rangle & \xrightarrow{Q} & B \\ \langle f, g \rangle & \xrightarrow{\quad} & g. \end{array}$$

Assim como o produto cartesiano, esta construção possui a seguinte propriedade “universal”: dado uma categoria \mathbf{C} e funtores

$$A \xleftarrow{R} C \xrightarrow{T} B,$$

existe um único funtor $F : C \rightarrow A \times B$ tal que $P \circ F = R$ e $Q \circ F = T$. A composição de funtores é definida analogamente à composição de flechas, isto é dados funtores $F_1 : A \rightarrow B$ e $F_2 : B \rightarrow C$, tem-se que $F_2 \circ F_1 : A \rightarrow C$ é dado por

$$\begin{aligned} A &\longmapsto F_1(F_2(A)) \\ f &\longmapsto F_1(F_2(f)). \end{aligned}$$

A propriedade universal, então é representada no seguinte diagrama:

$$\begin{array}{ccccc} & & C & & \\ & R \swarrow & \vdots F \downarrow & \searrow T & \\ A & \xleftarrow{P} & A \times B & \xrightarrow{Q} & B \end{array}$$

onde usamos a linha pontilhada para significar “existe e é único”. Explicitamente, este diagrama indica que para toda flecha h de \mathbf{C} , $F(h) = \langle R(h), T(h) \rangle$ e, por outro lado, esta construção torna F um funtor com a propriedade desejada. Note a semelhança com o produto cartesiano, isto ocorre porque ambas as construções são casos particulares de um tipo específico de **limite** em uma categoria. Trataremos deste assunto mais à frente.

Assim como o produto cartesiano define um produto de funções, também o produto de categorias define um produto de funtores. Dados dois funtores $U : A \rightarrow A'$ e $V : B \rightarrow B'$, definimos o produto $U \times V : A \times B \rightarrow A' \times B'$ por

$$\begin{aligned} \langle A, B \rangle &\longmapsto \langle U(A), V(B) \rangle \\ \langle f, g \rangle &\longmapsto \langle U(f), V(g) \rangle. \end{aligned}$$

Alternativamente, este funtor é descrito como o único funtor tal que o seguinte diagrama comute:

$$\begin{array}{ccccc} A & \xleftarrow{P} & A \times B & \xrightarrow{Q} & B \\ \downarrow U & & \vdots U \times V \downarrow & & \downarrow V \\ A' & \xleftarrow{P'} & A' \times B' & \xrightarrow{Q'} & B' \end{array}$$

Assim, o produto \times é, na verdade, um par de atribuições: para cada par de categorias $\langle \mathbf{B}, \mathbf{C} \rangle$, atribui uma categoria $\mathbf{B} \times \mathbf{C}$ e para cada par de funtores $\langle U, V \rangle$, um funtor $U \times V$. Além disso, quando as composições $U' \circ U$ e $V' \circ V$ estão definidas, tem-se que $(U' \times V') \circ (U \times V) = (U' \circ U) \times (V' \circ V)$. Assim, a operação \times é, também, um funtor. Mais precisamente, nos restringindo a categorias localmente pequenas, é um funtor:

$$\times : \mathbf{Cat} \times \mathbf{Cat} \rightarrow \mathbf{Cat}.$$

Por fim, esta discussão incluiu funtores $F : \mathbf{C} \rightarrow \mathbf{A} \times \mathbf{B}$ para uma categoria produto. Funtores $S : \mathbf{A} \times \mathbf{B} \rightarrow \mathbf{C}$, a partir de uma categoria produto, são chamados **bifuntores**. Tais funtores ocorrem frequentemente, por exemplo, produto cartesiano $X \times Y$ é a atribuição de objetos de um bifuntor $\mathbf{Set} \times \mathbf{Set} \rightarrow \mathbf{Set}$. Assim, a definição de produto fornece imediatamente a definição de “funtor de duas variáveis” assim como a definição do produto de espaços topológicos fornece a de função contínua de duas variáveis, por exemplo. Bifuntores também serão úteis na descrição de anyons.

3.3 – Transformações naturais

Passamos agora ao terceiro dos conceitos básicos em teoria de categorias: transformações naturais. O conceito de transformação natural pode ser um pouco surpreendente de início, pois tratam de mapas entre funtores. Na filosofia de teoria de categorias, estudamos objetos a partir de como se relacionam com outros, ou seja por meio das flechas entre estes objetos. Assim, flechas fornecem um meio de comparar objetos entre si. Similarmente, funtores fornecem meios de comparar categorias entre si e, agora, transformações naturais fornecem meios de comparar funtores entre si.

Para entender melhor a ideia de comparar funtores, considere dois funtores $F, G : \mathbf{A} \rightarrow \mathbf{B}$, onde \mathbf{A} é uma categoria discreta cujos objetos são números naturais, isto é, os objetos de \mathbf{A} são os números $0, 1, 2, \dots$ e as flechas são somente as flechas identidade para cada elemento, e \mathbf{B} é uma categoria arbitrária. Os funtores F e G , então, correspondem apenas a sequências, (F_0, F_1, F_2, \dots) e (G_0, G_1, G_2, \dots) de objetos em \mathbf{B} , já que \mathbf{A} não possui nenhuma flecha além das flechas identidade. Seria razoável, então, definir um mapa entre os funtores F e G como uma sequência de mapas

$$(F_0 \xrightarrow{\alpha_0} G_0, F_1 \xrightarrow{\alpha_1} G_1, F_2 \xrightarrow{\alpha_2} G_2, \dots)$$

Em outras palavras, comparamos, para cada objeto $A \in \mathbf{A}$, os objetos $F(A)$ e $G(A)$ de \mathbf{B} . Se \mathbf{A} possuísse flechas não triviais, compararíamos, para cada flecha $f : A \rightarrow A'$ de \mathbf{A} , as flechas $F(f) : F(A) \rightarrow F(A')$ e $G(f) : G(A) \rightarrow G(A')$ por meio dos mapas α_A e $\alpha_{A'}$, exigindo uma certa “compatibilidade” entre os mapas α_A e as flechas de \mathbf{A} . Vejamos,

então a definição formal de transformações naturais:

Definição 3.13: Sejam \mathbf{A} e \mathbf{B} categorias e $F, G : \mathbf{A} \rightarrow \mathbf{B}$ funtores covariantes entre estas categorias. Uma **transformação natural** $\alpha : F \rightarrow G$, entre os funtores F e G , é uma família

$$\left(F(A) \xrightarrow{\alpha_A} G(A) \right)_{A \in \mathbf{A}}$$

de mapas em \mathbf{B} tal que para toda flecha $f : A \rightarrow A'$ em \mathbf{A} , o diagrama

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(A') \\ \alpha_A \downarrow & & \downarrow \alpha_{A'} \\ G(A) & \xrightarrow{G(f)} & G(A') \end{array}$$

comute. Os mapas α_A são chamados **componentes** de α .

Uma transformação natural, α , usualmente é representada em diagramas como:

$$\begin{array}{ccc} & F & \\ & \curvearrowright & \\ A & \Downarrow \alpha & B. \\ & \curvearrowleft & \\ & G & \end{array}$$

Como transformações naturais são mapas, pode-se perguntar se é possível formar composições. A resposta é sim, dada duas transformações naturais, α e β :

$$\begin{array}{ccc} & F & \\ & \curvearrowright & \\ A & \xrightarrow{\quad} G \xrightarrow{\quad} B. & \\ & \curvearrowleft & \\ & H & \end{array}$$

formamos a composição

$$\begin{array}{ccc} & F & \\ & \curvearrowright & \\ A & \Downarrow \beta \circ \alpha & B. \\ & \curvearrowleft & \\ & G & \end{array}$$

definida por $(\beta \circ \alpha)_A = \beta_A \circ \alpha_A$ para todo $A \in \mathbf{A}$. Além disso, também existe uma transformação natural identidade

$$\begin{array}{ccc}
 & F & \\
 & \curvearrowright & \\
 A & & B \\
 & \curvearrowleft & \\
 & F & \\
 & \Downarrow \text{id}_F & \\
 & &
 \end{array}$$

para qualquer functor F , definida por $(\text{id}_F)_A = \text{id}_{F(A)}$. Assim, para quaisquer duas categorias \mathbf{A} e \mathbf{B} , existe uma categoria cujos objetos são os funtores $F : \mathbf{A} \rightarrow \mathbf{B}$ e cujas flechas são as transformações naturais entre estes funtores. Tal categoria é chamada de **categoria de funtores** de \mathbf{A} em \mathbf{B} e é escrita $[\mathbf{A}, \mathbf{B}]$ ou $\mathbf{B}^{\mathbf{A}}$. Como transformações naturais são flechas em uma categoria, podemos definir **isomorfismos naturais**, que correspondem a transformações naturais que são inversíveis. Definimos, formalmente como

Definição 3.14: Dadas duas categorias \mathbf{A} e \mathbf{B} , um **isomorfismo natural** entre funtores $F, G : \mathbf{A} \rightarrow \mathbf{B}$ é uma transformação natural α tal que todos as componentes α_A são inversíveis em \mathbf{B} .

Formalizamos, então, o conceito de “naturalidade” de modo a significar que: uma construção é natural se equivale a um isomorfismo natural, como definido. Vejamos um exemplo que trata de transformações naturais e o conceito de naturalidade e como apenas a existência de uma transformação natural não fornece uma construção que é natural.

Espaços vetoriais finitos

Considere a categoria FDVect_K de espaços vetoriais de dimensão finita sobre um corpo K . Todo espaço vetorial, V , possui um dual, V^* definido como o espaço de todas os mapas K -lineares sobre V , isto é, de todos os mapas $L : V \rightarrow K$, em outras palavras, $V^* = [V, K]$. O dual, V^* também é um espaço vetorial, com a mesma dimensão de V e, portanto, um objeto de FDVect_K e assim dá origem a um mapa

$$\begin{array}{ccc}
 \text{FDVect}_K & \longrightarrow & \text{FDVect}_K \\
 V & \longmapsto & [V, K] = V^*
 \end{array}$$

que leva cada espaço V em seu dual. Além disso, um mapa $l : V \rightarrow W$ entre espaços vetoriais, uma flecha em FDVect_K , dá origem a um mapa linear entre os duais de W e V , $l^* : W^* \rightarrow V^*$, dado por $l^*(m) = m \circ l$, onde $m \in W^* = [W, K]$ é um covetor de W e \circ , aqui, é a composição usual de mapas entre conjuntos. Deste modo, um mapa linear l gera o seguinte mapa entre flechas de FDVect_K :

$$\begin{array}{ccc} \text{FDVect}_K^{op} & \longrightarrow & \text{FDVect}_K \\ l & \longmapsto & l^* \end{array}$$

Podemos, então, agrupar estes dois mapas em um functor contravariante, $(-)^* : \text{FDVect}_K^{op} \rightarrow \text{FDVect}_K$, tal que

$$\begin{array}{ccc} V & \longmapsto & V^* \\ (l : V \rightarrow W) & \longmapsto & (l^* : W^* \rightarrow V^*). \end{array}$$

Para verificar que $(-)^*$ de fato é um functor, considere dois mapas

$$V \xrightarrow{l_1} W \xrightarrow{l_2} Z.$$

Tem-se

$$\begin{aligned} (l_1^* \circ l_2^*)(m) &= l_1^*(l_2^*(m)) = l_1^*(m \circ l_2) \\ &= (m \circ l_2) \circ l_1 \\ &= m \circ (l_2 \circ l_1) = (l_2 \circ l_1)^*(m). \end{aligned}$$

Além disto, tomando $Z = W$ e $l_2 = \text{id}_W$, tem-se, pela igualdade anterior que

$$(l_1^* \circ \text{id}_W^*) = (\text{id}_W \circ l_1)^* = l_1$$

e, similarmente, tomando $W = V$ e $l_1 = \text{id}_V$, tem-se:

$$(\text{id}_V^* \circ l_2^*) = (l_2 \circ \text{id}_V)^* = l_2,$$

logo, $\text{id}_V^* = \text{id}_{V^*}$, de modo que $(-)^*$ é, de fato, um functor. A ação de $(-)^*$ está ilustrada nos diagramas abaixo.

$$\begin{array}{ccc} \begin{array}{ccc} & W & \\ l_1 \nearrow & & \searrow l_2 \\ V & \xrightarrow{l_2 \circ l_1} & Z \end{array} & \longmapsto & \begin{array}{ccc} & W^* & \\ l_1^* \nwarrow & & \nearrow l_2^* \\ V^* & \xleftarrow{(l_2 \circ l_1)^*} & Z^* \end{array} \end{array}$$

Agora, definimos o functor $(-)^{**}$ como a composição $(-)^* \circ (-)^*$, de modo que

$$\begin{array}{ccc} V & \longmapsto & V^{**} \\ (l : V \rightarrow W) & \longmapsto & (l^{**} : V^{**} \rightarrow W^{**}) \end{array}$$

onde $l^{**}(v) = v \circ l^*$ e $v : V^* \rightarrow K$ pertence a V^{**} . É um resultado elementar de álgebra linear que existe um isomorfismo canônico entre os espaços vetoriais V e V^* , isto é, existe um mapa K -linear bijetivo, $\alpha : V \rightarrow V^{**}$, entre V e V^{**} . Este mapa é dado por

$$v \longmapsto ev_v,$$

onde $ev_v : V^* \rightarrow K$ é o mapa que avalia elementos de V^* em v , isto é:

$$\begin{aligned} ev_v : V^* &\longrightarrow K \\ (f : V \rightarrow K) &\longmapsto f(v). \end{aligned}$$

O fato de α ser canônica corresponde justamente ao fato de ser um isomorfismo natural entre id , o funtor identidade em FDVect_K e $(-)^{**}$. Para ver isto, considere o seguinte diagrama mostrando a condição de naturalidade:

$$\begin{array}{ccc} V & \xrightarrow{l} & W \\ \alpha_V \downarrow & & \downarrow \alpha_W \\ V^{**} & \xrightarrow{l^{**}} & W^{**} \end{array}$$

onde reescrevemos α como uma família de mapas indexada pelos objetos de FDVect_K , α_V . Para mostrar que $\alpha_V, V \in \text{FDVect}_K$ de fato é um isomorfismo natural, devemos mostrar que o diagrama acima comuta. Note que $\alpha_W \circ l : V \rightarrow W^*$ e leva $v \mapsto ev_{l(v)} : W^* \rightarrow K$, assim, se $v \in V$ e $W^* \ni f : W \rightarrow K$, tem-se:

$$\begin{aligned} (\alpha_W \circ l)(v)(f) &= [\alpha_W(l(v))](f) = ev_{l(v)}(f) = f(l(v)) \\ &= (f \circ l)(v) = ev_v(f \circ l) = ev_v(l^*(f)) \\ &= (ev_v \circ l^*)(f) = (l^{**}(ev_v))(f) = (l^{**} \circ \alpha_V(v))(f) = (l^{**} \circ \alpha_V)(v)(f). \end{aligned}$$

Como α_V é um mapa K -linear bijetivo para todo V , tem-se que α_V é um isomorfismo natural entre id e $(-)^{**}$.

Dissemos que o isomorfismo de espaços vetoriais α é canônico, também referido como natural, justamente porque a família α_V de isomorfismos forma um isomorfismo natural. Elaboremos sobre isto. Primeiro, note que os termos “natural” e “canônico” são usados, informalmente, desde antes do surgimento de teoria de categorias, significando algo como “não necessita de escolhas arbitrárias”. Por exemplo, dados dois conjuntos A e B , existe um isomorfismo canônico entre $A \times B$ e $B \times A$, dado por $(a, b) \mapsto (b, a)$. Similarmente, existe uma função canônica $A \times B \rightarrow A$ dada por $(a, b) \mapsto a$, a projeção canônica em A . Por outro lado, a função $f : B \rightarrow A$ definida por: escolha um elemento $a_0 \in A$ e faça $f(b) = a_0$ para todo $b \in B$ não é canônica, ou natural, porque envolve escolher um

elemento de A . Peça a duas pessoas para escolherem uma função $f : B \rightarrow A$ deste tipo e provavelmente escolherão funções distintas. Não existe senso em que um a_0 é melhor que outro, não existe escolha óbvia.

A situação é similar quando se trata de um espaço vetorial V e seu dual, V^* . Em virtude de possuírem a mesma dimensão, V e V^* são espaços vetoriais isomorfos, porém, qualquer isomorfismo entre V e V^* depende de uma escolha de base, de modo que qualquer um destes isomorfismos dificilmente seria canônico. Por outro lado, quando se trata do dual do dual, V^{**} , existe uma escolha óbvia, pois, para cada elemento $v \in V$, podemos definir a função

$$\begin{aligned} \hat{v} : V^* &\longrightarrow K \\ (f : V \rightarrow K) &\longmapsto f(v), \end{aligned}$$

exatamente o isomorfismo de espaços vetoriais definido anteriormente, apenas reescrito de modo a enfatizar v . Todo elemento de V^{**} pode ser escrito como \hat{v} para algum $v \in V$. Já vimos que este mapa é um isomorfismo entre V e V^{**} , porém, note que diferente do isomorfismo entre V e V^* , não é necessário especificar uma base em V para defini-lo, podemos tratar dos vetores diretamente. A questão, então, era: porque o segundo dual possui este comportamento “natural” enquanto o primeiro não? Obtivemos a resposta acima.

O segundo dual possui este comportamento natural porque existe um isomorfismo natural entre id e $(-)^{**}$, ou seja, os objetos de FDVect_K , V e V^{**} são naturalmente isomorfos, isto é, além de existirem flechas inversíveis entre eles, a família destas flechas indexadas por objetos V forma uma transformação natural. Por outro lado, não pode é possível que V e V^* sejam naturalmente isomorfos porque não pode existir nem mesmo transformações naturais entre os funtores id e $(-)^*$, já que são funtores cujas categorias fontes são distintas, id é um funtor cuja fonte é FDVect_K enquanto a fonte de $(-)^*$ é FDVect_K^{op} . Por este motivo, entende-se que isomorfismos naturais são a formalização do conceito de “naturalidade” ou “canonicidade”. Mantenha em mente que exibimos apenas um exemplo, no entanto esta conclusão foi alcançada após se estudar diversos casos distintos onde isto ocorre.

Por fim, notamos que apesar de ser possível formalizar a ideia de “naturalidade”, afinal este foi um dos motivos pelos quais teoria de categorias foi inventada, isto não significa que as palavras “natural” e “canônica” são sempre usadas neste sentido rigoroso. Além disso, uma construção ser natural não significa que é única. Se olharmos puramente para a definição de isomorfismo natural, uma construção ser natural significa que ela é preservada sob flechas entre objetos de uma categoria. Neste sentido, então, diz respeito à possibilidade de ser definida apenas em termos de propriedades gerais dos objetos sob estudo, não dependendo de nenhuma particularidade de um ou outro objeto.

3.4 - Universalidade e limites

Neste ponto já introduzimos três conceitos básicos: categorias, funtores e transformações naturais. O único conceito que mencionamos e ainda não definimos é o de **universalidade**, que definiremos nesta seção. Antes disto, no entanto, alguns comentários. Dissemos, anteriormente, que teoria de categorias surgiu para sustentar o estudo de transformações naturais. Isto não é inteiramente verdade, pode-se dizer que o estudo de transformações naturais surgiu para sustentar o conceito de **adjunções** [37]. Não trataremos deste assunto aqui, porém mencionamos que adjunções fornecem uma maneira de entender o conceito de **universalidade** que mencionamos brevemente no início deste capítulo.

Outro modo de entender universalidade é por meio de **funtores representáveis** e pelo **lema de Yoneda**, que talvez seja o resultado mais importante em teoria de categorias básica. Também não trataremos deste assunto aqui, apesar de que nos depararemos com o lema de Yoneda ao tratar de anyons. Em geral, estes dois conceitos restantes são os que faltam para completar um curso introdutório em teoria de categorias. No entanto, o conceito de universalidade pode ser compreendido razoavelmente bem sem usar nenhuma destas duas vias. O preço a se pagar é que a definição de universalidade fornecida pode não ser muito própria se estudar construções universais, isto é, determinar se uma construção é universal ou não, por exemplo. Apesar disto, será suficiente para definir algumas estruturas que serão necessárias para tratar de anyons, por isto decidimos que esta abordagem é mais adequada.

Tratemos o conceito de universalidade junto com o de **limites**. Começaremos pela definição de **flecha universal** e então a usamos para definir limites em uma categoria C de modo geral. Esta definição é um tanto abstrata e por isto passaremos, logo em seguida, a tratar de exemplos básicos de limites. Estes exemplos são justamente aqueles dos quais precisaremos no próximo capítulo além de que servirão para ilustrar porque a ideia de universalidade que estamos tratando. Pois bem, comecemos com alguns conceitos preliminares.

O primeiro passo é notar que existe um functor esquecimento $U : \text{Cat} \rightarrow \text{Grph}$, isto é, um functor esquecimento da categoria de todas as categorias pequenas para a categoria de todos os grafos pequenos. Um grafo, G , é consiste em um conjunto \mathbb{V} de vértices ou nodos, um conjunto \mathbb{A} de arestas. Um aresta é simplesmente um conjunto $\{a, b\}$ de vértices. Um grafo direcionado é um grafo junto com dois mapas

$$\text{Fon} : \mathbb{A} \rightarrow \mathbb{V},$$

$$\text{Alv} : \mathbb{A} \rightarrow \mathbb{V}$$

que atribuem a cada aresta uma fonte e um alvo. Note a semelhança com uma categoria.

O funtor U , então, simplesmente esquece que flechas compostas são compostas e flechas identidades são identidades e produz um grafo direcionado onde os vértices são os objetos da categoria e as arestas, as flechas. Abaixo indicamos uma categoria de três objetos, $\mathbf{3}$, e o grafo direcionado subjacente:

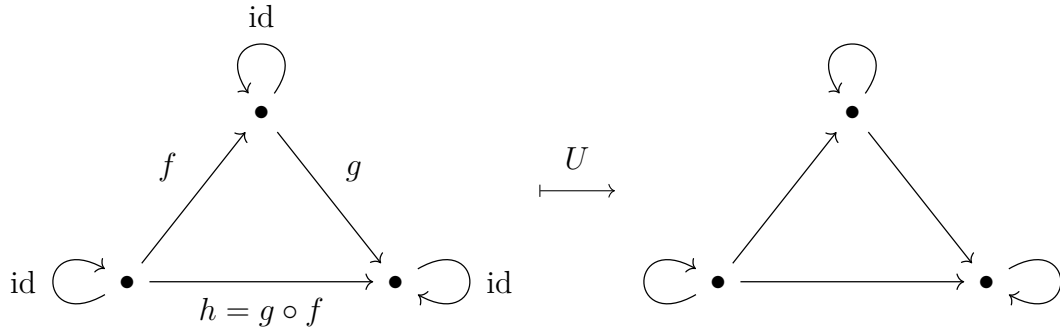


Figura 3.4: Grafo direcionado subjacente à categoria de três objetos.

Note que a categoria de três objetos possui exatamente seis flechas, as três flechas identidade e exatamente uma flecha entre cada um dos três pares possíveis de elementos. Note que a flecha h é obrigatoriamente a composição de f e g para que tenhamos, de fato uma categoria. Por outro lado, o grafo subjacente não se importa com o fato de h ser ou não a composição de f e g , inclusive, nem mesmo nomeamos as arestas no grafo.

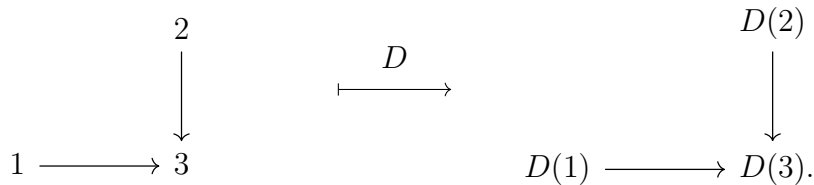
Podemos, então, pensar o grafo como o diagrama subjacente de $\mathbf{3}$. Usando esta ideia, dado um grafo G , podemos definir diagramas de tipo ou formato G ou G -diagramas em uma categoria \mathbf{C} pela seguinte construção: para cada vértice v , considere uma família de objetos $D(v) \in \mathbf{C}$ e para cada aresta a , uma família de flechas $D(a)$ em \mathbf{C} tal que, cada aresta

$$v_1 \xrightarrow{a} v_2$$

produza uma flecha

$$D(v_1) \xrightarrow{D(a)} D(v_2)$$

na categoria \mathbf{C} . Note a semelhança desta construção com um funtor. Por exemplo, considerando-se o grafo P abaixo, obtemos o seguinte diagrama em uma categoria \mathbf{C} :



Dizemos que o grafo P é um modelo para o P -diagrama em \mathbf{C} .

Com esta ideia, então, definimos formalmente um G -diagrama em uma categoria arbitrária \mathbf{C} , trocando o grafo G por uma categoria pequena \mathbf{G} e realizando o mesmo processo

descrito acima. Continuamos visualizando a categoria \mathbf{G} pelo seu grafo subjacente, G , porém, agora, um G -diagrama define um funtor, D , entre \mathbf{G} e \mathbf{C} . Formalmente:

Definição 3.15: Sejam \mathbf{C} uma categoria arbitrária, \mathbf{G} uma categoria pequena e G seu grafo subjacente. Um G -**diagrama** ou um **diagrama de formato** G em \mathbf{C} é um funtor $D : \mathbf{G} \rightarrow \mathbf{C}$.

Note que a coleção de todos os G -diagramas em uma categoria \mathbf{C} dá origem a uma categoria de funtores, $\mathbf{C}^{\mathbf{G}}$, onde os objetos são G -diagramas e as flechas são transformações naturais entre estes diagramas. Note que para uma dada categoria \mathbf{C} em geral existem inúmeros diagramas possíveis. Por exemplo, considerando a categoria \mathbf{P} cujo grafo subjacente é o grafo P anterior e uma categoria \mathbf{C} que possui, digamos, seis objetos, então existem, possivelmente, cinco P -diagramas possíveis, um para cada escolha de conjuntos de três objetos possíveis de \mathbf{C} . Se existirão, de fato cinco P -diagramas em \mathbf{C} , depende das flechas de \mathbf{C} .

Agora podemos definir flechas universais. A seguir, apresentamos a definição formal de flecha universal segundo Mac Lane [38] e então apresentaremos uma espécie de definição alternativa segundo Awodey [37] para conectar a primeira definição com a ideia de diagramas apresentada anteriormente.

Definição 3.16: Sejam $D : \mathbf{G} \rightarrow \mathbf{C}$ um funtor e C um objeto de \mathbf{C} . Uma **flecha universal** de C a D é um par $\langle V, f_V \rangle$, onde V é um objeto de \mathbf{G} e $f_V : C \rightarrow D(V)$ é uma flecha em \mathbf{C} de C a $D(V)$ satisfazendo à seguinte propriedade, comumente chamada de **propriedade universal**: para cada par $\langle G, f \rangle$ de objetos $G \in \mathbf{G}$ e flechas $f : C \rightarrow D(G)$ de \mathbf{C} , existe uma única flecha $m : V \rightarrow G$ tal que $D(m) \circ f_V = f$. Diagramaticamente:

$$\begin{array}{ccc}
 & D(V) & \\
 f_V \nearrow & \downarrow D(m) & \\
 C & & \\
 f \searrow & D(G) & \\
 & & \\
 & & \begin{array}{c} V \\ \downarrow m \\ G \end{array}
 \end{array}$$

Notamos que pode-se obter uma segunda definição, a de uma **flecha universal** de D a C , considerando-se o dual da definição anterior. Além disso, na definição acima, mantivemos a notação consistente com a discussão anterior sobre diagramas. Agora, para entender melhor esta definição, passemos à abordagem de Awodey que acabará nos levando à definição de limite. Segundo Awodey, todo diagrama D em uma categoria \mathbf{C} propõe um problema porém, usualmente estamos mais interessados na solução do que no problema de modo que raramente falamos sobre este. Cada diagrama, inclusive, propõe

dois problemas, relacionados por dualidade e, conseqüentemente, duas soluções.

Definição 3.17: Seja $D : \mathbf{G} \rightarrow \mathbf{C}$ um G -diagrama em uma categoria \mathbf{C} . Uma **solução esquerda** e **solução direita** correspondem, respectivamente, a um objeto C de \mathbf{C} e uma família de flechas de \mathbf{C} :

$$C \xrightarrow{\alpha(V)} D(V); \quad D(V) \xrightarrow{\alpha^*(V)} C$$

indexadas pelos objetos, V , de \mathbf{G} — vértices de G — tal que para cada flecha, $a : V_1 \rightarrow V_2$ de \mathbf{G} — aresta no grafo subjacente — os triângulos abaixo comutam.

$$\begin{array}{ccc}
 & D(V_1) & \\
 \alpha(V_1) \nearrow & \downarrow D(a) & \\
 C & & \\
 \alpha(V_2) \searrow & & \\
 & D(V_2) &
 \end{array}
 \quad
 \begin{array}{ccc}
 D(V_1) & & \\
 \downarrow D(a) & & \\
 D(V_2) & &
 \end{array}
 \quad
 \begin{array}{ccc}
 & D(V_1) & \\
 & \downarrow D(a) & \alpha^*(V_1) \searrow \\
 & D(V_2) & \nearrow \alpha^*(V_2) \\
 & & C.
 \end{array}$$

Uma solução esquerda é geralmente chamada de **cone** e uma solução direita de **cocone**.

Note que uma solução é um objeto e uma coleção de flechas que “comutam” com o diagrama D . Para cada elemento V de \mathbf{G} que dá origem a um vértice do G -diagrama em \mathbf{C} e para cada aresta $a : V_1 \rightarrow V_2$ de G , existem flechas α de C a V_1 e V_2 que comutam com a flecha $D(a)$ gerada pela aresta a . Além disso, um diagrama pode ter várias, uma ou nenhuma solução esquerda e várias, uma ou nenhuma solução direita. Em geral não existe relação entre soluções esquerdas e direitas para um diagrama, apesar de serem conceitos duais.

Por fim, uma **solução universal** é uma solução que é a “melhor possível” entre as soluções de um dado tipo, esquerda ou direita, quando existem. Pode-se entender uma solução universal como o meio mais otimizado de encontrar o objeto C , quando ele existe. Vejamos a definição:

Definição 3.18: Seja $D : \mathbf{G} \rightarrow \mathbf{C}$ um G -diagrama em uma categoria \mathbf{C} . Uma **solução universal esquerda**, usualmente chamada de **limite**, e **solução universal direita**, usualmente chamada de **colimite**, correspondem, respectivamente, a uma solução particular

$$L \xrightarrow{\lambda(V)} D(V); \quad D(V) \xrightarrow{\lambda^*(V)} L$$

tal que para cada solução

$$C \xrightarrow{\alpha(V)} D(V); \quad D(V) \xrightarrow{\alpha^*(V)} C$$

existe uma única flecha

$$C \xrightarrow{m} L; \quad L \xrightarrow{m} C$$

chamada de **flecha mediadora**, tal que para cada vértice V , os triângulos abaixo comutam.

Para colocar em perspectiva a definição de limite, em particular o que a flecha mediadora significa, é interessante considerar toda a definição de solução, ou cone. Em outras palavras é interessante ter uma visão do todo, como ilustrada no diagrama abaixo:

Assim, um limite é uma solução particular tal que qualquer outra solução pode ser fatorada por ela. Olhando para o diagrama acima, talvez você perceba a similaridade da definição de produto cartesiano de conjuntos e do produto de categorias com este conceito. Dica: ignore a flecha $D(a)$. Caso não, elaboraremos sobre isto adiante.

Antes de passarmos aos exemplos resta, então, explicitar como o conceito de flecha universal e de limite estão relacionados. Considere então a categoria $\mathbf{C}^{\mathbf{G}}$ de G -diagramas em \mathbf{C} . Definimos o funtor diagonal

$$\Delta : \mathbf{C} \longrightarrow \mathbf{C}^{\mathbf{G}}$$

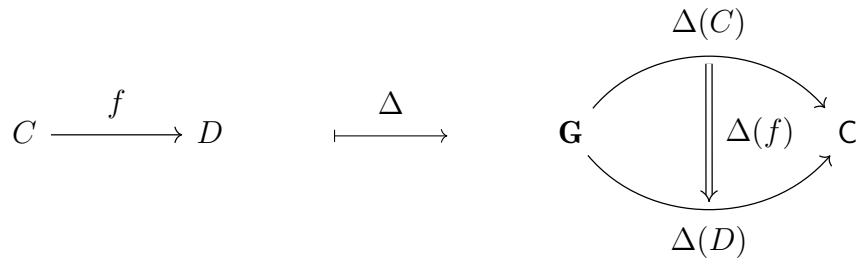
como aquele que envia cada objeto C no funtor constante $\Delta(C) : \mathbf{G} \rightarrow \mathbf{C}$ dado por:

$$\begin{aligned}
 V &\longmapsto \Delta(C)(V) = C \\
 a &\longmapsto \Delta(C)(a) = \text{id}_C
 \end{aligned}$$

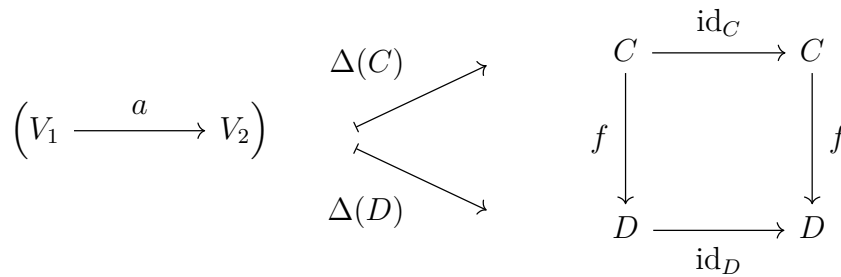
e que leva cada flecha $f : C \rightarrow D$ de \mathbf{C} na transformação natural $\Delta(f) : \Delta(C) \rightarrow \Delta(D)$ que é dada pela família

$$(f : C \rightarrow D)_{V \in \mathbf{G}}.$$

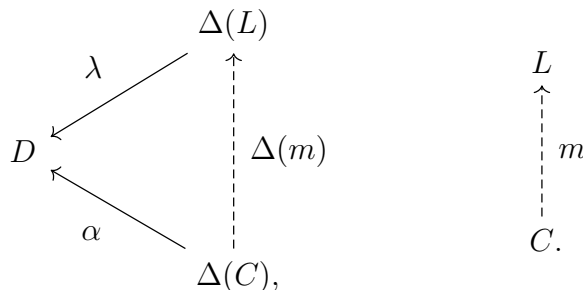
Diagramaticamente tem-se:



onde, para dois objetos V_1, V_2 de \mathbf{G} e uma flecha $a : V_1 \rightarrow V_2$, tem-se:



Note que os G -diagramas, $D : \mathbf{G} \rightarrow \mathbf{C}$, são objetos de $\mathbf{C}^{\mathbf{G}}$, assim, um **limite** é uma flecha universal de $\Delta : \mathbf{C} \rightarrow \mathbf{C}^{\mathbf{G}}$ a D . Note que estamos usando o conceito dual da definição 3.16, ou seja, temos um par $\langle L, \lambda : \Delta(L) \rightarrow D \rangle$, onde u é uma transformação natural, flecha de $\mathbf{C}^{\mathbf{G}}$, tal que para qualquer transformação natural $\alpha : \Delta(C) \rightarrow D$ de $\mathbf{C}^{\mathbf{G}}$, existe uma única flecha de \mathbf{C} , $m : C \rightarrow L$ tal que o triângulo abaixo comute.



Como $\Delta(L), \Delta(C), D : \mathbf{G} \rightarrow \mathbf{C}$ são funtores e $\Delta(m), \lambda, \alpha$ transformações naturais, podemos expandir o triângulo acima. Porém, antes disso, note que como $\Delta(C)$ para $C \in \mathbf{C}$ é o funtor constante, tem-se que uma transformação natural $\alpha : \Delta(C) \rightarrow D$ consiste em flechas $\alpha_i : C \rightarrow D(V_i)$ para todo objeto $V_i \in \mathbf{G}$, isto nos permite fazer escrever:

$$\begin{array}{ccc}
 \begin{array}{ccc}
 C & \xrightarrow{\alpha_1} & D(V_1) \\
 \text{id}_C \downarrow & & \downarrow D(a) \\
 C & \xrightarrow{\alpha_2} & D(V_2)
 \end{array} & \longrightarrow &
 \begin{array}{ccc}
 & & D(V_1) \\
 & \nearrow \alpha_1 & \downarrow D(a) \\
 C & & \\
 & \searrow \alpha_2 & \\
 & & D(V_2)
 \end{array}
 \end{array}$$

Ou seja, cada transformação natural de $\Delta(C) \rightarrow D$ para algum $C \in \mathcal{C}$ é um cone ou uma solução esquerda para o diagrama D . Deste modo, tem-se que a condição de um limite seja um flecha universal, $\langle L, u \rangle$, corresponde precisamente ao diagrama:

$$\begin{array}{ccccc}
 & & & & D(V_1) \\
 & & & \nearrow \tau_1 & \downarrow D(a) \\
 & & & \nearrow \lambda_1 & \\
 C & \xrightarrow{m} & L & & \\
 & & \searrow \lambda_2 & & \\
 & & \searrow \tau_2 & & D(V_2)
 \end{array}$$

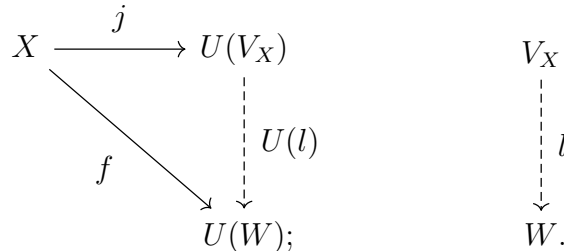
que indica que o limite $\langle L, \lambda \rangle$ é uma solução esquerda universal para o diagrama D , pois é um cone como o argumento acima mostrou e, além disto, existe uma única m , a flecha mediadora, para cada outro cone $\langle C, \alpha \rangle$ tal que o diagrama acima comute, já que $\langle L, \lambda \rangle$ é um limite. A situação é análoga para o colímite: ele consiste em uma flecha universal $\langle L, \lambda : D \rightarrow \Delta(L) \rangle$ de D a Δ o que transforma todos os cones na argumentação acima em co-cones.

Toda a discussão acima foi bastante abstrata e é difícil entender qual o objetivo de falar de limites de G -diagramas em uma categoria, especialmente se este for teu primeiro encontro com o assunto. Normalmente costuma-se introduzir alguns limites particulares primeiros e depois tratar da definição geral. Escolhemos o caminho contrário porque queríamos tratar do conceito de universalidade em geral primeiro para depois ver porque algumas construções, alguns limites, correspondem a construções universais. Agora, então, finalizamos este capítulo com três exemplos, um de universalidade apenas e dois sobre limites e a universalidade que acompanha.

Base de um espaço vetorial

Considere a categoria \mathbf{Vect}_K de espaços vetoriais sobre um corpo K e cujas flechas são mapas lineares entre espaços vetoriais. Existe um funtor esquecimento $U : \mathbf{Vect}_K \rightarrow \mathbf{Set}$ que envia cada espaço vetorial V em seu conjunto subjacente. Além disso, para cada

conjunto X , existe um espaço vetorial V_X dado pelo conjunto de todas as combinações K -lineares formais de elementos $x \in X$. A função que envia cada $x \in X$ no mesmo x visto como um elemento de V_x é uma flecha $j : X \rightarrow U(V_X)$. Além disso, sabe-se que para cada espaço vetorial W , toda função $f : X \rightarrow U(W)$ pode ser estendida para um único mapa linear $l : V_X \rightarrow W$. Em outras palavras tem-se



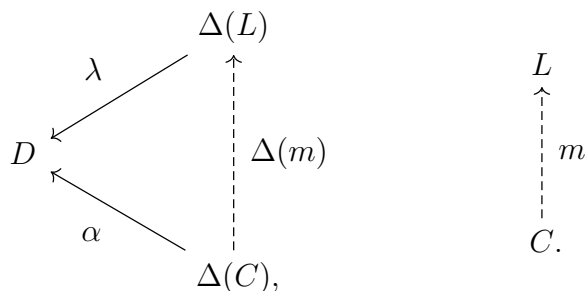
O resultado de que l existe e é único corresponde justamente a afirmar que $\langle V_X, j \rangle$ é uma flecha universal de X a U . A interpretação deste resultado é bastante simples. Por combinações K -lineares formais, queremos dizer que V_X é um espaço vetorial cuja base é X , isto é, o espaço vetorial gerado por X . Assim, este resultado corresponde ao fato bem conhecido de que para se especificar uma transformação linear, l , entre dois espaços vetoriais V_X e W , basta especificar a imagem da base de V_X sob l . Por outro lado, então, qualquer função $f : X \rightarrow U(W)$ determina uma transformação linear entre V_X e W , já que determina a imagem da base de V_X em W .

3.4.1 - Produto e coproduto

Considere a categoria \mathbf{P} que consiste apenas em dois objetos e suas flechas identidades. O grafo subjacente de \mathbf{P} é



Um P -diagrama, $D : \mathbf{P} \rightarrow \mathbf{C}$, em uma categoria \mathbf{C} consiste em dois elementos C_1 e C_2 de \mathbf{C} . Vejamos qual a forma de um limite de D . Como um limite é uma flecha universal do funtor diagonal $\Delta : \mathbf{C} \rightarrow \mathbf{C}^{\mathbf{P}}$ a um diagrama D , que é um objeto de $\mathbf{C}^{\mathbf{P}}$, então consideramos um par $\langle L, \lambda : \Delta(L) \rightarrow D \rangle$. Isto significa que para qualquer objeto C de \mathbf{C} e flecha $\alpha : \Delta(C) \rightarrow D$ em $\mathbf{C}^{\mathbf{P}}$ existe uma única flecha $m : C \rightarrow L$ tal que



Agora, considere a transformação natural α . Como D e $\Delta(C)$ são funtores de \mathbf{P} em \mathbf{C} , tem-se que, para os dois únicos elementos de \mathbf{P} esta transformação natural equivale a:

$$\begin{array}{ccc}
 \begin{array}{ccc}
 C & \xrightarrow{\alpha_1} & C_1 \\
 \text{id}_C \downarrow & & \\
 C & \xrightarrow{\alpha_2} & C_2
 \end{array} & \longrightarrow & \begin{array}{ccc}
 & & C_1 \\
 & \nearrow^{\alpha_1} & \\
 C & & \\
 & \searrow_{\alpha_2} & \\
 & & C_2
 \end{array}
 \end{array}$$

Onde C_1 e C_2 são a imagem do funtor D em \mathbf{C} . Note que isto é um cone ou uma solução esquerda. Os cones do digrama de formato P tem uma peculiaridade que é: como P é uma categoria com apenas dois objetos e nenhuma flecha entre eles, então existe flecha entre C_1 e C_2 em \mathbf{C} , na perspectiva do diagrama D , ao menos, para que o triangulo comute. Relembrando a definição de cone, vê-se que este é apenas um caso especial onde não existe nenhuma flecha não trivial na categoria que dá o formato do diagrama. Assim, não há condição de comutação nos cones de D .

Continuando com a universalidade, então, tem-se que um limite de D é um cone:

$$\begin{array}{ccc}
 & & C_1 \\
 & \nearrow^{\lambda_1} & \\
 L & & \\
 & \searrow_{\lambda_2} & \\
 & & C_2
 \end{array}$$

tal que, para todo outro cone $\langle C, \alpha \rangle$ de D , existe uma flecha $m : C \rightarrow L$, tal que o diagrama abaixo comute:

$$\begin{array}{ccccc}
 & & C & & \\
 & \swarrow^{\alpha_1} & | & \searrow_{\alpha_2} & \\
 & & m & & \\
 & & \downarrow & & \\
 C_1 & \xleftarrow{\lambda_1} & L & \xrightarrow{\lambda_2} & C_2.
 \end{array}$$

Repare a semelhança deste diagrama com aquele descrevendo o produto cartesiano de conjuntos ou o produto de categorias. O objeto limite, L , de D é chamado de **produto** e, seguindo a notação, escrito como $L = C_1 \times C_2$. As funções λ_1 e λ_2 são chamadas de projeções. Se considerarmos um P -diagrama em \mathbf{Set} , obtemos exatamente o produto cartesiano: trocando a notação um pouco, tem-se:

$$\begin{array}{ccccc}
 & & C & & \\
 & f \swarrow & | & \searrow g & \\
 A & \xleftarrow{\pi_1} & A \times B & \xrightarrow{\pi_2} & B.
 \end{array}$$

Fica estabelecido, então, em que sentido o produto cartesiano de conjuntos é universal. Ele corresponde a um limite para diagramas do formato \mathbf{P} em \mathbf{Set} . Do mesmo jeito, trocando a categoria \mathbf{C} , obtemos o conceito de produto de espaços topológicos, produto de grupos e assim por diante. Tais produtos geralmente são chamados produtos diretos. Lembramos, no entanto, que soluções esquerdas não necessariamente existem, então nem toda categoria possui produtos ou ainda, em uma dada categoria apenas certos pares de objetos podem possuir produtos.

Além da universalidade, afirmamos também que o produto cartesiano em \mathbf{Set} é natural e, como vimos no começo, de fato existe uma bijeção

$$[C, A \times B] \cong [\langle C, C \rangle, \langle A, B \rangle] = [C, A] \times [C, B].$$

A existência e naturalidade desta bijeção é garantida pelo seguinte lema que apresentamos sem prova [38]:

Proposição 3.6: Para um funtor $F : D \rightarrow C$, um par $\langle U, u : C \rightarrow F(U) \rangle$ é universal de C a F se, somente se a flecha que envia $f : U \rightarrow D$, para $D \in D$ na flecha $F(f) \circ u : C \rightarrow F(D)$, isto é, a flecha

$$(f : U \rightarrow D) \longrightarrow (F(f) \circ u : C \rightarrow F(D))$$

estabelece um isomorfismo

$$D[U, D] \cong C[C, F(D)].$$

Este isomorfismo é natural. Por outro lado, dados $U \in D$ e $C \in C$, qualquer isomorfismo natural deste tipo é determinado por uma única flecha $u : C \rightarrow F(U)$ tal que $\langle U, u \rangle$ é uma flecha universal de C a F .

Note que a proposição trata de flechas de C a F , porém limites são flechas de F a C . Assim, consideramos a proposição dual, aquela com as flechas invertidas. Como no caso de \mathbf{Set} , o produto é um par $\langle P, \pi \rangle$, o funtor é o funtor diagonal $\Delta : \mathbf{Set} \rightarrow \mathbf{Set}^{\mathbf{P}}$ e o objeto C é o P -diagrama, D . De modo que o resultado da proposição fornece um isomorfismo

natural entre:

$$[-, P] : \mathbf{Set} \rightarrow \mathbf{Set}$$

e

$$[\Delta(-), D] : \mathbf{Set} \rightarrow \mathbf{Set}.$$

Note que estes funtores são funtores hom em \mathbf{Set} . Além disso, como o P -diagrama D é um funtor $\mathbf{P} \rightarrow \mathbf{Set}$ que simplesmente seleciona dois objetos de \mathbf{Set} , escrevemos $D = \langle A, B \rangle$. Como P é o objeto limite de D , $P = A \times B$, de modo que o isomorfismo é

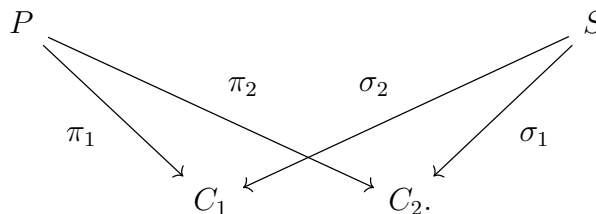
$$[-, A \times B] \cong [\Delta(-), \langle A, B \rangle]$$

ou ainda, para um dado objeto C :

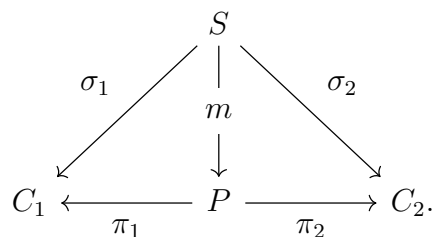
$$[C, A \times B] \cong [\Delta(C), \langle A, B \rangle] = [\langle C, C \rangle, \langle A, B \rangle] = [C, A] \times [C, B].$$

Lembre-se que como $[-, P]$ e $[\Delta(-), D]$ são funtores, eles também atuam sobre flechas em \mathbf{Set} . Nos restringimos a escrever apenas a ação destes funtores em objetos pois a notação, $[-, P]$, nos parece própria para tratar de objetos mas pode gerar confusões quando usada para flechas, apesar desta ser a notação usual para funtores hom.

Por fim, o último tópico sobre produtos é que eles são únicos a menos de isomorfismos. Para mostrar isto, considere dois produtos, P e S em uma categoria \mathbf{C} . Denotamos suas projeções por π_i e σ_i , $i = 1, 2$, respectivamente. Tem-se:



Note que o produto P corresponde a uma solução esquerda que é universal. A condição de universalidade garante que para qualquer outra solução esquerda, exista uma única flecha mediadora que faz o cone comutar. Em particular, tomando a outra solução esquerda como o outro produto, tem-se:



Este diagrama fornece as condições

$$\pi_2 \circ m = \sigma_2 \quad \text{e} \quad \pi_1 \circ m = \sigma_1.$$

Do mesmo modo, trocando P e S de lugar, tem-se uma flecha mediadora n tal que

$$\sigma_2 \circ n = \pi_2 \quad \text{e} \quad \sigma_1 \circ n = \pi_1.$$

Isto implica que

$$(\sigma_2 \circ n) \circ m = \sigma_2 \implies \sigma_2 \circ (n \circ m) = \sigma_2$$

$$(\sigma_1 \circ n) \circ m = \sigma_1 \implies \sigma_1 \circ (n \circ m) = \sigma_1.$$

Agora, note que a flecha $n \circ m$ é uma flecha de $S \rightarrow S$ que faz o diagrama de produto

$$\begin{array}{ccc} & S & \\ \sigma_1 \swarrow & | & \searrow \sigma_2 \\ C_1 & \xleftarrow{\sigma_1} S \xrightarrow{\sigma_2} & C_2 \\ & n \circ m \downarrow & \end{array}$$

comutar, ou seja, é uma flecha mediadora e é única. Como a flecha identidade id_S também faz este diagrama comutar, tem-se $n \circ m = \text{id}_S$. Um argumento análogo, substituindo agora o segundo par de equações na primeira, mostra que $m \circ n = \text{id}_P$. Logo as flechas m e n formam um par de flechas inversas e os dois produtos P e S são isomorfos.

Por este motivo então podemos falar d'O produto cartesiano, d'O produto direto de grupos, d'O produto de espaços topológicos e assim por diante. Note, por exemplo que para dois conjuntos A e B , existem dois produtos possíveis $B \times A$ e $A \times B$, no entanto existe um isomorfismo entre eles dado por $(b, a) \mapsto (a, b)$.

Por último existe o conceito dual ao de produto, chamado **coproduto**. Ele é obtido invertendo-se todas as flechas em diagramas de produto e possui propriedades similares ao produto no aspecto categórico. Isto é, a proposição 3.6 também se aplica, desta vez a própria, não sua dual, e o coproduto também é único a menos de isomorfismos. Ele é o colimite do P -diagrama D . Quando o produto é escrito $C_1 \times C_2$, o coproduto geralmente é denotado por $C_1 + C_2$, mas as vezes também denotam-se cada um por $C_1 \sqcap C_2$ e $C_1 \sqcup C_2$ respectivamente. Um diagrama para o coproduto $C_1 \sqcup C_2$ é:

$$\begin{array}{ccccc} & & C & & \\ & j_1 \nearrow & \uparrow m & \nwarrow j_2 & \\ C_1 & \xrightarrow{i_1} & C_1 \sqcup C_2 & \xleftarrow{i_2} & C_2 \end{array}$$

As flechas i_1 e i_2 que são as duais das projeções são chamadas de **inclusões canônicas** ou inserções ou injeções. Este último não é um bom nome pois evoca a ideia de que

correspondem a funções injetivas, o que não é necessariamente o caso. Um exemplo de coproduto é a união disjunta em **Set**. A união disjunta de dois conjuntos A e B é obtida marcando-se cada elemento segundo o conjunto a qual pertence. Definimos as inclusões $i_1 : A \rightarrow A \sqcup B$ e $i_2 : B \rightarrow A \sqcup B$ por

$$a \mapsto (a, 1) \quad \text{e} \quad b \mapsto (b, 2)$$

respectivamente. A união disjunta então é dada pela união das imagens destas inclusões, isto é

$$A \sqcup B = \{ (a, 1) \mid a \in A \} \cup \{ (b, 2) \mid b \in B \}.$$

No capítulo 2 encontramos a união disjunta durante a construção do fibrado na quantização de anyons, seção 2.3.1. Além de **Set**, outro exemplo de coproduto é a **soma direta** de grupos abelianos. Por fim, pode-se ter uma situação onde o mesmo objeto, P , é tanto o produto quanto o coproduto de um par de objetos, A e B , neste caso, P é chamado de **biproduto** e geralmente representado por $A \oplus B$. Nos depararemos com biprodutos na descrição de anyons.

3.4.2 - Equalizador e coequalizador

Equalizadores e coequalizadores são outro tipo de limite e colimite, respectivamente, em uma categoria \mathbf{C} . Assim, sua construção é bastante similar à de produtos e coprodutos, apenas alteramos o formato do diagrama em \mathbf{C} . Para equalizadores e coequalizadores, o diagrama é um funtor de $\mathbf{E} \rightarrow \mathbf{C}$ onde o grafo subjacente de \mathbf{E} é

$$\bullet \rightrightarrows \bullet$$

Deste modo, repetimos o processo que realizamos para o produto. Um E -diagrama é um funtor $D : \mathbf{E} \rightarrow \mathbf{C}$ e um **equalizador** é uma flecha universal $\langle E, \epsilon : \Delta(E) \rightarrow D \rangle$ do funtor diagonal $\Delta : \mathbf{C} \rightarrow \mathbf{C}^{\mathbf{E}}$ para o objeto D . Ou seja, para cada objeto $C \in \mathbf{C}$ e flecha $\alpha : \Delta(C) \rightarrow D$, existe uma única flecha $m : C \rightarrow E$ tal que o diagrama abaixo comute:

$$\begin{array}{ccc}
 & \Delta(E) & \\
 \epsilon \swarrow & \uparrow & \\
 D & \Delta(m) & E \\
 \alpha \swarrow & \uparrow & \uparrow \\
 & \Delta(C) & C
 \end{array}$$

Note que como o produto, a forma do diagrama do equalizador também só possui dois elementos, ou seja, um E -diagrama seleciona dois elementos da categoria \mathbf{C} . Diferentemente do produto, no entanto, agora existem flechas entre estes dois elementos. A transformação natural α do diagrama acima produz o seguinte cone do E -diagrama D :

$$\begin{array}{ccc}
 C & \xrightarrow{\alpha_1} & C_1 \\
 \text{id}_C \downarrow & & \downarrow g \\
 C & \xrightarrow{\alpha_2} & C_2 \\
 & & \downarrow f
 \end{array}
 \longrightarrow
 \begin{array}{ccc}
 & & C_1 \\
 \alpha_1 \nearrow & & \downarrow g \\
 C & & C_2 \\
 \alpha_2 \searrow & & \downarrow f
 \end{array}$$

Nomeamos as duas flechas entre os elementos C_1 e C_2 f e g por conveniência. Estas flechas são a imagem, sob D das duas flechas não triviais de \mathbf{E} . Note que a condição de comutação deste cone implica que

$$f \circ \alpha_1 = \alpha_2 = g \circ \alpha_1.$$

Dizemos que a flecha α_1 **torna igual** ou **equaliza** as flechas f e g . Deste modo, um cone de um E -diagrama geralmente é apresentado como um diagrama abaixo, já que a flecha α_2 é igual às duas compostas.

$$C \xrightarrow{\alpha} C_1 \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} C_2.$$

onde α é a flecha que equaliza o par f, g .

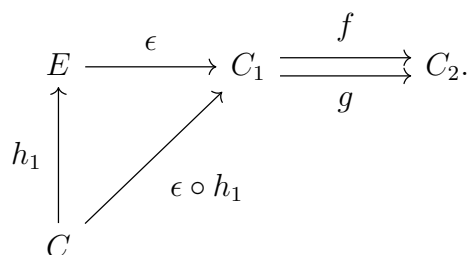
A condição de universalidade então diz que um equalizador é um par $\langle E, \epsilon \rangle$, onde ϵ é uma flecha que equaliza f e g , tal que, para cada par $\langle C, \alpha \rangle$ que equaliza f e g , existe uma flecha mediadora $m : C \rightarrow E$ que faz o diagrama abaixo comutar.

$$\begin{array}{ccc}
 E & \xrightarrow{\epsilon} & C_1 \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} C_2 \\
 \uparrow m & \nearrow \alpha & \\
 C & &
 \end{array}$$

Duas propriedades importantes de equalizadores são as seguintes: todo equalizador é uma flecha mônica e todo equalizador é único a menos de isomorfismos, tal como produtos e coprodutos. Para provar a primeira destas propriedades, primeiro note que a comutação do diagrama acima implica que

$$\epsilon \circ m = \alpha.$$

Assim, supomos que $\epsilon \circ h_1 = \epsilon \circ h_2$, para $h_1, h_2 : C \rightarrow E$. Como ϵ é um equalizador, tem-se, em particular, o seguinte diagrama:



que implica que

$$f \circ (\epsilon \circ h_1) = (f \circ \epsilon) \circ h_1 = (g \circ \epsilon) \circ h_1 = g \circ (\epsilon \circ h_1)$$

ou seja, $\epsilon \circ h_1$ equaliza o par f e g . Ou seja, existe uma única flecha $m : C \rightarrow E$ tal que

$$\epsilon \circ m = \epsilon \circ h_1.$$

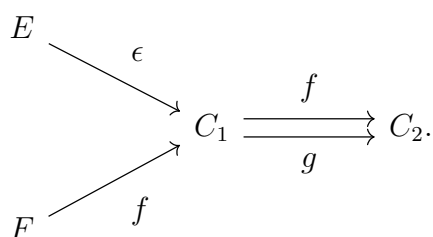
Pela unicidade de m , tem-se $m = h_1$. Do mesmo modo, $\epsilon \circ h_2$ equaliza f e g , logo temos uma única flecha $n : C \rightarrow E$ tal que

$$\epsilon \circ n = \epsilon \circ h_2$$

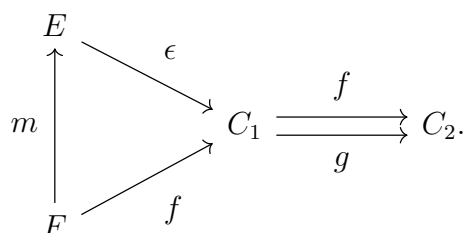
de modo que $h_1 = m = n = h_2$. Logo

$$\epsilon \circ h_1 = \epsilon \circ h_2 \implies h_1 = h_2.$$

Para demonstrar a segunda propriedade, considere um par de equalizadores, $\langle E, \epsilon \rangle$ e $\langle F, f \rangle$ das flechas $f, g : C_1 \rightarrow C_2$, como ilustrado no diagrama a seguir:



Pela definição, para qualquer flecha $\alpha : C \rightarrow C_1$ que equaliza o par f, g , existem flechas $m : C \rightarrow E$ e $n : C \rightarrow F$, únicas, tais que $\epsilon \circ m = \alpha$ e $f \circ n = \alpha$. Como ambas ϵ e f equalizam f e g , em particular existem $m : F \rightarrow E$ tal que o diagrama abaixo comute.

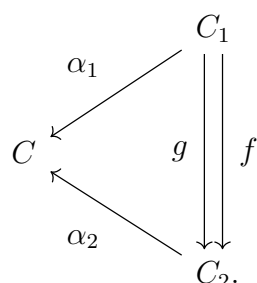


Ou seja, tem-se que $\epsilon \circ m = f$. Do mesmo modo, existe flecha $n : E \rightarrow F$ tal que $f \circ n = \epsilon$. Substituindo-se a última equação na primeira, obtém-se:

$$(f \circ n) \circ m = f \circ (n \circ m) = f.$$

Como f é uma flecha de um equalizador, tem-se que $n \circ m = \text{id}_F$ pois já que f equaliza f e g , existe uma flecha $m' : F \rightarrow F$, única, tal que $f \circ m' = f$. Como a identidade em F satisfaz esta relação, tem-se $\text{id}_F = m' = n \circ m$. Do mesmo modo obtém-se que $m \circ n = \text{id}_E$, logo m e n são um par de flechas inversas e F e E são isomorfos.

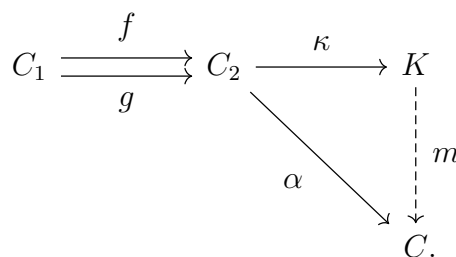
A noção dual de um equalizador é um coequalizador, uma flecha universal $\langle K, \kappa : D \rightarrow \Delta(K) \rangle$ do E -diagrama D para o funtor diagonal Δ . Um cocone de D está representado no diagrama que comuta a seguir.



Assim como no caso do equalizador, simplificamos este diagrama como

$$C_1 \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} C_2 \xrightarrow{\alpha} C.$$

Novamente, dizemos que α torna igual o par f, g . Também diremos que α coequaliza o par f, g . Um coequalizador então é o par $\langle K, \kappa \rangle$ tal que para cada flecha $\alpha : C_2 \rightarrow C$ que coequaliza o par f, g , existe uma única flecha $m : K \rightarrow C$ tal que o diagrama abaixo comute.



Sendo o dual do equalizador, o coequalizador também possui propriedades similares. Em particular, assim como o equalizador ele é único a menos de isomorfismos e toda flecha de um coequalizador é uma flecha épica.

Por fim, vejamos dois exemplos, o equalizador e o coequalizador em **Set**. O equalizador em **Set** é bastante simples. Sejam $f, g : A \rightarrow B$ funções e S o conjunto dado por

$$S = \{ a \in A \mid f(a) = g(a) \}.$$

Isto é, o conjunto de elementos nos quais f e g concordam. A inserção

$$S \xhookrightarrow{i} A$$

é o equalizador de f, g . Para ver isto, suponha que $h : X \rightarrow A$ equalize f e g . Neste caso, tem-se, para todo $x \in X$:

$$f(h(x)) = (f \circ h)(x) = (g \circ h)(x) = g(h(x))$$

de modo que $h(x) \in S$ para todo $x \in X$. Assim, a função $m : X \rightarrow S$ dada por $m(x) = h(x)$ é a flecha mediadora requerida. Diagramaticamente, tem-se:

$$\begin{array}{ccc} S & \xhookrightarrow{i} & A \xrightarrow{f} B \\ & \nearrow h & \xrightarrow{g} \\ X & & \end{array}$$

m (vertical dashed arrow from X to S)

O coequalizador em **Set** é um pouco mais complicado. Para isto precisamos da seguinte construção preliminar. Seja S um conjunto arbitrário e \sim uma relação de equivalência em S . Esta relação particiona em S em suas classes de equivalência $[s]$. Representamos o conjunto de todas tais classes por S/\sim e a sobrejeção canônica $s \mapsto [s]$ por $\sigma : S \rightarrow S/\sim$. Agora, considere uma função

$$S \xrightarrow{h} X.$$

O **núcleo** de h é a relação \approx em S dada por

$$s_1 \approx s_2 \iff h(s_1) = h(s_2)$$

para $s_1, s_2 \in S$. Esta relação é, trivialmente, uma relação de equivalência em S .

Agora, suponha que \approx inclua \sim , isto é:

$$s_1 \sim s_2 \implies h(s_1) = h(s_2)$$

para $s_1, s_2 \in S$. Sob estas condições existe um triângulo comutativo

$$\begin{array}{ccc} S & \xrightarrow{\sigma} & S/\sim \\ & \searrow h & \downarrow h^\# \\ & & X \end{array}$$

para alguma função única $h^\#$. Esta função é dada por

$$h^\#([s]) = s$$

para algum $s \in S$. Não provaremos aqui a existência e unicidade de $h^\#$ no entanto.

Prosseguindo, para formar um equalizador, então, usamos a construção acima. Sejam $f, g : A \rightarrow B$ um par de funções e \rightsquigarrow uma relação em B dada por

$$b_1 \rightsquigarrow b_2 \iff \exists a \in A. b_1 = f(a) \wedge b_2 = g(a)$$

para $b_1, b_2 \in B$ e seja \sim a relação de equivalência em B gerada por \rightsquigarrow . Esta relação não é simples de se descrever, porém isto não importa. Usando a construção anterior, tem-se que a sobrejeção canônica

$$B \xrightarrow{\sigma} B/\sim$$

é o coequalizador de f, g , pois o diagrama abaixo comuta.

$$\begin{array}{ccccc} A & \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} & B & \xrightarrow{\sigma} & B/\sim \\ & & & \searrow h & \downarrow h^\# \\ & & & & X \end{array}$$

Note que a condição de que o núcleo de h contenha \sim garante, exatamente que h coequaliza f e g .

Capítulo 4

Sobre anyons e categorias

O capítulo anterior tratou exclusivamente de categorias, sem nenhuma menção a anyons. Neste capítulo, finalmente, juntaremos as duas coisas para tratar de anyons usando categorias. Antes de começarmos, dois comentários são necessários. Primeiro, escolhemos fazer o capítulo anterior como está porque achamos necessário incluir exemplos para auxiliar no entendimento e tentar situar onde teoria de categorias está em relação ao resto da matemática. Fizemos isto porque usualmente, estudantes de física, para quem este trabalho é voltado, não costumam ter contato com teoria de categorias em geral, especialmente dado que este trabalho visa ser uma introdução ao assunto de anyons e sua descrição via categorias. Deste modo, tivéssemos apenas colocado as definições e pouca ou nenhuma discussão, sentiríamos que estaríamos contrariando os objetivos do trabalho.

Segundo, o capítulo anterior não é de nenhuma maneira exaustivo sobre teoria de categorias apesar de cobrir quase todos os conceitos principais da teoria básica. Em particular isto significa que ele não fornece prontamente todos os conceitos necessários para a descrição de anyons. Assim, neste capítulo introduziremos mais conceitos categoriais, porém o faremos de modo subordinado à descrição de anyons. Pode-se pensar a discussão que será feita aqui como análoga à apresentação de espaços de Hilbert em cursos de mecânica quântica. A definição formal de espaços de Hilbert, em particular o espaço das funções quadrado integráveis requer uma quantidade razoável de conhecimento usualmente além do esperado de estudantes no primeiro curso de mecânica quântica. No entanto, apenas sabendo o que são espaços vetoriais e a noção de produto interno é possível ter uma ideia bastante boa e suficiente para o que se segue. A situação é análoga aqui. Todos os conceitos apresentados podem ser obtidos a partir daqueles no capítulo anterior, porém, agora, torna-se mais vantajoso apresentá-los concorrentemente ao seu uso na descrição de anyons.

Antes de começarmos, discutamos um pouco sobre o comportamento de anyons e o que é necessário para formalizá-lo matematicamente. A primeira coisa a se notar é que existem diferentes tipos de anyons, isto é, anyons com diferentes parâmetros estatísticos ou **cargas**, como diremos neste capítulo. Assim, uma descrição de anyons deve ser capaz de acomodar diversos objetos distintos. Devido à possibilidade de fusão, um sistema anyônico dificilmente conterá apenas um tipo de anyon. Como a palavra objeto sugere,

categorias são adequadas para isto. Além disto, o trançamento de anyons é parte crucial da física destas partículas e, como vimos, apesar de corresponderem a loops no espaço de configurações, o resultado físico não necessariamente permanece inalterado. Isto sugere que precisamos de uma noção de igualdade “mais fraca”. Em particular, é conveniente falar de objetos isomórficos mas que não são necessariamente iguais, outro ponto forte de teoria de categorias.

Além disto, uma descrição de anyons deve capturar tanto o aspecto cinemático, isto é, o trançamento de anyons, quanto o aspecto dinâmico, que requer conectar o primeiro com as regras usuais da mecânica quântica. Formalizar o aspecto cinemático requer uma estrutura algébrica capaz de capturar as cargas de anyons, suas regras de fusão e suas trajetórias em $2 + 1$ dimensões. No capítulo 2, desenvolvemos o assunto até chegar ao grupo de tranças para descrever estas partículas. No entanto, se olharmos com atenção, veremos que o grupo de tranças corresponde a uma descrição parcial devido à existência de regras de fusão.

Considerando o exemplo que usamos para introduzir a fusão de anyons, lembre-se que naquele caso em particular, trançar um anyon de estatística ou carga θ ao redor de dois outros de mesma carga corresponde a trançar o primeiro ao redor de um anyon de carga 4θ e então **rotacionar** o anyon de carga 4θ por um ângulo de π . O fato de podermos falar em rotacionar um anyon ao redor de si mesmo sugere que ele é um objeto extenso e, deste modo, o grupo de tranças como está não é suficiente, pois ele trata apenas de tranças em fios que correspondem às linhas de mundo das partículas. Para capturar a rotação devemos falar de “fitas de mundo”, basicamente extendendo as linhas de mundo de modo que possamos torce-las para capturar o grau de liberdade extra.

Pode parecer estranho considerar que todos os anyons são objetos extensos a partir deste argumento, mas lembre-se também que o mundo em que vivemos é tridimensional. Isto significa que, até onde sabemos, anyons só podem ser obtidos como partículas emergentes em sistemas de muitas partículas. Em outras palavras, só podem ser obtidos como modos coletivos, de modo que é bastante razoável considerar que são objetos extensos. O argumento fica mais plausível quando levamos em conta a natureza topológica destas partículas. As propriedades de um dado conjunto de anyons dependem das propriedades topológicas **globais** do sistema em questão.

Dividiremos este capítulo em diversas seções menores, cada uma abordando um aspecto da descrição formal de anyons e ao fim trataremos de um modelo anyônico específico como exemplo. Este capítulo é fundamentado principalmente nos trabalhos de Panangaden e Paquette [40] e Blass e Gurevich [42]. Curiosamente a ordem em que estes trabalhos construíram a descrição categoria de anyons é inversa. Nos ateremos mais à abordagem de Blass e Gurevich. Panangaden e Paquette, no entanto fornecem um resumo muito útil dos componentes da descrição categorial de anyons.

A descrição categoria de anyons requer:

- Um sistema de rótulos ou tipos para capturar a noção de carga anyônica.
- Um modo de expressar a composição de anyons, isto é, o anyon obtido através de uma fusão. Isto induz uma noção de produto na categoria pois obtemos um novo objeto a partir de outros. A categoria resultante possui uma estrutura dita **monoidal**.
- Uma estrutura que capture as linhas de mundo de um sistema de anyons, representadas pelo grupo de tranças. Assim, nossa categoria monoidal deve ter uma **estrutura de tranças**.
- Uma noção de carga conjugada, isto é, uma carga A^* tal que a fusão de um anyon de carga A com um de carga A^* resulte na carga trivial, $\mathbf{1}$. Isto induz uma estrutura que chamamos de **estrutura rígida** na categoria.
- Uma noção capaz de capturar o fato de anyons são objetos estendidos, como discutido é linhas de mundo não são suficientes para capturar seus movimentos, é necessário usar fitas. A descrição algébrica disto foi dada muito antes de matemáticos estarem cientes de anyons e fornece uma **estrutura de fita** para a categoria.
- Uma estrutura que descreva as regras de fusão e que insira todo este formalismo algébrico no contexto de espaços de Hilbert. Isto é obtido com uma **estrutura semissimples**.

Antes de prosseguirmos um comentário sobre a ideia de carga conjugada. Tratar a estatística de um anyon como uma espécie de carga é uma ideia que surge no tratamento via teoria topológica de campos [15]. A origem da relação de estatística com o conceito de carga remonta à opção que não seguimos no processo de quantização de anyons — escolher não anular o potencial de calibre e o tensor \mathcal{F} . Ao usar a analogia entre estes dois objetos com o potencial vetor e o campo eletromagnético, é possível obter-se uma descrição de anyons como compósitos de partículas e tubos de fluxo magnético. Esta descrição é, mais propriamente, uma descrição sobre como obter anyons fisicamente do que uma descrição sobre o conceito de anyons e é devida a Wilczek [4].

Nesta abordagem de Wilczek e posteriormente via teoria topológica de campos relaciona-se a estatística da partícula com o quanta de fluxo magnético envolvido pelo tubo de fluxo que, por sua vez, depende da carga elétrica da partícula que forma o compósito com o tubo de fluxo. Desta relação, então, surge uma ideia de conservação da carga topológica de anyons já que ela está relacionada, em última instância, à uma carga elétrica. Esta discussão não é trivial e passa por questões como fracionalização da carga elétrica, por exemplo. De todo modo, a apresentamos superficialmente aqui apenas para contextualizar a ideia de carga conjugada, até agora ausente.

Por fim, para entender a necessidade destas estruturas precisamos revisitar as regras de fusão de anyons e estudá-las com mais detalhes. Partimos das informações definidoras que fornecemos no capítulo 2. Dado um sistema com partículas anyônicas, podemos fundir duas destas partículas de modo a obter uma partícula cujo comportamento estatístico é, de alguma forma, derivado do comportamento estatístico das duas partículas originais. No caso de anyons abelianos, a fusão acarreta simplesmente em um novo anyon que possui estatística obtida de um modo simples a partir da estatística dos anyons originais. Por exemplo, a fusão de dois anyons θ gera um anyon de estatística igual a 4θ . Em geral [15], para um modelo com anyons abelianos de estatística θ é:

$$n^2\theta \otimes k^2\theta = (n+k)^2\theta$$

com a limitação que $n^2\theta, k^2\theta, (n+k)^2\theta \in [0, 2\pi]$ e $n, k \in \mathbb{Z}$. Assim, $n = 1$ corresponde a um único anyon, $n = 2$ ao anyon obtido pela fusão de dois de estatística θ e assim por diante.

O caso de anyons não abelianos é mais complexo. Além disto, a literatura sobre este tipo de anyons tratados em termos de mecânica quântica básica aparenta ser escassa, beirando ao inexistente, apesar de ser, em princípio, possível como segure Lerda [26]. Por isto, a discussão a seguir deve ser entendida mais como uma análise heurística para entender o conceito do que uma descrição própria e formal. Isto se dá, inclusive, pela questão de que anyons só podem existir em nosso mundo tridimensional como partículas emergentes. Usar mecânica quântica básica para estudá-los é um modo útil e relativamente simples de entender as propriedades fundamentais deste tipo de partículas mas é inadequado se pretendemos falar de qualquer aplicação real.

Partimos então da discussão final do capítulo 2. Segundo Nayak *et al.* [15], anyons não abelianos são aqueles onde uma permutação transforma a função de onda segundo uma representação não unidimensional, isto é representações de dimensão maior ou igual a dois, do grupo de tranças. Isto está de acordo com a afirmação de Lerda [26] de que: ao se considerar funções de onda multidimensionais, é possível obter anyons que se transformam segundo representações de dimensões superiores do grupo de tranças quando permutados. O diferente sobre estes anyons, segundo Blass e Gurevich [42] é que os elementos do espaço de Hilbert que descreve este sistema não representam configurações em si, mas sim modos distintos em que pode-se fundir anyons de modo a obter uma dada configuração.

Podemos começar a entender isto através do seguinte raciocínio. Suponha que no processo de quantização que realizemos na seção 2.3.1, os espaços \mathfrak{h}_c tivessem dimensão maior que um. Neste caso à cada configuração c corresponderia um espaço vetorial de dimensão $n > 1$ e seriam todas degeneradas. Supomos então, em alinhamento com Lerda, que o procedimento de quantização restante resultaria em funções de onda que evoluem segundo uma representação n -dimensional do grupo de tranças. Repare, no entanto que

cada subespaço corresponde a uma única configuração c . Em analogia com a representação da posição, seria como se a cada autoket $|x\rangle$ correspondesse um subespaço degenerado. Mas, se todos os vetores neste espaço representam uma partícula na mesma configuração c — em analogia $|c\rangle$ —, então o que representam os graus de liberdade deste espaço?

Seguindo a afirmação de Blass e Gurevich, devemos concluir que estão relacionados à fusão de anyons. Exceto por permutações, não podemos mover os anyons, pois passaríamos de uma configuração a outra. Lembre-se: uma configuração c guarda consigo a posição relativa dos anyons, se esta posição mudar por meio de uma transformação que não é uma permutação passamos a outra configuração e, portanto, outro subespaço. Assim, uma possibilidade é encarar estes graus de liberdade como modos distintos de caracterizar este conjunto de anyons, em outras palavras, modos distintos de considerar a fusão. Para tornar as coisas um pouco mais palpáveis, fornecemos um exemplo.

Anyons de Fibonacci, são um modelo particular de anyons que estudaremos formalmente ao fim deste capítulo. O modelo consiste em um anyon não abeliano de tipo \mathcal{T} e o anyon de estatística trivial, $\mathbf{1}$. As regras de fusão deste modelo são:

$$\begin{aligned}\mathcal{T} \otimes \mathcal{T} &= \mathcal{T} \oplus \mathbf{1}; \\ \mathcal{T} \otimes \mathbf{1} &= \mathbf{1} \otimes \mathcal{T} = \mathcal{T}; \\ \mathbf{1} \otimes \mathbf{1} &= \mathbf{1}.\end{aligned}$$

Considere então um sistema de três anyons, τ do tipo \mathcal{T} . Como discutido, anteriormente, a fusão destes anyons não significa a formação de uma nova partícula a partir das antigas, em particular, o sistema sempre terá os três anyons τ . No entanto, as regras de fusão nos dizem que um sistema de três anyons do tipo \mathcal{T} é dado por

$$\mathcal{T} \otimes \mathcal{T} \otimes \mathcal{T} = \mathcal{T} \oplus (\mathcal{T} \otimes \mathbf{1}) = (\mathcal{T} \otimes \mathcal{T}) \oplus (\mathcal{T} \otimes \mathbf{1}) = \mathcal{T} \oplus \mathbf{1} + \mathcal{T} = 2\mathcal{T} \oplus \mathbf{1}.$$

No fim, este sistema corresponde a aquele com um único anyon de tipo \mathcal{T} ou de tipo $\mathbf{1}$. No entanto existem dois modos distintos de ver a fusão deste sistema como um anyon de tipo \mathcal{T} . O primeiro processo de fusão pode ter resultado em $\mathbf{1}$ ou o primeiro pode ter resultado em \mathcal{T} e o segundo em \mathcal{T} . É a este espaço de possibilidades que corresponde o espaço de Hilbert de anyons não abelianos.

Note que a fusão de anyons não abelianos possui semelhanças formais com a combinação de momento angular de duas partículas. Neste sentido a estatística ou tipo ou carga dos anyons pode ser pensado como um número quântico similar ao spin, por exemplo. Quando duas partículas são combinadas, o momento angular da partícula resultante é transformado segundo o produto tensorial das representações de $SU(2)$. Este produto não necessariamente é irredutível, e é necessário, então, decompô-lo em uma soma de representações irredutíveis. Assim, se combinarmos uma partícula de spin j e outra de spin

k , obtém-se um sistema que pode estar em um estado de spin qualquer entre $|j - k|$ e $j + k$. Por exemplo, “fundir” duas partículas de spin meio pode gerar uma partícula de spin zero ou uma partícula de spin um. Formalmente poderíamos escrever o seguinte:

$$\mathcal{J} \otimes \mathcal{K} = \mathcal{J} - \mathcal{K} \oplus \dots \oplus \mathcal{J} + \mathcal{K}.$$

Entenda a palavra formalmente aqui como simbolicamente e não “de modo rigoroso”. Na expressão acima, $\mathcal{J} \otimes \mathcal{K}$ é o operador spin para a nova partícula composta e a operação \oplus simplesmente separa as possibilidades. A situação é similar para anyons não abelianos. Se \mathcal{A} e \mathcal{B} são tipos de anyons, tem-se

$$\mathcal{A} \otimes \mathcal{B} = \mathcal{A} - \mathcal{B} \oplus \dots \oplus \mathcal{A} + \mathcal{B}.$$

onde $\mathcal{A} \pm \mathcal{B}$ e os fatores intermediários são anyons de um novo tipo. No caso do momento angular, as operações \otimes e \oplus correspondem ao produto tensorial e soma direta de representações. No caso de anyons, precisamos ainda formalizar estas operações de modo que capturem todas as propriedades relevantes de anyons. Isto é o que faremos usando categorias. Começaremos pela operação \otimes .

4.1 – Monoidalidade

A palavra monoidalidade é um neologismo que usaremos para nos referir a algo que possui propriedades monoidais, isto é propriedades derivadas daquelas de um monoide. Começamos a definir precisamente o que isto significa. No início do capítulo 3, apresentamos uma definição categorial de grupos e de monoides. Em termos de teoria de categorias, a vantagem de definições deste tipo é que podemos tratar de grupos, monoides e outras estruturas algébricas grandes, isto é, estruturas cuja coleção subjacente de elementos não forma um conjunto. No nosso caso não precisamos nos preocupar muito com isto pois uma das condições que imporemos sobre as categorias representando modelos anyônicos é que possuam um número finito de objetos sendo, conseqüentemente, pequenas o pelo menos essencialmente pequenas — equivalentes a uma categoria pequena. Categorias ditas **monoidais** surgiram para estudar monoides em categorias. Vejamos a definição:

Definição 4.1: Uma **categoria monoidal** \mathbf{M} , formalmente $\langle \mathbf{M}, \otimes, \mathbf{I}, \alpha, \lambda, \rho \rangle$ é uma categoria \mathbf{M} equipada com um bifuntor $\otimes : \mathbf{B} \times \mathbf{B} \rightarrow \mathbf{B}$, um objeto $\mathbf{I} \in \mathbf{B}$, um isomorfismo natural α dado por

$$\alpha_{A,B,C} : A \otimes (B \otimes C) \cong (A \otimes B) \otimes C$$

tal que o diagrama

$$\begin{array}{ccc}
 & A \otimes (B \otimes (C \otimes D)) & \\
 \text{id}_A \otimes \alpha_{B,C,D} \swarrow & & \searrow \alpha_{A,B,C \otimes D} \\
 A \otimes ((B \otimes C) \otimes D) & & (A \otimes B) \otimes (C \otimes D) \\
 \alpha_{A,B \otimes C,D} \downarrow & & \downarrow \alpha_{A \otimes B,C,D} \\
 (A \otimes (B \otimes C)) \otimes D & \xrightarrow{\alpha_{A,B,C} \otimes \text{id}_D} & ((A \otimes B) \otimes C) \otimes D
 \end{array}$$

comuta para todos os objetos $A, B, C, D \in \mathbf{M}$ e mais dois isomorfismos naturais, λ e ρ dados por

$$\lambda_A : \mathbf{I} \otimes A \cong A \quad \text{e} \quad \rho_A : A \otimes \mathbf{I} \cong A$$

tais que o diagrama triangular abaixo comute para todos os objetos $A, B \in \mathbf{M}$:

$$\begin{array}{ccc}
 A \otimes (\mathbf{1} \otimes B) & \xrightarrow{\alpha_{A,\mathbf{1},B}} & (A \otimes \mathbf{1}) \otimes B \\
 \text{id}_A \otimes \lambda_B \searrow & & \swarrow \rho_A \otimes \text{id}_B \\
 & A \otimes B &
 \end{array}$$

Esta é uma definição um tanto extensa mas não é complicada. Analisemos uma parte de cada vez. Primeiro, notamos o bifuntor \otimes . Salientamos que o objeto $A \otimes B$ **não** é um produto ou coproduto dos objetos A e B no sentido categorial. Meramente escolhemos algum objeto de \mathbf{M} e para ser a imagem do par $\langle A, B \rangle$ sob \otimes . Para que $A \otimes B$ fosse um produto ou coproduto seria necessário que existissem flechas projeção ou inclusão e que satisfizessem a propriedade de universalidade. Não exigimos nada disto aqui. Salientamos este ponto apenas porque até agora, todos os bifuntores apresentados estavam relacionados a um produto ou coproduto em categorias, como o produto cartesiano ou a união disjunta. O símbolo \times na definição do bifuntor representa a construção da categoria produto.

O isomorfismo natural α traduz a associatividade da operação \otimes . Ao invés de exigirmos que os objetos $A \otimes (B \otimes C)$ e $(A \otimes B) \otimes C$ sejam iguais, exigimos apenas que sejam isomorfos. Se trocássemos todos os isomorfismos por igualdades estritas, teríamos uma categoria monoidal **estrita**. O primeiro diagrama, chamado de **identidade pentagonal** ou **diagrama pentagonal** apenas estabelece as propriedades que se espera de uma operação associativa. Como $A \otimes (B \otimes C) \cong (A \otimes B) \otimes C$, esperamos que valha

$(A \otimes (B \otimes C)) \otimes D \cong ((A \otimes B) \otimes C) \otimes D$ e similarmente para um objeto fora dos parênteses à esquerda. As flechas $\alpha_{A,B,C} \otimes \text{id}_D$ e $\text{id}_A \otimes \alpha_{B,C,D}$ são obtidas pela functorialidade de \otimes .

Por fim, o objeto \mathbf{I} faz o papel de identidade para a operação \otimes e assim como a associatividade, exigimos que $\mathbf{I} \otimes A$ e $A \otimes \mathbf{I}$ sejam apenas isomorfos a A , donde surgem os isomorfismos naturais λ e ρ . O diagrama triangular então é meramente a compatibilidade da identidade com a associatividade de \otimes : se $A \otimes \mathbf{1} \cong A$ esperamos que $(A \otimes \mathbf{1}) \otimes B \cong A \otimes B$ e similarmente para a identidade à esquerda.

De certo modo, assumimos que um monoide é a estrutura algébrica mais básica que nos interessa. Grupos são monoides onde cada elemento possui um inverso com respeito à operação de monoide, grupos abelianos são grupos cuja operação é comutativa, anéis são grupos abelianos em uma operação e monoides em outra e ambas as operações são “compatíveis” — distributividade — e assim por diante. Enfraquecer qualquer uma das propriedades de monoide não nos interessa. Sendo assim, categoriais monoidais são a base para obtermos categorias cujas coleções de flechas possuem estrutura. Fazemos isto pelo processo de **enriquecimento**. Uma categoria, \mathbf{C} , é \mathbf{M} -enriquecida quando para quaisquer objetos A e B de \mathbf{C} , a coleção de flechas $[A, B]$ é um objeto de \mathbf{M} . A operação \otimes então fornece uma composição adequada de morfismos.

A estrutura extra das coleções $[A, B]$ existe se os objetos da categoria \mathbf{M} -enriquecida, \mathbf{C} , possuem esta estrutura. Por exemplo, se forem monoides. Para elucidar isto, repetimos aqui a definição categorial de monoide usando a operação \otimes em \mathbf{M} . Um objeto M é um monoide em \mathbf{M} se existem flechas $\star : M \otimes M \rightarrow M$ e $\eta : \mathbf{1} \rightarrow M$ tais que

$$\begin{array}{ccc} (M \otimes M) \otimes M & \xleftarrow{\alpha_{M,M,M}} & M \otimes (M \otimes M) \xrightarrow{\text{id}_M \otimes \star} M \otimes M \\ \star \otimes M \downarrow & & \downarrow \star \\ M \otimes M & \xrightarrow{\quad \star \quad} & M; \end{array}$$

e

$$\begin{array}{ccccc} \mathbf{I} \otimes M & \xrightarrow{\eta \otimes \text{id}_M} & M \otimes M & \xleftarrow{\text{id}_M \otimes \eta} & M \otimes \mathbf{I} \\ & \searrow \lambda_M & \downarrow \star & \swarrow \rho_M & \\ & & M & & \end{array}$$

comutam. A partir desta estrutura para um objeto M , adicionando mais condições podemos definir M como um grupo, um anel, um espaço vetorial e assim por diante.

Bem, como dissemos, o bifunctor \otimes será a formalização da operação \otimes para anyons, isto é, da fusão de anyons. As propriedades exibidas por \otimes em uma categoria monoidal são

exatamente aquelas que esperamos da fusão de anyons. Em particular a ordenação dos parênteses é de grande importância. Para anyons não abelianos, a ordem de fusão importa porque, enquanto as representações unidimensionais e unitárias do grupo de tranças são necessariamente abelianas, as representações de ordem superior não necessariamente o são. Assim, a fusão $\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$ não necessariamente é igual a $(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}$ para anyons de tipo \mathcal{A} , \mathcal{B} e \mathcal{C} . Mas ainda esperamos que estes objetos estejam relacionados de alguma forma, donde entra o isomorfismo natural. Em particular veremos que esta relação é obtida a partir de trançamento dos anyons. Denotando por \mathfrak{A} — esta é a letra A no estilo fraktur — a categoria de um sistema de anyons arbitrário, enunciamos o primeiro axioma:

Axioma 1: A categoria \mathfrak{A} é uma categoria monoidal.

Ainda tratando da monoidalidade de \mathfrak{A} , podemos esperar que a relação que exista entre $\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$ e $(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}$ seja obtida a partir de trancamentos, em particular, que o isomorfismo $\alpha_{A,B,C}$ corresponda a algum trancamento de anyons. Se \otimes é a fusão de anyons então podemos supor, também que seja comutativo. Isto é, que exista um isomorfismo natural

$$\mathcal{A} \otimes \mathcal{B} \cong \mathcal{B} \otimes \mathcal{A}.$$

Em particular este isomorfismo corresponderia precisamente a um gerador do grupo de tranças, aquele que troca o anyon de tipo \mathcal{A} com o de tipo \mathcal{B} . A existência deste isomorfismo transformaria \mathfrak{A} exatamente no que é chamado de **categoria modular trançada**.

Definição 4.2: Uma categoria monoidal \mathfrak{T} é dita **trançada** se existe um isomorfismo natural σ dado por

$$\sigma_{A,B} : A \otimes B \cong B \otimes A$$

tal que o diagrama

$$\begin{array}{ccccc}
 & & A \otimes (B \otimes C) & \xrightarrow{\sigma_{A,B \otimes C}} & (B \otimes C) \otimes A & & \\
 & \nearrow^{\alpha_{A,B,C}^{-1}} & & & & \searrow^{\alpha_{B,C,A}^{-1}} & \\
 (A \otimes B) \otimes C & & & & & & B \otimes (C \otimes A) \\
 & \searrow_{\sigma_{A,B} \otimes \text{id}_C} & & & & \nearrow_{\text{id}_B \otimes \sigma_{A,C}} & \\
 & & (B \otimes A) \otimes C & \xrightarrow{\alpha_{B,A,C}^{-1}} & B \otimes (A \otimes C) & &
 \end{array}$$

e o mesmo diagrama trocando-se σ por σ^{-1} comutam para todo $A, B, C \in \mathbb{T}$ e

$$\begin{array}{ccc}
 A \otimes \mathbf{1} & \xrightarrow{\rho_A} & A \\
 \sigma_{A,1} \downarrow & & \nearrow \lambda_A \\
 \mathbf{1} \otimes A & &
 \end{array}$$

comute para todo A em \mathbb{T} .

O segundo diagrama, chamado de **identidade hexagonal** ou **diagrama hexagonal** corresponde a uma espécie de compatibilidade entre os isomorfismos σ e α . Como σ corresponde à comutatividade, este diagrama nos diz que se $A \otimes B \cong B \otimes A$ então devemos ter $(A \otimes B) \otimes C \cong B \otimes (C \otimes A) \cong (B \otimes C) \otimes A$ e assim por diante. Em outras palavras, trocar os elementos de lugar com diversas configurações de parênteses ainda produz objetos isomorfos. Similarmente, a o primeiro diagrama apenas nos diz que $\mathbf{1} \otimes A \cong A \otimes \mathbf{1}$. Apesar disto, estes diagramas possuem uma interpretação mais interessante do que simplesmente associatividade, especialmente porque exige-se que a identidade hexagonal também valha para as flechas inversas σ^{-1} . Podemos interpretar as flechas σ como tranças.

Para ver isto, considere que os objetos A, B são, por exemplo fios, alinhados da esquerda para a direita. A operação \otimes poderia ser interpretada como colocando os fios A e B lado a lado, por exemplo. Neste as flechas σ corresponderiam precisamente aos geradores do grupo de tranças, trocando A e B de lugar. Esta troca pode ser feita no sentido horário ou anti-horário, se escolhermos σ como correspondente à troca no sentido anti-horário, então σ^{-1} corresponde ao sentido horário e vice-versa. Com esta interpretação, também, as flechas α corresponderiam a alguma trança entre três elementos. De fato, esta interpretação é sustentada por um teorema devido a Joyal e Street [40, 38].

Este teorema diz que para cada isomorfismo construído a partir da composição dos isomorfismos α, λ, ρ e σ , corresponde uma trança e estes isomorfismos são iguais se, somente se suas tranças associadas são iguais. Não provaremos este teorema aqui pois acreditamos que não há muito a se ganhar nas technicalidades. Diremos apenas que é esperado já que a ideia de categoria trançada surgiu justamente como um modo de transportar ou generalizar as propriedades da categoria de todos os grupos de tranças, **Braid**, para outras categorias. Notamos que a categoria **Braid** é definida como a categoria cujos objetos são números naturais n e cujas flechas são as tranças em n fios, $n \rightarrow n$. Assim, os isomorfismos σ fornecem uma estrutura de trançamento no sentido mais genuíno a uma categoria e então enunciamos mais um axioma:

Axioma 2: A categoria \mathfrak{A} é uma categoria trançada.

Exigindo que \mathfrak{A} seja trançada conseguimos capturar quase toda a cinemática de anyons. O único aspecto que ausente é a extensão destas partículas. Como vimos, em se tratando da fusão de anyons, é possível que eles adquiram uma fase devido a rotações em torno de seu eixo o que corresponde à permutações dos anyons que compõem o anyon resultante. Em outras palavras, anyons possuem um grau de liberdade a mais, equivalente a serem partículas extensas e trançar suas **linhas** de mundo não é suficiente para capturar toda a sua cinemática. Lidaremos com esta questão agora.

4.2 - Rigidez e extensão

Completar a descrição cinemática de anyons na categoria \mathfrak{A} exige que revisitemos as regras de fusão de anyons e exploremos um pouco mais este grau de liberdade extra que eles possuem. A discussão que fizemos no final do capítulo 2, deixa a entender que este grau de liberdade só existe para anyons que são obtidos a partir de fusão. Além disso, é questionável até mesmo dizer que é um grau de liberdade genuíno de uma partícula porque a fusão de anyons não necessariamente forma um estado ligado, é apenas um agrupamento de anyons segundo suas distâncias relativas. Estas questões são resolvidas se considerarmos a abordagem de Wilczek [4] onde anyons são vistos como compostos de cargas e tubos de fluxo magnético. Neste caso o grau de liberdade extra é a posição relativa da carga ao tubo de fluxo e a fase resultante é proveniente do efeito Aharonov-Bohm.

Como discutimos, esta abordagem é baseada na alternativa que não tomamos no processo de quantização — não tomar o potencial de calibre como nulo. Apesar disto, vejamos se podemos espremer mais informação da alternativa que tomamos. Para justificar a existência deste grau de liberdade extra para todos os anyons, não só aqueles resultantes de fusão, partimos da seguinte observação sobre anyons abelianos: como o parâmetro estatístico é definido a menos de soma ou subtração de múltiplos inteiros de 2π , fusões progressivas de anyons de estatística θ acabam “dando a volta” e retornando ao anyon inicial de estatística θ . Para ver isto, lembre-se que as regras de fusão para anyons abelianos são dadas por [15]:

$$n^2\theta \otimes k^2\theta = (n + k)^2\theta$$

onde n e k são o número de anyons de tipo θ que forma os anyons sendo fundidos.

Assim, escrevamos $\theta = \pi/m$ para algum inteiro positivo m . A condição de m ser inteiro é meramente conveniência, a princípio, $m \in [1/2, \infty)$ já que $\theta \in [0, 2\pi]$. Além disto, supondo m par, temos que o resultado da fusão de $m - 1$ anyons de estatística π/m

é dado por

$$(m-1)^2 \frac{\pi}{m} = m\pi - 2\pi + \frac{\pi}{m} \equiv m\pi \frac{\pi}{m} \equiv \frac{\pi}{m}$$

onde \equiv é a identificação de que θ é definido a menos de soma de múltiplos inteiros de 2π e na última igualdade usamos o fato de m ser par. Um resultado análogo pode ser obtido se m for ímpar, porém é necessária a fusão de $2m-1$ anyons. Assim, o anyon fundamental do nosso sistema anyônico pode, ele próprio, ser obtido a partir de fusões sucessivas de si mesmo. Se este é o caso, não podemos realmente saber se estamos falando de um anyon realmente fundamental, se é que isto existe, ou um anyon composto. Acreditamos ser mais lógico, então, assumir que todos os anyons envolvidos são compostos de modo que todos possuem este grau de liberdade rotacional.

Mais interessante que isto, é que, ainda assumindo que m é par, temos:

$$m^2 \frac{\pi}{m} = m\pi \equiv 0.$$

Ou seja, o anyon trivial, $\mathbf{1}$ — que possui estatística bosônica — também pode ser obtido como uma fusão de anyons θ . Damos este nome ao anyon com estatística bosônica pois ele atua, em termos matemáticas como a identidade para a operação \otimes e em termos físicos como o vácuo do sistema, pois

$$\mathbf{1} \otimes k^2\theta = k^2\theta = k^2\theta \otimes \mathbf{1}.$$

Nesta expressão fica claro porque o anyon trivial corresponde à identidade e, em termos físicos, o associamos com o vácuo justamente porque ele corresponde a uma “tela em branco”. É efetivamente como se não houvesse nenhum anyon no sistema. Isto é interessante porque temos, então que o vácuo do sistema na verdade é resultado da fusão e portanto, necessariamente composto de anyons. Isto nos empurra de imediato à ideia do mar de Dirac e indica que a descrição mais adequada para anyons é realmente alguma teoria de campos. Reservaremos esta discussão para o capítulo final, no entanto.

Se adotarmos, então a visão do mar de Dirac e dissermos que o vácuo é um mar de anyons, podemos então obter a ideia de um anti-anyon, ou um anyon com a estatística oposta. Por definição, se a é um tipo de anyon de um dado tipo, então seu anti-anyon seria um a^* tal que

$$a \otimes a^* = \mathbf{1} = a^* \otimes a.$$

No caso de anyons abelianos de estatística $\theta = \pi/m$, isto corresponde à exigência que $a = n^2\theta$, $a^* = k^2\theta$ tal que $(n+k)^2\pi/m \equiv 0$. Em virtude das restrições sobre θ , n e k em geral não possuem uma relação tão simples quanto $n = -k$. Como vimos acima, se m for par, $(m-1)^2\pi/m$ é o anti-anyon de π/m . Apesar disto a estatística de $(m-1)^2\pi/m$ é equivalente à de π/m . Isto, inclusive, reforça todo o tratamento categorial em termos de isomorfismos, pois mesmo que o anyon $(m-1)^2\theta$ tenha a mesma estatística que θ eles

evidentemente não são iguais pois θ não é seu próprio anti-anyon.

Como esta discussão revela, havia mais informação a ser extraída da abordagem que adotamos. Estas informações, inclusive, completam o que precisávamos para terminar de a parte cinemática da nossa descrição categorial. Daremos continuidade a esta descrição agora, começando com a rotação de anyons. Linhas de mundo não são suficientes para capturar este grau de liberdade porque são unidimensionais, não há como torcer uma linha, para capturar isto é necessário estender estas linhas para fitas. Deste modo a rotação de um anyon ao redor é representada por uma torção na fita.

Definição 4.3: Uma **torção** em uma categoria monoidal trançada, \mathcal{C} é um isomorfismo natural θ dado pela família de isomorfismos

$$\theta_A : A \rightarrow A$$

para cada objeto $A \in \mathcal{C}$ tal que o diagrama

$$\begin{array}{ccc}
 A \otimes B & \xrightarrow{\theta_{A \otimes B}} & A \otimes B \\
 \theta_A \otimes \theta_B \downarrow & & \uparrow \sigma_{B,A} \\
 A \otimes B & \xrightarrow{\sigma_{A,B}} & B \otimes A
 \end{array}$$

comuta para todos os objetos A e B de \mathcal{C} .

Para visualizar precisamente o que este diagrama significa é interessante introduzir uma espécie de cálculo gráfico. Como as flechas correspondem a transformações de anyons, associamos a cada flecha uma linha de mundo representando a evolução temporal que ela representa. Em particular, já vimos que as flechas σ estão associadas aos geradores do grupo de trança. O diagrama acima está graficamente representando na figura 4.1 adiante.

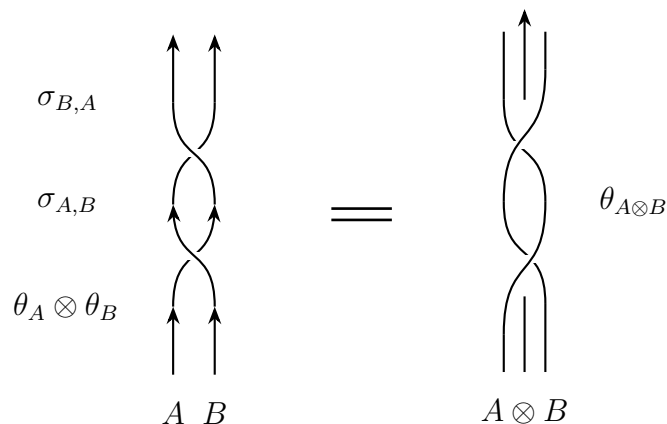


Figura 4.1: Representação visual da torção $\theta_{A \otimes B}$.

Com esta representação, é possível ver claramente que apesar de $\theta_{A \otimes B}$ ser uma flecha de $A \otimes B$ em $A \otimes B$, claramente é distinta da flecha identidade e de algum elemento do grupo de tranças. Ela representa uma rotação completa dos anyons A e B um ao redor do outro, essencialmente rotacionando o anyon $A \otimes B$ por 2π . Esta transformação não é necessariamente a identidade porque os anyons constituintes adquirem uma fase conforme suas estatísticas. Em particular, o caso da rotação de um único anyon, A , pode ser obtido se considerarmos a torção $\theta_{A \otimes \mathbf{1}}$, isto é, rotacionar A em torno do anyon trivial, $\mathbf{1}$. Equivalentemente, podemos pensar A como composto de outros anyons, como discutido anteriormente. A primeira alternativa é, em geral, mais conveniente. Com isto, exigimos, então que \mathfrak{A} seja equipada com uma torção:

Axioma 3: A categoria \mathfrak{A} é equipada com uma torção.

O próximo passo é formalizar a noção de anti-anyon. Para isto, estabelecemos uma operação de conjugação, $(-)^*$. Esta operação é formalizada pelo conceito de objeto **dual**:

Definição 4.4: Seja \mathbf{C} é uma categoria monoidal e A um objeto de \mathbf{C} . O **dual** de A é um objeto $A^* \in \mathbf{C}$ junto com duas flechas $i_A : \mathbf{I} \rightarrow A \otimes A^*$ e $e_A : A^* \otimes A \rightarrow \mathbf{I}$ tais que

$$\begin{array}{ccc}
 A^* \otimes \mathbf{I} & \xrightarrow{\text{id}_{A^*} \otimes i_A} & A^* \otimes (A \otimes A^*) \\
 \downarrow \sigma_{A^* \otimes \mathbf{I}} & & \downarrow \alpha_{A^*, A, A^*} \\
 \mathbf{I} \otimes A^* & \xleftarrow{e_A \otimes \text{id}_{A^*}} & (A^* \otimes A) \otimes A^*
 \end{array}
 \qquad
 \begin{array}{ccc}
 \mathbf{I} \otimes A & \xrightarrow{i_A \otimes \text{id}_A} & (A \otimes A^*) \otimes A \\
 \downarrow \sigma_{\mathbf{I} \otimes A} & & \downarrow \alpha_{A, A^*, A}^{-1} \\
 A \otimes \mathbf{I} & \xleftarrow{\text{id}_A \otimes e_A} & A \otimes (A^* \otimes A)
 \end{array}$$

comutam. Uma categoria monoidal \mathbf{C} é dita **rígida** se todo objeto possui um dual.

Essencialmente as flechas i e e representam a criação e aniquilação de um par anyon anti-anyon, respectivamente. Lembre-se \mathbf{I} é a identidade da operação monoidal \otimes . A criação e aniquilação de pares está de acordo com a ideia do mar de Dirac para anyons. Além disso, note que como definimos a fusão de anyons é um processo reversível. Anyons são fundidos quando se aproximam de modo que sua distância relativa é muito menor que suas distâncias a todos os outros anyons. Do mesmo modo, se afastarmos os anyons podemos separá-los. Assim, se o vácuo é resultado da fusão de vários anyons, a criação de pares consiste em separar um anyon e um anti-anyon do vácuo. O vácuo permanece inalterado porque a fusão do par os aniquila, essencialmente como se a “estatística total” do sistema fosse conservada, como uma espécie de carga.

As flechas i e e nos permitem descrever $(-)^*$ como um funtor $\mathbf{C} \rightarrow \mathbf{C}$. Note que dado uma flecha $f : A \rightarrow B$, podemos construir uma flecha $f^* : B^* \rightarrow A^*$ como

$$B^* \xrightarrow{\text{id}_B \otimes i_A} B^* \otimes (A \otimes A^*) \xrightarrow{\text{id}_{B^*} \otimes f \otimes \text{id}_{A^*}} (B^* \otimes B) \otimes A^* \xrightarrow{e_B \otimes \text{id}_{A^*}} A^*$$

onde implicitamente usamos que $B^* \cong \mathbf{I} \otimes B^*$, $B^* \otimes (A \otimes A^*) \cong (B^* \otimes A) \otimes A^*$ e $\mathbf{I} \otimes A^* \cong A$. Se quiséssemos ser explícitos, deveríamos compor os mapas acima com a flecha associadora α , especificamente com sua inversa e com as flechas, ditas unitoras, λ^{-1} e λ . Com esta construção, definimos o funtor contravariante: $(-)^* : \mathbf{C} \rightarrow \mathbf{C}$ pelas atribuições

$$A \longmapsto A^*;$$

$$(f : A \rightarrow B) \longmapsto (f^* : B^* \rightarrow A^*).$$

Por fim, notamos que se \mathbf{C} possui uma estrutura de tranças e de torção, é possível compatibilizar as flechas i e e com σ e θ . Isto é obtido exigindo-se que o seguinte diagrama comute:

$$\begin{array}{ccc} \mathbf{1} & \xrightarrow{i_A} & A \otimes A^* \\ \downarrow i_A & & \downarrow \text{id}_A \otimes \theta_{A^*} \\ A \otimes A^* & \xrightarrow{\theta_A \otimes \text{id}_{A^*}} & A \otimes A^* \end{array}$$

A compatibilização da estrutura rígida com de tranças e as torções gera diversos resultados sobre uma categoria [43], em particular, implica que $(A^*)^* = A^{**} \cong A$ naturalmente. Ou seja, o dual do dual de um objeto A é isomorfo a A , isto é muito similar ao que vimos sobre espaços vetoriais. Inclusive, categorias monoidais trançadas rígidas compatíveis com torções são chamadas categorias de fita e geralmente são definidas exigindo-se um isomorfismo natural entre A^{**} e A que satisfaz certas propriedades. As torções, então, são construídas em termos deste isomorfismo [40]. Apresentaremos uma definição equivalente, partindo de torções [43]:

Definição 4.5: Uma categoria \mathbf{F} é chamada **categoria de fita** se é uma categoria monoidal trançada com uma torção e que possui uma estrutura rígida, $\langle (-)^*, i, e \rangle$ compatível.

Isto nos leva ao nosso último axioma tratando da parte cinemática de anyons:

Axioma 4: A categoria \mathfrak{A} é uma categoria de fita.

4.3 – Semissimplicidade

A última parte que falta é tratar da questão dinâmica de anyons, relacionando esta descrição categórica com a descrição usual da mecânica quântica que se utiliza de espaços

de Hilbert. Esta conexão é feita por meio do conceito de **semissimplicidade** de uma categoria. Relembrando a expressão formal para a fusão de anyons temos:

$$\mathcal{A} \otimes \mathcal{B} = \mathcal{A} - \mathcal{B} \oplus \dots \oplus \mathcal{A} + \mathcal{B}$$

onde os tipos de anyons $\mathcal{A} - \mathcal{B}$, $\mathcal{A} + \mathcal{B}$ e intermediários são “tipos irreduzíveis”. Como dissemos, isto é em analogia com o que acontece com o momento angular de partículas compostas que não necessariamente é dado por uma representação irreduzível de $SU(2)$ e portanto é necessário fatorá-lo. A semissimplicidade captura precisamente esta ideia e ainda nos fornece espaços vetoriais naturalmente.

Para falarmos de semissimplicidade, no entanto, é necessário passarmos por vários conceitos preliminares, cada um acrescentando uma parte da solução final. Por este motivo, a discussão que segue será tão extensa quanto a discussão apresentada até agora neste capítulo já que é necessário discutir todos os conceitos intermediários em algum nível de detalhe. Começaremos por tratar de categorias ditas **pré-aditivas** e **aditivas**. Estas categorias essencialmente nos fornecem a operação \oplus da expressão acima. Depois trataremos, sem muitos detalhes, de categorias **pré-abelianas** e **abelianas** e por fim, de categorias semissimples.

4.3.1 - Categorias aditivas

Começamos com a definição de categorias pré-aditivas:

Definição 4.6: Uma categoria \mathbf{A} é dita **pré-aditiva** ou **ab** se para quaisquer pares de objetos $A, B \in \mathbf{A}$, a coleção de flechas $[A, B]$ é um grupo abeliano.

Chamamos atenção para o fato de que todas as coleções $[A, B]$ serem grupos abelianos não implica que a categoria \mathbf{A} é localmente pequena pois estes grupos não necessariamente são pequenos, isto é, a coleção de seus elementos não necessariamente forma um conjunto. Grupos que não são pequenos podem ser definidos de modo categorial como feito no exemplo no início do capítulo 3. É instrutivo explorar um pouco mais sobre como categorias pré-aditivas podem ser construídas.

Segundo Mac Lane [38], categorias deste tipo são construídas do seguinte modo: considere a categoria \mathbf{Ab} de grupos abelianos não necessariamente pequenos, uma categoria pré-aditiva consiste em uma coleção de objetos, para cada par de objetos A, B , um grupo abeliano $\mathcal{A}[A, B]$, para cada tripla de objetos A, B, C um homomorfismo de grupos abelianos

$$\mathcal{A}[B, C] \otimes \mathcal{A}[A, B] \rightarrow \mathcal{A}[A, C]$$

chamado de composição e para cada objeto A , um homomorfismo $\mathbb{Z} \rightarrow \mathcal{A}[A, A]$ de modo

a satisfazerem as propriedades associativa e de identidade para a composição. A operação \otimes é o produto tensorial de grupos abelianos. A categoria \mathbf{Ab} é uma categoria monoidal e a construção descrita acima corresponde exatamente à construção de uma **categoria enriquecida** sobre \mathbf{Ab} . Relembramos o que mencionamos brevemente quando abordamos categorias monoidais: uma categoria enriquecida é tal que suas as flechas são objetos de categoria monoidal onde o bifunctor \otimes é usado para definir a composição de flechas. Acima, denotamos a coleção de flechas de A a B por $\mathcal{A}[A, B]$ para evidenciar como objeto de \mathbf{Ab} , no que segue retomamos a notação usual, $[A, B]$.

No caso particular de categorias pré-aditivas, a composição obtida a partir de um produto tensorial é bilinear. Isto é, dados $f, f' \in [A, B]$ e $g, g' \in [B, C]$, tem-se

$$(g + g') \circ (f + f') = g \circ f + g \circ f' + g' \circ f + g' \circ f'.$$

Além disso cada coleção $[A, B]$ possui um elemento zero — a unidade do grupo — escrito $0 : A \rightarrow B$ chamado de flecha zero. Note que flecha zero aqui não corresponde à flecha zero de uma categoria que possui um elemento nulo, é simplesmente o nome da flecha correspondente à identidade do grupo. Note que a bilinearidade da composição implica que a composição com uma flecha zero é, também, uma flecha zero:

$$g \circ 0 = g \circ (0 + 0) = g \circ 0 + g \circ 0 \implies g \circ 0 = 0.$$

Notamos que a notação acima está sobrecarregada pois usamos 0 para indicar a flecha zero entre quaisquer dois elementos A e B e é necessário manter em mente a qual grupo de flechas pertence 0 .

Categorias pré-aditivas gozam da seguinte propriedade:

Proposição 4.1: Se O é um objeto de uma categoria pré-aditiva, \mathbf{A} , então as seguintes propriedades são equivalentes:

- O é inicial;
- O é final;
- $\text{id}_O = 0 : O \rightarrow O$;
- $[O, O] = \mathbf{0}$, onde $\mathbf{0}$ é o grupo trivial.

Prova: Se O é inicial, então existe um único mapa $O \rightarrow O$, logo $\text{id}_O = 0$ e $[O, O] = \mathbf{0}$. Se $\text{id}_O = 0$ então qualquer flecha $f : A \rightarrow O$ satisfaz

$$f = \text{id}_O \circ f = 0 \circ f = 0 : A \rightarrow O$$

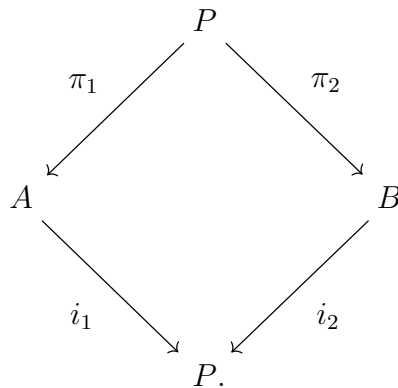
de modo que existe uma única flecha, 0 de $A \rightarrow O$ e O é final. Por dualidade, se O é final, O é inicial. Por fim, pelos axiomas de categoria se $[O, O] = \mathbf{0}$, tem-se que $0 = \text{id}_O$ obrigatoriamente e, se $\text{id}_O = 0$, supomos que exista $f : O \rightarrow O$, assim

$$f = \text{id}_O \circ f = 0 \circ f = 0 : O \rightarrow O.$$

Ou seja, $\text{id}_O = 0 \implies [O, O] = \mathbf{0}$. Isto completa a prova. □

Note, em particular que este resultado implica que se A é pré-aditiva e possui um elemento inicial ou final, ela possui um elemento zero, já que o elemento inicial será também final e vice-versa. Se $\mathbf{0}$ é o elemento zero de A , então as flechas únicas $A \rightarrow \mathbf{0}$ e $\mathbf{0} \rightarrow B$ são as flechas zero de $[A, \mathbf{0}]$ e $[\mathbf{0}, B]$, isto é, são o elemento identidade do grupo. Além disso, a flecha $0 : A \rightarrow B$ obtida da composta destas duas flechas zero é única e é a flecha zero — elemento identidade — de $[A, B]$.

A próxima propriedade sobre categorias pré-aditivas e a que mais nos interessa trata de biprodutos em A . Mencionamos brevemente biprodutos quando tratamos de produtos e coprodutos, seção 3.4.1. Um biproduto é um produto que também é um coproduto, ou seja, dados dois objetos A, B , seu biproduto é o objeto P equipado com flechas π_1, π_2, i_1, i_2 como mostrado no diagrama abaixo:



Além disso, algumas condições devem ser satisfeitas pelas projeções e inclusões. Estas propriedades essencialmente asseguram a compatibilidade entre o produto e o coproduto, lembre-se que eles são definidos a menos de isomorfismos. A compatibilidade consiste essencialmente em garantir que o qualquer produto de objetos A e B é isomorfo a qualquer coproduto destes mesmos objetos. Existem múltiplos modos de assegurar esta compatibilidade [44], porém todos são equivalentes. Apresentaremos a versão usada quando falamos de categorias pré-aditivas e que foi a definição original de biproduto segundo Mac Lane [38]:

$$\pi_1 \circ i_1 = \text{id}_A; \quad \pi_2 \circ i_2 = \text{id}_B; \quad i_1 \circ \pi_1 + i_2 \circ \pi_2 = \text{id}_P. \quad (4.1)$$

O isomorfismo entre o produto e o coproduto que assegura a existência do biproduto é

o que nos interessa aqui pois ele produz um modo de determinar flechas entre biprodutos finitos, isto é, biprodutos de mais de dois objetos, por meio de uma matriz de flechas entre cada um dos objetos. Isto nos será muito útil. Vejamos então este resultado. Primeiro:

Proposição 4.2: Se A é pré-aditiva, então dois objetos A e B possuem um produto em A se, e somente se, possuem um coproduto em A . Alternativamente, dois objetos possuem um produto ou coproduto se, e somente se, possuem um biproduto.

Prova: Primeiro, suponha que dois objetos A e B possuam um biproduto. Então, pelo diagrama de biproduto

$$A \begin{array}{c} \xrightarrow{i_1} \\ \xleftarrow{\pi_1} \end{array} P \begin{array}{c} \xleftarrow{i_2} \\ \xrightarrow{\pi_2} \end{array} B$$

e pelas condições do biproduto, equações 4.1, tem-se:

$$\begin{aligned} \pi_1 \circ i_2 &= \pi_1 \circ \text{id}_P \circ i_2 = \pi_1 \circ (i_1 \circ \pi_1 + i_2 \circ \pi_2) \circ i_2 \\ &= \pi_1 \circ i_1 \circ \pi_1 \circ i_2 + \pi_1 \circ i_2 \circ \pi_2 \circ i_2 \\ &= \text{id}_A \circ \pi_1 \circ i_2 + \pi_1 \circ i_2 \circ \text{id}_B \\ &= \pi_1 \circ i_2 + \pi_1 \circ i_2 \end{aligned}$$

donde $\pi_1 \circ i_2 = 0$. Similarmente, $\pi_2 \circ i_1 = 0$. Agora, considere qualquer diagrama

$$A \xleftarrow{f_1} C \xrightarrow{f_2} B.$$

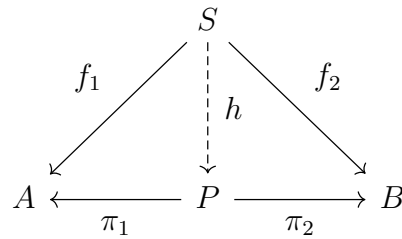
A soma $h = i_1 \circ f_1 + i_2 \circ f_2 : C \rightarrow P$, satisfaz

$$\begin{aligned} \pi_1 \circ h &= \pi_1 \circ (i_1 \circ f_1 + i_2 \circ f_2) \\ &= \pi_1 \circ i_1 \circ f_1 + \pi_1 \circ i_2 \circ f_2 \\ &= \text{id}_A \circ f_1 + 0 \circ f_2 = f_1 \end{aligned}$$

e similarmente, $\pi_2 \circ h = f_2$. Por outro lado, dado $h' : C \rightarrow P$ tal que $\pi_i \circ h' = f_i$ para $i = 1, 2$ então

$$\begin{aligned} h' &= \text{id}_P \circ h' = (i_1 \circ \pi_1 + i_2 \circ \pi_2) \circ h \\ &= i_1 \circ \pi_1 \circ h + i_2 \circ \pi_2 \circ h \\ &= i_1 \circ f_1 + i_2 \circ f_2 = h. \end{aligned}$$

Ou seja, para qualquer objeto C e flechas $f_1 : C \rightarrow A$ e $f_2 : C \rightarrow B$ existe uma única flecha $h : C \rightarrow P$ tal que o diagrama



comute. Ou seja,

$$A \xleftarrow{\pi_1} P \xrightarrow{\pi_2} B$$

de fato é um produto. O mapa $h \mapsto \langle f_1, f_2 \rangle$ é um isomorfismo

$$[C, P] \cong [C, A] \oplus [C, B]$$

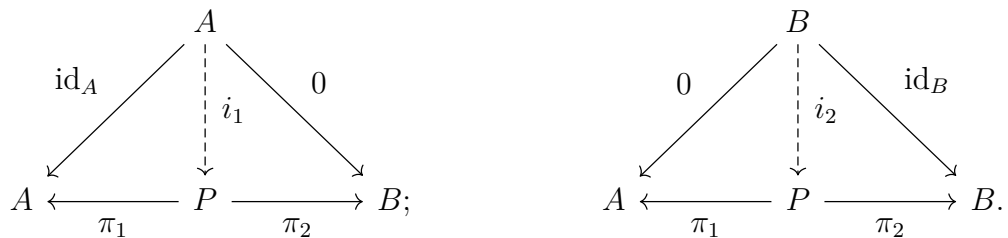
onde \oplus é a soma direta de grupos abelianos. A soma direta de grupos abelianos \mathcal{A} e \mathcal{B} é o grupo $\mathcal{A} \oplus \mathcal{B}$ cujo conjunto subjacente é o produto cartesiano $\mathcal{A} \times \mathcal{B}$ e cuja operação de grupo é:

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 \star a_2, b_1 \star b_2)$$

onde \star e $*$ são as operações de grupo de \mathcal{A} e \mathcal{B} respectivamente. A soma direta, inclusive, é um exemplo de biproduto em \mathbf{Ab} . Continuando, se, por outro lado, tem-se um produto

$$A \xleftarrow{\pi_1} P \xrightarrow{\pi_2} B$$

por definição existe uma única flecha $i_1 : A \rightarrow P$ com componentes $\pi_1 \circ i_1 = \text{id}_A$ e $\pi_2 \circ i_1 = 0$ e uma única flecha $i_2 : B \rightarrow P$ com componentes $\pi_1 \circ i_2 = 0$ e $\pi_2 \circ i_2 = \text{id}_B$. Estas flechas são exatamente as flechas mediadoras para os diagramas:



Note que a existência das flechas zero decorre do fato de \mathbf{A} ser pré-aditiva. Assim, tem-se

$$\begin{aligned}
 \pi_1 \circ (i_1 \circ \pi_1 + i_2 \circ \pi_2) &= \pi_1 \circ i_1 \circ \pi_1 + \pi_1 \circ i_2 \circ \pi_2 = \pi_1 + 0 = \pi_1; \\
 \pi_2 \circ (i_1 \circ \pi_1 + i_2 \circ \pi_2) &= \pi_2 \circ i_1 \circ \pi_1 + \pi_2 \circ i_2 \circ \pi_2 = 0 + \pi_2 = \pi_2.
 \end{aligned}$$

Ou seja, $i_1 \circ \pi_1 + i_2 \circ \pi_2 : P \rightarrow P$ é a flecha única com componentes π_1 e π_2 e portanto, pelas propriedades de produto, é a flecha identidade id_P pois tem-se:

$$\begin{array}{ccc}
 & P & \\
 \pi_1 \swarrow & & \searrow \pi_2 \\
 A & \xleftarrow{\pi_1} & P & \xrightarrow{\pi_2} & B \\
 & & \downarrow \text{id}_P & & \\
 & & P & &
 \end{array}$$

Por fim, as flechas π_1, π_2, i_1 e i_2 formam um diagrama de biproduto e satisfazem às condições 4.1, logo, se dois objetos de \mathbf{A} possuem um produto, eles possuem um biproduto. Note, também, que as flechas i_1 e i_2 são justamente as inclusões de um coproduto. □

Assim como o produto e o coproduto, o biproduto de dois objetos é determinado a menos de isomorfismos. Além disso, assim como o produto e o coproduto, ele define um bifunctor em \mathbf{A} quando existe para todos os pares de objetos A e B de \mathbf{A} . Considere uma categoria \mathbf{C} onde todos os pares de objetos possuem um produto. Representando o produto por \times , esta operação define um bifunctor de $\mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C}$, onde o símbolo \times em $\mathbf{C} \times \mathbf{C}$ se refere ao produto de categorias. Este bifunctor é dado por:

$$\begin{aligned}
 \langle A, B \rangle &\longmapsto A \times B \\
 \langle f : A \rightarrow A', g : B \rightarrow B' \rangle &\longmapsto (f \times g : A \times A' \rightarrow B \times B').
 \end{aligned}$$

Tem-se o seguinte diagrama que comuta:

$$\begin{array}{ccccc}
 A & \xleftarrow{\pi_1} & A \times B & \xrightarrow{\pi_2} & B \\
 f \downarrow & & \downarrow f \times g & & \downarrow g \\
 A' & \xleftarrow{\pi'_1} & A' \times B' & \xrightarrow{\pi'_2} & B'
 \end{array}$$

ou seja, a flecha $f \times g$ satisfaz:

$$\pi'_1 \circ (f \times g) = f \circ \pi_1 \quad \text{e} \quad \pi'_2 \circ (f \times g) = g \circ \pi_2.$$

Similarmente, uma categoria \mathbf{C} onde quaisquer pares de objetos possuem um coproduto também define um bifunctor, $\sqcup : \mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C}$ por:

$$\begin{aligned}
 \langle A, B \rangle &\longmapsto A \sqcup B \\
 \langle f : A \rightarrow A', g : B \rightarrow B' \rangle &\longmapsto (f \sqcup g : A \sqcup A' \rightarrow B \sqcup B').
 \end{aligned}$$

diagramaticamente, tem-se:

$$\begin{array}{ccccc}
 A & \xrightarrow{i_1} & A \sqcup B & \xleftarrow{i_2} & B \\
 \downarrow f & & \downarrow f \sqcup g & & \downarrow g \\
 A' & \xrightarrow{i'_1} & A' \sqcup B' & \xleftarrow{i'_2} & B'
 \end{array}$$

ou seja

$$(f \sqcup g) \circ i_1 = i'_1 \circ f \quad \text{e} \quad (f \sqcup g) \circ i_2 = i'_2 \circ g.$$

Como o biproduto é simultaneamente um produto e um coproduto, em uma categoria pré-aditiva — para que continuemos com a mesma definição de biproduto — onde todos os pares de objetos possuem biproduto, representado por \oplus , ele também define um bifuntor: $\oplus : \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$. Este bifuntor leva pares de objetos em seu biproduto

$$\langle A, B \rangle \longmapsto A \oplus B$$

e leva pares de flechas entre objetos em flechas entre o biproduto destes objetos. No entanto, como o biproduto é simultaneamente um produto e um coproduto, existem duas possibilidades: podemos usar a atribuição de flechas do bifuntor \times ou do bifuntor \sqcup . Correspondendo às seguintes condições:

$$\pi'_1 \circ (f \oplus g) = f \circ \pi_1 \quad \text{e} \quad \pi'_2 \circ (f \oplus g) = g \circ \pi_2$$

ou

$$(f \oplus g) \circ i_1 = i'_1 \circ f \quad \text{e} \quad (f \oplus g) \circ i_2 = i'_2 \circ g.$$

Pelas propriedades do biproduto ambas são equivalentes. Para ver isto, considere as equações 4.1 e $\pi_1 \circ i_2 = 0$ e $\pi_2 \circ i_1 = 0$, que decorrem destas equações. Suponha que definamos o bifuntor como o bifuntor produto, tem-se

$$\begin{aligned}
 (f \oplus g) \circ i_1 &= \text{id}_{A' \oplus B'} \circ (f \oplus g) \circ i_1 = (i'_1 \circ \pi'_1 + i'_2 \circ \pi'_2) \circ (f \oplus g) \circ i_1 \\
 &= i'_1 \circ \pi'_1 \circ (f \oplus g) \circ i_1 + i'_2 \circ \pi'_2 \circ (f \oplus g) \circ i_1 \\
 &= i'_1 \circ f \circ \pi_1 \circ i_1 + i'_2 \circ g \circ \pi_2 \circ i_1 \\
 &= i'_1 \circ f \circ \text{id}_A + i'_2 \circ g \circ 0 = i'_1 \circ f.
 \end{aligned}$$

Do mesmo modo, $(f \oplus g) \circ i_2 = i'_2 \circ g$. Por outro lado, definindo-se o bifuntor \oplus segundo

o bifuntor do coproduto, tem-se:

$$\begin{aligned}
 \pi'_1 \circ (f \oplus g) &= \pi'_1 \circ (f \oplus g) \circ \text{id}_{A \oplus B} = \pi'_1 \circ (f \oplus g) \circ (i_1 \circ \pi_1 + i_2 \circ \pi_2) \\
 &= \pi'_1 \circ (f \oplus g) \circ i_1 \circ \pi_1 + \pi'_1 \circ (f \oplus g) \circ i_2 \circ \pi_2 \\
 &= \pi'_1 \circ i'_1 \circ f \circ \pi_1 + \pi'_1 \circ i'_2 \circ g \circ \pi_2 \\
 &= \text{id}_{A'} \circ f \circ \pi_1 + 0 \circ g \circ \pi_2 = f \circ \pi_1.
 \end{aligned}$$

Similarmente, $\pi'_2 \circ (f \oplus g) = g \circ \pi_2$. Diagramaticamente isto corresponde ao diagrama comutativo a seguir. Note que este diagrama representa, essencialmente uma transformação natural entre os bifuntores \oplus definidos como o bifuntor produto, \times e como o bifuntor coproduto, \sqcup . Além disso, como as flechas que compõem a transformação natural são as flechas identidade

$$A \oplus B \xrightarrow{\text{id}_{A \oplus B}} A \oplus B,$$

esta transformação natural é um isomorfismo natural. Eis o diagrama:

$$\begin{array}{ccc}
 A \oplus B & \xrightarrow{f \times g} & A' \oplus B \\
 \text{id}_{A \oplus B} \downarrow & & \downarrow \text{id}_{A' \oplus B} \\
 A \oplus B & \xrightarrow{f \sqcup g} & A' \oplus B'.
 \end{array}$$

Com este resultado podemos falar apenas d'O bifuntor \oplus . Agora, no início, dissemos que a o bripoduto é a propriedade que mais nos interessa ao tratar de semissimplicidade. Isto se dá porque o bripoduto fornece uma espécie de cálculo matricial para as flechas da categoria. Cálculo aqui no sentido de “calcular” não no sentido de cálculo diferencial e integral. Para ver como isto surge, basta notar que, em vista do funtor bripoduto, uma flecha para ou de um bripoduto é determinada inteiramente em termos de suas componentes. Uma flecha

$$C \xrightarrow{f} A \oplus B$$

é determinada por suas componentes

$$\pi_1 \circ f = f_1 \quad \text{e} \quad \pi_2 \circ f = f_2$$

donde podemos escrevê-la como

$$f = \begin{pmatrix} f_1 & f_2 \end{pmatrix}.$$

Do mesmo modo, uma flecha

$$A \oplus B \xrightarrow{g} D$$

é determinada por suas componentes

$$g \circ i_1 = g_1 \quad \text{e} \quad g \circ i_2 = g_2$$

de modo que escrevemos

$$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}.$$

Agora, considere uma flecha entre dois biprodutos:

$$\begin{array}{ccccc} A & \xrightleftharpoons[i_1]{\pi_1} & A \oplus B & \xrightleftharpoons[i_2]{\pi_2} & B \\ \downarrow f & & \downarrow f \oplus g & & \downarrow g \\ C & \xrightleftharpoons[i'_1]{\pi'_1} & C \oplus D & \xrightleftharpoons[i'_2]{\pi'_2} & D \end{array}$$

a flecha $f \otimes g$ é determinada por simultaneamente por suas componentes de produto e coproduto. Pelo diagrama tem-se:

$$\begin{aligned} \pi'_1 \circ (f \oplus g) \circ i_1 &= f_{11} : A \rightarrow C \\ \pi'_2 \circ (f \oplus g) \circ i_1 &= f_{21} : A \rightarrow D \\ \pi'_1 \circ (f \oplus g) \circ i_2 &= f_{12} : B \rightarrow C \\ \pi'_2 \circ (f \oplus g) \circ i_2 &= f_{22} : B \rightarrow D \end{aligned}$$

Com este resultado, podemos escrever, por exemplo, a flecha identidade

$$\text{id}_{A \oplus B} = i_1 \circ \pi_1 + i_2 \circ \pi_2$$

como

$$\text{id}_{A \oplus B} = \begin{pmatrix} \text{id}_A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \text{id}_B \end{pmatrix} = \begin{pmatrix} \text{id}_A & 0 \\ 0 & \text{id}_B \end{pmatrix}.$$

Note que a composição se comporta exatamente como um produto tensorial, o que é esperado já que é oriunda do produto tensorial de grupos abelianos. Apesar disso, alertamos que nem todas as propriedades usuais encontradas no produto tensorial de espaços vetoriais podem estar disponíveis pois estamos tratando do produto tensorial apenas em grupos. Posteriormente, equiparemos estes grupos com uma multiplicação por escalar de modo a obter espaços vetoriais propriamente.

Por fim, é possível iterar o biproduto e definir biprodutos de um número arbitrário de elementos. Lembrando da construção do produto e do coproduto como limites, isto

corresponde a considerar diagramas cujo grafo subjacente possui n elementos e apenas as arestas identidade. Assim, dados C_1, \dots, C_n em uma categoria \mathbf{C} que possui biproduto para todos os pares de elementos, tem-se que o biproduto de n elementos é caracterizado por diagramas

$$C_i \xrightarrow{i_j} \bigoplus_j C_j \xrightarrow{\pi_k} C_K$$

para $j, k = 1, 2, \dots, n$ e equações

$$i_1 \circ \pi_1 + \dots + i_n \circ \pi_n = \text{id}_{\bigoplus_j C_j}$$

e

$$\pi_k \circ i_j = \delta_{kj} = \begin{cases} 0 & \text{se } k \neq j; \\ \text{id}_{C_j} & \text{se } k = j. \end{cases}$$

Além disso, dados D_1, D_2, \dots, D_m em \mathbf{C} , existe um isomorfismo [38]

$$\left[\bigoplus_j C_j, \bigoplus_k D_k \right] \cong \bigoplus_{j,k} [C_j, D_k].$$

onde o último símbolo \bigoplus representa o produto direto de grupos abelianos. Este isomorfismo implica que toda flecha

$$\bigoplus_j C_j \xrightarrow{f} \bigoplus_k D_k$$

é determinada por uma matriz $m \times n$ de componentes

$$f_{kj} = \pi'_k \circ f \circ i_j : C_j \rightarrow D_k.$$

Desta equação também vê-se que a composição de flechas, então é dada simplesmente pela multiplicação das matrizes correspondente às flechas. Dado:

$$\bigoplus_j C_j \xrightarrow{f} \bigoplus_k D_k \xrightarrow{g} \bigoplus_l F_l$$

tem-se

$$\begin{aligned} (g \circ f)_{lj} &= \pi''_l \circ (g \circ f) \circ i_j = \pi''_l \circ g \circ \text{id}_{\bigoplus_k D_k} \circ f \circ i_j \\ &= \sum_k \pi''_l \circ g \circ i'_k \circ \pi'_k \circ f \circ i_j = \sum_k f_{lk} \circ g_{kj}. \end{aligned}$$

Agora note que, se uma categoria pré-aditiva possui os produtos finitos, isto é, para qualquer número de objetos C_i o produto $\times_i C_i$ existe, então ela possui todos os coprodutos e todos os biprodutos finitos, já que é pré-aditiva. Além disso, ela possui um objeto

final, dado pelo produto de zero objetos [38]. Portanto, possui também um objeto inicial e um objeto nulo. Assim, fazemos a seguinte definição:

Definição 4.7: Uma categoria \mathbf{A} é dita **aditiva** se for pré-aditiva e se possuir um objeto zero e para todos os pares de objetos, A e B de \mathbf{A} , o biproduto $A \oplus B$ existir.

Note que a definição é mais forte do que o necessário, no entanto, é a definição padrão. Podemos agora passar para a próxima etapa onde definimos categorias ditas **abelianas**. Para falar de categorias abelianas serão necessários mais alguns conceitos e então trabalharemos como fizemos para categorias aditivas: primeiro exploraremos o significado destes conceitos e então definiremos formalmente categorias abelianas. Assim com o caso de categorias aditivas, categorias abelianas serão aquelas que possuem uma propriedade útil definida a partir destes conceitos.

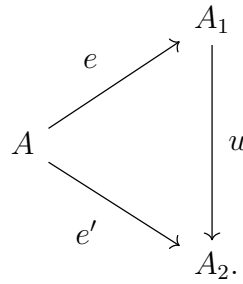
4.3.2 - Categorias abelianas

O primeiro conceito preliminar e seu dual que trataremos são, respectivamente **subobjetos** e **objetos quociente**. O conceito de subobjeto é essencialmente uma generalização categorial do conceito de “objetos que estão contidos em outros objetos” como subconjuntos, subgrupos, subespaços e assim por diante. Similarmente, objetos quocientes são uma generalização do conceitos de conjuntos quocientes, grupos quocientes, espaços quocientes e assim por diante. Naturalmente, como em teoria de categorias a estrutura interna dos objetos não é o que nos interessa na maioria dos casos, a definição destes conceitos é feita por meio de flechas. Começamos por estabelecer uma relação:

Definição 4.8: Dadas duas flechas mônicas $m : A_1 \rightarrow A$ e $m' : A_2 \rightarrow A$ de uma categoria \mathbf{C} , dizemos que m' está contida em m , representado por $m' \leq m$ se existe $u : A_2 \rightarrow A_1$ tal que o diagrama abaixo comute:

$$\begin{array}{ccc}
 & A_1 & \\
 & \uparrow & \searrow m \\
 & u & \\
 & A_2 & \nearrow m' \\
 & & A
 \end{array}$$

Dualmente, uma flecha épica $e' : A \rightarrow A_2$ contém outra, $e : A \rightarrow A_1$, se existe $u : A_1 \rightarrow A_2$ tal que o diagrama



comuta e escrevemos $e' \geq e$.

Pela definição, uma flecha mónica, m' , está contida em outra, m , se ela “fatora através dela”, isto é, se existe u tal que

$$m' = m \circ u.$$

Similarmente, uma flecha épica e' contém outra, e , se fatora através dela, ou seja, se existe u tal que

$$e' = u \circ e.$$

As duas definições são similares, sendo duais mas exigem um pouco de atenção, por este motivo, reservaremos \leq para falar de flechas mónicas, com $m' \leq m$ significando que m' está contida em m e \geq para falar de flechas épicas com $e' \geq e$. Evitaremos usar $e \leq e'$ para significar que uma flecha épica e está contida em e' e similarmente para flechas mónicas para evitar confusão. Apesar do costume de entender \leq e \geq como “dois lados” de uma mesma relação, aqui as usamos como relações distintas. Por fim, note que a flecha u é única. Para ver isto, suponha que tenhamos duas flechas u, v tais que $m' = m \circ u$ e $m' = m \circ v$, então podemos escrever

$$m' = m \circ u \implies m \circ v = m \circ u \implies v = u$$

pois m é mónica. Similarmente, se tivermos $e' = u \circ e$ e $e' = v \circ e$, então

$$e' = u \circ e \implies v \circ e = u \circ e \implies v = u.$$

Antes de prosseguirmos, note algumas propriedades úteis sobre flechas mónicas e épicas. Primeiro, dadas duas flechas mónicas $m : A \rightarrow B$ e $m' : B \rightarrow C$, a composição $m' \circ m : A \rightarrow C$ também é mónica. Para ver isto suponha que $(m' \circ m) \circ f = (m' \circ m) \circ g$ para flechas $f, g : O \rightarrow A$ de \mathbf{C} , então

$$\begin{aligned} (m' \circ m) \circ f = (m' \circ m) \circ g &\implies m' \circ (m \circ f) = m' \circ (m \circ g) \\ &\implies m \circ f = m \circ g \implies f = g \end{aligned}$$

já que ambas m e m' são mónicas. Similarmente, se $e : A \rightarrow B$ e $e' : B \rightarrow C$ são

épicas, a composição $e' \circ e : A \rightarrow C$ também é. Dadas flechas $f, g : C \rightarrow D$ tais que $f \circ (e' \circ e) = g \circ (e' \circ e)$, tem-se:

$$\begin{aligned} f \circ (e' \circ e) = g \circ (e' \circ e) &\implies (f \circ e') \circ e = (g \circ e') \circ e \\ &\implies f \circ e' = g \circ e' \implies f = g \end{aligned}$$

pois e e e' são épicas. Além disso, dada uma flecha mónica $m = h \circ u$, tem-se que u também é mónica. Para ver isto, suponha que $u \circ f = u \circ g$ para flechas f e g de \mathbf{C} . Então

$$u \circ f = u \circ g \implies h \circ u \circ f = h \circ u \circ g \implies m \circ f = m \circ g \implies f = g$$

pois m é mónica. Do mesmo modo, se $e = u \circ h$ é épica, então u é épica. Supondo $f \circ u = g \circ u$ tem-se

$$f \circ u = g \circ u \implies f \circ u \circ h = g \circ u \circ h \implies f \circ e = g \circ e \implies f = g$$

pois e' é épica. Com estas propriedades podemos estabelecer o seguinte resultado:

Proposição 4.3: As relações \leq entre flechas mónicas e \geq entre flechas épicas são reflexivas e transitivas.

Prova:

- A reflexividade é direta, pois para qualquer flecha, $f : A \rightarrow B$, tem-se $f = f \circ \text{id}_A = \text{id}_B \circ f$. Assim, se f for mónica, $f \leq f$ e se for épica, $f \geq f$.
- Para a transitividade, considere flechas mónicas $m : A_1 \rightarrow A$, $m' : A_2 \rightarrow A$ e $m'' : A_3 \rightarrow A$ tais que $m'' \leq m'$ e $m' \leq m$ então existem flechas $u : A_2 \rightarrow A_1$ e $u' : A_3 \rightarrow A_2$ tais que $m'' = m' \circ u'$ e $m' = m \circ u$ donde tem-se que $m'' \leq m$ com a flecha $v = u \circ u' : A_3 \rightarrow A_1$. De fato

$$m \circ u \circ u' = m' \circ u' = m''.$$

A situação é análoga para flechas épicas, $e : A \rightarrow A_1$, $e' : A \rightarrow A_2$ e $e'' : A \rightarrow A_3$ tais que $e' \geq e$ e $e'' \geq e'$. Logo \leq e \geq são transitivas.

□

Com as relações \leq e \geq em \mathbf{C} , definimos as seguintes relação de equivalência entre as flechas mónicas e épicas de \mathbf{C} :

Definição 4.9: Dadas flechas mónicas m e m' e épicas e e e' de \mathbf{C} , definimos relações por:

$$m \equiv m' \iff m \leq m' \wedge m' \leq m$$

e

$$e \equiv e' \iff e \geq e' \wedge e' \geq e.$$

Usaremos o mesmo símbolo, \equiv para as duas relações mas mantenha em mente que a princípio tratam-se de relações diferentes, uma vale para flechas mónicas e outra para flechas épicas.

As duas relações definidas acima são relações de equivalência. Elas são reflexivas e transitivas pois \leq e \geq o são e, ambas são simétricas pois

$$m \leq m' \wedge m' \leq m \iff m' \leq m \wedge m \leq m'$$

e similarmente para \geq . Esta definição é equivalente a afirmar que as fontes de duas flechas mónicas tais que $m \equiv m'$ são isomorfas e similarmente, que os alvos de duas flechas épicas $e \equiv e'$ são isomorfos. Para ver isto, lembramos que se $m' \leq m$ existe uma **única** flecha u tal que $m' = m \circ u$ e similarmente para flechas épicas. Então, considere $m : A_1 \rightarrow A$ e $m' : A_2 \rightarrow A$ tais que $m' \leq m$ e $m \leq m'$, então $m' = m \circ u$ e $m = m' \circ u'$ para flechas $u : A_2 \rightarrow A_1$ e $u' : A_1 \rightarrow A_2$ únicas. Logo

$$m' = m \circ u \implies m' = m' \circ u' \circ u \implies \text{id}_{A_2} = u' \circ u.$$

do mesmo modo obtém-se que $\text{id}_{A_1} = u \circ u'$, onde usamos o fato que m e m' são flechas mónicas. As flechas u e u' , então, formam um par de flechas inversas e tem-se $A_1 \cong A_2$. A prova para flechas épicas é análoga. Isto nos permite reescrever a definição acima como

Definição 4.10: Dadas flechas mónicas $m : A_1 \rightarrow A$ e $m' : A_2 \rightarrow A$, definimos

$$m \equiv m' \iff A_1 \cong A_2.$$

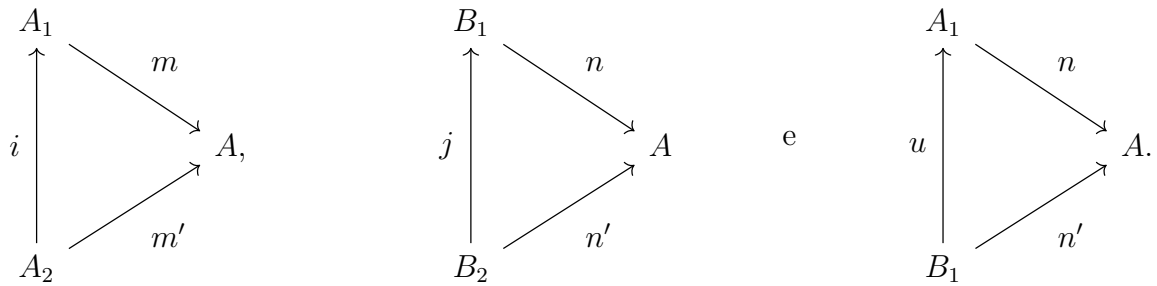
Similarmente, dadas flechas épicas $e : A \rightarrow A_1$ e $e' : A \rightarrow A_2$,

$$e \equiv e' \iff A_1 \cong A_2.$$

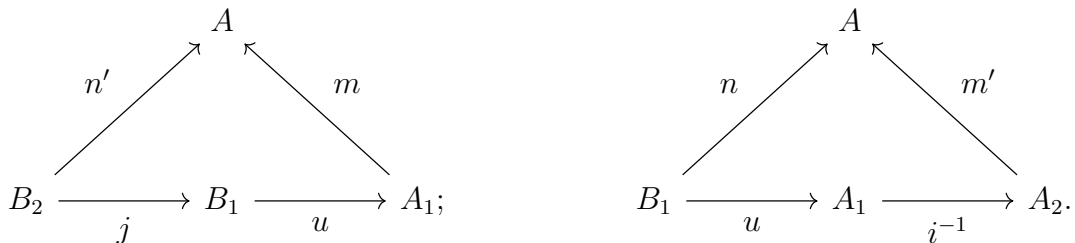
Como relações de equivalência, estas relações particionam as coleções de flechas mónicas e épicas de uma categoria \mathbf{C} . As classes de equivalência destas flechas são precisamente os subobjetos e objetos quocientes de \mathbf{C} . Formalmente

Definição 4.11: Dado um objeto A de uma categoria \mathbf{C} , um **subobjeto** $[m]$ de A é uma classe de equivalência de flechas mónicas, $m : X \rightarrow A$ sob a relação \equiv e um **objeto quociente**, $[e]$, de A é uma classe de equivalência de flechas épicas, $e : A \rightarrow X$, sob a relação \equiv em \mathbf{C} .

As relações \leq e \geq se estendem para as coleções de subobjetos e objetos quocientes, respectivamente, definido uma ordem parcial na coleção de modo que podemos ordenar subobjetos e objetos quocientes por inclusão, isto é, podemos falar de um subobjeto contido em outro e de objetos quocientes que contêm outros. Estendemos \leq e \geq para subobjetos e objetos quocientes da seguinte forma: considere flechas mónicas m, m' e n, n' tais que $m \equiv m'$, e $n \equiv n'$ e $n \leq m$. Diagramaticamente tem-se



Onde i, j são isomorfismos únicos e u é única. Nestas condições, podemos construir os seguintes diagramas que comutam:



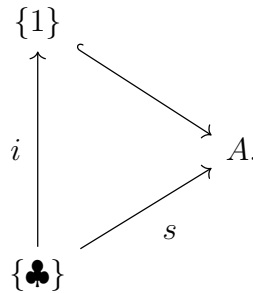
Deste diagramas concluímos que $n' \leq m$, $n \leq m'$ e também que $n' \leq m'$, apesar de não termos desenhado este diagrama. Ou seja, se $n \leq m$ então $[n] \leq [m]$ pois todos os subobjetos estão conectados por pares de flechas inversíveis. O mesmo resultado vale para objetos quocientes.

Para ver que \leq e \geq são de fato ordens parciais, note que as propriedades reflexivas e transitivas de \leq e \geq são transportadas para subobjetos e objetos quocientes diretamente em vista das definições de \equiv . Além disso, a antissimetricidade que decorre do fato que: se $m \leq m'$ e $m' \leq m$, então $m \equiv m'$ de modo que $[m] = [m']$. Mas, se $m \leq m'$, tem-se que $[m] \leq [m']$ e similarmente, $m' \leq m \implies [m'] \leq [m]$, logo

$$([m] \leq [m']) \wedge ([m'] \leq [m]) \implies [m] = [m'].$$

O mesmo resultado vale para \geq e objetos quocientes.

Por fim, antes de passarmos ao próximo conceito, vejamos como subobjetos, por exemplo, correspondem às subestruturas que estamos familiarizados, como subconjuntos ou subgrupos. Em Set , considere um conjunto $A = \{1, 2, 3\}$. O conjunto $\{1\}$ é um subconjunto de A , assim como $\{2\}$ e $\{3\}$. Mas não só isto, no sentido categorial, $\{4\}$ e $\{\clubsuit\}$ também são subconjuntos de A pois todos estes são isomorfos ao conjunto de um elemento $\{\bullet\}$. Tem-se, por exemplo:



A inclusão de $\{1\}$ em A e a função s são injetivas e portanto representam flechas mônicas. A função s simplesmente seleciona um objeto de A , $s(\clubsuit)$ e a inclusão de $\{1\}$ em A pode ser entendida como fazendo o mesmo, ela seleciona o elemento $1 \in A$. O isomorfismo i então é meramente uma tradução: $\clubsuit \mapsto 1$. A correspondência entre esta construção e o conceito de subconjuntos é feita através das seguintes observações:

- i Para conjuntos finitos, dois conjuntos são isomorfos se, somente se, possuem o mesmo número de elementos.
- ii O conjunto A possui quatro subconjuntos não isomorfos, \emptyset , $\mathbf{1} = \{\bullet\}$, $\mathbf{2} = \{\bullet, \bullet\}$ e $\mathbf{3} = \{\bullet, \bullet, \bullet\}$.
- iii O objeto A possui quatro subobjetos distintos, $\emptyset \rightarrow A$, $\mathbf{1} \rightarrow A$, $\mathbf{2} \rightarrow A$ e $\mathbf{3} \rightarrow A$.

Assim, existe uma correspondência bijetiva entre subobjetos e subconjuntos em Set . O mesmo argumento funciona para subgrupos e subespaços de uma espaço vetorial, notando agora que as flechas são homomorfismos de grupos e operadores lineares, respectivamente. A abstração em falar de classes de equivalência de subconjuntos isomorfos surge porque sem esta abstração não seríamos capazes de capturar a ideia de subconjunto com esta construção, por exemplo. Nem a de subgrupos e subespaços. Essencialmente, se falássemos de subobjetos apenas como flechas mônicas ao invés de classe de equivalência, acabaríamos com subobjetos demais. Por exemplo, o conjunto $A = \{1\}$ possui apenas dois subconjuntos, mas, possuiria infinitos subobjetos: $\{2\} \rightarrow A$, $\{\clubsuit\} \rightarrow A$, $\{1\} \rightarrow A$ e assim por diante. Assim, não teríamos uma bijeção entre subobjetos e subconjuntos em Set .

Os próximos conceitos necessários para falar de categorias abelianas são o de núcleo e conúcleo respectivamente. Começamos por definir o que é um núcleo e conúcleo da diferença:

Definição 4.12: Dados dois mapas $f, g : A \rightarrow B$ em uma categoria \mathbf{C} , dizemos que o par $\langle K, \kappa : K \rightarrow A \rangle$ é o núcleo da diferença de f e g se é o equalizador de f e g . Similarmente, um par $\langle F, \phi : B \rightarrow F \rangle$ é o conúcleo da diferença de f e g se é o coequalizador de f e g .

Primeiro notamos que esta definição não afirma que o núcleo ou conúcleo da diferença de quaisquer duas flechas de uma dada categoria \mathbf{C} existe. Porém, se existirem, sabemos que a flecha do núcleo é uma flecha mónica e a flecha do conúcleo, épica, em virtude de serem um equalizador e um coequalizador respectivamente. Isto nos permite relacionar estes dois conceitos com subobjetos e objetos quocientes. Pelas propriedades de equalizador, para qualquer flecha $h : O \rightarrow A$ que equaliza $f, g : A \rightarrow B$, tem-se

$$\begin{array}{ccccc}
 K & \xrightarrow{\kappa} & A & \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} & B \\
 \uparrow m & & \nearrow h & & \\
 O & & & &
 \end{array}$$

A existência da flecha mediadora m implica, em particular que, para qualquer h que equaliza f e g e também é mónica, $h \leq \kappa$. Assim, o núcleo da diferença de f e g representa o maior subobjeto que equaliza f e g . Similarmente, para qualquer flecha h' que coequaliza f e g e que também é épica, tem-se $h' \geq \phi$ de modo que o conúcleo da diferença é o menor objeto quociente que coequaliza f e g . Representamos estes subobjetos e objetos quocientes pela notação especial $[\kappa] = \text{Ker}(f - g)$ e $[\phi] = \text{Cok}(f - g)$. Note que $f - g$ é apenas uma notação conveniente, não é necessário que exista qualquer operação como soma definida para as flechas f e g .

Se a categoria \mathbf{C} possui um objeto zero, $\mathbf{0}$, então para quaisquer dois elementos A e B existe uma flecha única

$$A \xrightarrow{\mathbf{0}} B$$

chamada flecha zero de A em B dada pela composição

$$A \longrightarrow \mathbf{0} \longrightarrow B.$$

Usando a flecha zero, definimos o núcleo e o conúcleo de uma flecha f como:

Definição 4.13: Em uma categoria \mathbf{C} com um objeto zero, o núcleo de uma flecha $f : A \rightarrow B$ é o núcleo da diferença de f e $0 : A \rightarrow B$ e o conúcleo de f é o conúcleo da diferença de f e 0 .

Neste caso, denotamos o subobjeto representado pelo núcleo de f apenas por $\text{Ker}(f)$ e o objeto quociente representado pelo conúcleo por $\text{Cok}(f)$. Com estes conceitos definimos:

Definição 4.14: Uma categoria aditiva \mathbf{C} é dita **abeliana** se

- é **pré-abeliana**, isto é, se toda flecha f de \mathbf{C} possui um núcleo e um conúcleo;
- toda flecha mónica de \mathbf{C} é o núcleo de alguma outra flecha de \mathbf{C} e toda flecha épica é o conúcleo de alguma outra flecha de \mathbf{C} .

Categorias abelianas são surpreendentemente ricas de resultados [38, 45]. No que segue apresentaremos apenas os resultados relevantes para este trabalho.

4.3.3 - Categorias semissimples

Por fim, estamos em posição de definir categorias semissimples. Começamos por definir:

Definição 4.15: Um objeto S de uma categoria abeliana, \mathbf{C} é dito **simples** se $S \not\cong \mathbf{0}$ e toda flecha cujo alvo é S ou é uma flecha zero ou é inversível.

Esta definição corresponde a dizer que um objeto simples não possui nenhum subobjeto não trivial, ou seja, seus subobjetos são apenas $[\mathbf{0} \rightarrow S]$ e $[S \rightarrow S]$. De fato, suponha que S possuísse um subobjeto $[m]$, $m : X \rightarrow S$ tal que $m \leq [S \rightarrow S]$, então devemos ter que $m = s \circ u$ onde $u : X \rightarrow S$ porém, pela simplicidade de S , u ou é uma flecha zero ou é um isomorfismo. Se $u = 0$ então $m = 0$ de modo que $[m] = [\mathbf{0} \rightarrow S]$. Por outro lado, se u for um isomorfismo, então $X \cong S$ e $[m] = [S \rightarrow S]$. Com isto, definimos então uma categoria **semissimples**:

Definição 4.16: Uma categoria abeliana \mathbf{S} é dita **semissimples** se para todo objeto A de \mathbf{S} tem-se

$$A \cong \bigoplus_{j \in \mathbf{J}} N_j S_j$$

onde \mathbf{J} é a coleção das classes de isomorfismos de objetos simples de \mathbf{S} , S_j é um objeto simples e $N_j \in \mathbb{N}$ tal que $N_j \neq 0$ apenas para uma quantidade finito de classes de isomorfismos, j .

A semissimplicidade captura a ideia da fusão de anyons ser fatorada em uma soma de objetos irreduzíveis. Lembre-se que o spin é descrito conforme uma representação irreduzível de $\text{SU}(2)$ e quando combinamos duas partículas, realizamos o produto tensorial das representações para cada partícula e a representação resultante não é necessariamente

irredutível de modo que podemos escrevê-la como uma soma direta de representações irredutíveis. Agora, a definição de uma representação irredutível, I é uma representação tal que suas únicas sub-representações são a representação trivial 0 , onde todos os elementos do grupo atuam como a transformação identidade e a própria I . Isto corresponde exatamente a dizer que representações irredutíveis são objetos simples em uma categoria adequada.

Uma categoria semissimples, então é aquela na qual todo objeto ou é simples ou pode ser escrito como uma soma de objetos simples. Este é precisamente o comportamento que queríamos. Os tipos de anyons são os objetos simples de \mathfrak{A} . O fato de N_j poder ser diferente de zero para mais de um j indica que um anyon pode ser a soma de vários outros anyons possíveis. Quando integrarmos esta operação com a operação de fusão, \otimes , veremos que isto corresponde exatamente ao fato de que a fusão de dois anyons pode possuir mais de um resultado possível. Com isto enunciamos:

Axioma 5: A categoria \mathfrak{A} é semissimples.

O último passo agora é integrar as operações \otimes e \oplus para que possamos entender como a fusão de anyons pode ser escrita como uma soma de objetos simples e como esta soma se relaciona com a fatoração dos constituintes. Ao compatibilizar estas duas operações, estaremos acrescentando mais estrutura às coleções de morfismos de \mathfrak{A} transformando-as em espaços vetoriais que identificaremos como os espaços de Hilbert do sistema anyônico, aqueles que guardam os diferentes modos através dos quais podemos fundir dois anyons não-abelianos. Temos a seguinte definição:

Definição 4.17: Uma **categoria de fita semissimples**, \mathbf{C} é uma categoria semissimples equipada com uma estrutura de fita onde a unidade de \otimes , $\mathbf{1}$ é simples, o bifuntor \otimes é bilinear e onde, para cada objeto simples $S \in \mathbf{C}$, $[S, S] \cong \mathbb{K}$, um corpo de característica zero.

A característica de um corpo é o número de vezes que é possível somar-se a identidade, $1 \in \mathbb{K}$ de modo a obter zero, isto é, o número de fatores na soma

$$1 + 1 + \cdots + 1 = 0.$$

Diz que um corpo possui característica zero se esta soma nunca é igual a zero. O corpo dos reais e o dos complexos possui característica zero. O corpo de dois elementos, $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ possui característica igual a dois pois nele tem-se que $1 + 1 = 0$.

Por fim, note que $[S, S]$ em um espaço vetorial sobre \mathbb{K} , pois todo corpo é, trivialmente, um espaço vetorial sobre si mesmo. Mais explicitamente, o bifuntor \otimes oriundo da estrutura monoidal de \mathbf{C} define a multiplicação em $[S, S]$, pois para flechas $f, g \in [S, S]$, podemos

formar as flecha $f \otimes g$ e $g \otimes f$, também de $[S, S]$. Note que estas duas flechas podem ser relacionadas pelo isomorfismo σ . Além disto, é possível mostrar [45] que em uma categoria de fita semissimples, tem-se:

Proposição 4.4: Se \mathbf{C} é uma categoria de fita semissimples existem isomorfismos naturais tais que

$$A \otimes (B \oplus C) \cong (A \otimes B) \oplus (A \otimes C) \quad \text{e} \quad (A \oplus B) \otimes C \cong (A \otimes C) \oplus (B \otimes C)$$

e

$$(A \oplus B)^* \cong A^* \oplus B^* \quad \text{e} \quad \mathbf{1}^* \cong \mathbf{1}.$$

Com isto, temos o suficiente para escrever as regras de fusão pois como \mathbf{C} é semissimples, todo objeto é escrito como uma soma de objetos simples, assim, usando a distributividade de \otimes , temos:

$$A \otimes B \cong A \otimes \left(\bigoplus_k N_k S_k \right) \cong \bigoplus_k N_k (A \otimes S_k) \cong \bigotimes_{j,k} N_j N_k (S_j \otimes S_k) \cong \bigotimes_i N_{j,k}^i S_i.$$

A passagem na última igualdade é explicada da seguinte forma: conforme os índices i, j variam, o produto $S_j \otimes S_k$ resulta em um novo elemento que por sua vez pode ser fatorado numa soma de elementos simples. Em geral, estes elementos possuirão termos em comum em sua fatoração. Os coeficientes $N_{j,k}^i$ contam quantas vezes estes termos aparecem e assim, podemos escrever o produto $A \otimes B$ em termos da soma mostrada. Essencialmente isto significa que o objeto S_i pode aparecer na fatoração de vários produtos $S_j \otimes S_k$ distintos, em outras palavras, isto significa que há múltiplas maneiras de se “obter” o objeto S_i . Este comportamento é exatamente o comportamento observado na fusão de anyons e o que queríamos formalizar, assim, enunciamos o penúltimo axioma:

Axioma 6: A categoria \mathfrak{A} é uma categoria de fita semissimples. Além disto, para todo objeto $\mathcal{A} \in \mathfrak{A}$, tem-se que $[\mathcal{A}, \mathcal{A}] \cong \mathbf{C}$.

4.4 - Modularidade

O último axioma necessário para completar a descrição categorial de anyons corresponde a dizer que apenas certas categorias de fita semissimples correspondem a modelos anyônicos. Tais categorias são chamadas de **categorias tensoriais modulares**. Vejamos a definição:

Definição 4.18: Uma categoria, \mathcal{C} , é dita **tensorial modular** se é uma categoria de fita semissimples tal que:

- **Finitude:** o conjunto de índices \mathbf{J} de classes de isomorfismos de objetos semissimples é finito. Por convenção adotamos $S_0 = \mathbf{1}$.
- **Modularidade:** para $i, j \in \mathbf{J}$ a matriz

$$(s)_{ij} = [\lambda_{\mathbf{1}} \circ (e_{S_i} \otimes e_{S_j}) \circ (\text{id}_{S_i} \otimes \sigma_{S_j, S_i^*} \otimes \text{id}_{S_j^*}) \circ (\text{id}_{S_i} \otimes \sigma_{S_i^*, S_j} \otimes \text{id}_{S_j^*}) \circ (i_{S_i} \otimes i_{S_j}) \circ \lambda_{\mathbf{1}}^{-1}]_{ij}$$

é inversível.

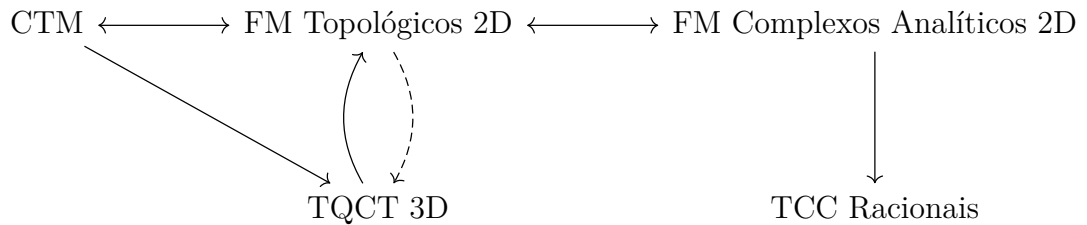
A definição de categorias tensoriais modulares apesar de não muito complexa é um tanto difícil de justificar e pode parecer bastante artificial. Começamos pela condição de finitude, ela afirma que existe apenas uma quantidade finita de classes de isomorfismos de objetos simples na categoria. Em termos de anyons isto implica que existe uma quantidade finita de tipos simples de anyons em um dado modelo. Panangaden e Paquette [40] afirmam que, a princípio, não há motivo para crer que isto é verdade na natureza, no entanto, computação quântica — talvez o principal motivador do estudo de anyons — pode ser feita apenas com um número finito e esta condição nos garante alguns resultados convenientes. Por exemplo, com finitude as categorias descrevendo anyons tornam-se equivalentes a categorias pequenas e isto nos permite usar o lema de Yoneda, como faremos a seguir, e obter uma simplificação muito útil para trabalhar com estes modelos.

Particularmente, acreditamos também que apesar de não ser estritamente proibido, um número infinito de classes de isomorfismos de objetos simples distintos parece irrazoável. Se retomarmos a analogia com momento angular, significaria uma partícula composta que pode carregar valores de spin infinitamente altos. Porém, experiência mostra que o mais comum é encontrar partículas com valores de spin 0, 1/2 e 1 na natureza, especialmente se forem partículas fundamentais, papel que seria ocupado por anyons em duas dimensões. Por estes motivos, então adotamos a finitude para categorias descrevendo anyons. A condição de modularidade é mais complicada.

Na definição da matriz acima, fizemos uso, implicitamente, do fato de que $A^{**} \cong A$ em uma categoria de fita para abreviar um pouco a definição da matriz s . A flecha que define a matriz corresponde ao processo de criar dois pares anyon anti-anyon, trançar os anyons S_i^* e S_j duas vezes segundo σ e então aniquilá-los. Esta matriz corresponde ao que é chamado de **enlace de Hopf**, *Hopf link*. Ao que tudo indica [40, 42], a exigência de modularidade surge muito mais por implicações algébricas do que por motivações físicas. Isto não significa que não existam motivações físicas, afinal, como dissemos no início do capítulo 3, as teorias de campo que descrevem anyons são equivalente a categorias tensoriais modulares, não categorias de fita semissimples.

Explorar a fundo esta equivalência para justificar a exigência de modularidade está além do escopo do trabalho. Mesmo estudar suas consequências algébricas exigiria a introdução de vários outros conceitos que dependem de um conhecimento muito mais aprofundado sobre teoria de categorias e suas relações com outras estruturas algébricas complexas como anéis de fusão e álgebras de Verlinde, dois conceitos particularmente relacionados com a descrição de anyons por meio de teoria de campos conformes. Não trataremos destes assuntos aqui, porém, apresentamos suas principais consequências segundo o trabalho de Panangaden e Paquette [40]:

- Como dito, a estrutura tensorial modular garante a equivalência da descrição categórica de anyons como mostrada no diagrama do capítulo 3. Por conveniência, o repetimos:



onde *CTM* representa categorias tensoriais modulares, *FM* funtores monoidais, *TQCT*, teoria quântica de campos topológica e *TCC* teoria de campo conformes.

- Através de um teorema chamado **fórmula de Verlinde**, estabelece-se que os coeficientes de fusão $N_{i,j}^k$ podem ser completamente determinados por meio da matriz s através de outra matriz chamada de matriz S .
- Especificar completamente a estrutura de fita semissimples de uma categoria de anyons, isto é, seus objetos simples, as identidades hexagonais e pentagonal a torção e regras de fusão e também a matriz S determina completamente as propriedades topológicas de anyons.

Em particular, o último ponto nos diz que uma vez estabelecido que anyons podem ser descritos por categorias tensoriais modulares, se inventarmos um conjunto de tipos de anyons e regras de fusão entre estes anyons de modo a satisfazer as condições exigidas acima, podemos, teoricamente, transitar entre estas equivalências e obter uma teoria de campos topológica para este modelo de anyons.

Em particular esta possibilidade foi a principal motivadora destes trabalho e discutiremos isto no capítulo final. E note que especificar estas estruturas corresponde essencialmente a falar que esta lista de partículas de fato são anyons. A lista e as regras de fusão a princípio são os elementos fundamentais. As identidades hexagonais e pentagonal correspondem a dizer que estas partículas evoluem segundo o grupo de tranças sob permutações. A torção, como vimos captura o grau de liberdade extra gerado pela fusão de

anyons e a matriz S é essencialmente determinada por todas estas coisas [40]. Em outras palavras, temos uma espécie de receita para criar modelos anyônicos. Finalizamos a seção sobre modularidade aqui e guardamos as discussões sobre esta receita para o capítulo final. Passemos agora a um tópico um pouco mais concreto.

4.5 – Simplificações

A descrição categorial de anyons está, neste ponto, concluída. A partir do que apresentamos até agora é possível descrever sistemas anyônicos, significando que, especificadas as regras os tipos de anyons e suas regras de fusão, podemos obter, por meio das restrições impostas pelas identidades hexagonais e pentagonais e da torção na categoria, como estes anyons evoluem sob quaisquer permutações e torções — rotações em torno de si mesmos. Especificamente, podemos calcular as transformações às quais a função de onda está sujeita conforme realizamos estas operações. Apesar disto, é possível simplificar toda a construção que apresentamos até agora, em particular, colocando-a em termos de estruturas mais familiares: espaços vetoriais.

Esta simplificação é obtida através do **lema de Yoneda**. Mencionamos brevemente o lema de Yoneda na seção 3.4, onde dissemos que ele é um dos resultados mais importantes de teoria de categorias básica. Essencialmente, o lema de Yoneda nos diz que objetos de uma dada categoria são determinados, a menos de isomorfismos, pelas flechas que tem como alvo este objeto. Note que isto está completamente de acordo com a filosofia geral de teoria de categorias de que as flechas são o que realmente importam. Determinar um objeto desta maneira essencialmente corresponde a determinar o objeto não por sua estrutura interna, se ele possuir alguma, mas por sua posição em relação a outros objetos. Objetos finais e iniciais, são exemplos de como objetos podem ser determinados por suas flechas. Denotando por H^A o funtor $\text{hom}[A, -] : \mathbf{C} \rightarrow \mathbf{Set}$, temos, formalmente:

Lema 4.1 – Yoneda: Se \mathbf{C} é uma categoria localmente pequena, dado um funtor $F : \mathbf{C} \rightarrow \mathbf{Set}$ e um objeto A de \mathbf{C} , existe uma bijeção entre as transformações naturais $H^A \rightarrow F$ e os objetos $F(A)$. Isto é

$$\text{Nat}(H^A, F) \cong F(A).$$

onde $\text{Nat}(H^A, F)$ é a coleção de transformações naturais de H^A e F .

Provas do lema de Yoneda podem ser encontradas nas referências [36, 38]. As consequências do lema são vastas. Particularmente útil é um corolário do lema quando o funtor F é também um funtor hom , H^B para outro objeto B de \mathbf{C} . Neste caso, o lema nos diz que as transformações naturais de H^A para H^B possuem uma correspondência bi-

jetiva com as flechas $f : B \rightarrow A$. Esta correspondência dá origem a um funtor $\mathbf{C} \rightarrow \mathbf{Set}^{\mathbf{C}^{\text{op}}}$ chamado **imersão de Yoneda** e nos informa que a categoria \mathbf{C} é “equivalente” à categoria de funtores contravariantes de \mathbf{C} a \mathbf{Set} . Não exploraremos rigorosamente isto, mas “equivalente” aqui significa que \mathbf{C} é uma subcategoria plena de $\mathbf{Set}^{\mathbf{C}^{\text{op}}}$. O efeito prático, para nós é que cada objeto A de \mathbf{C} é equivalente, sob a imersão de Yoneda, a um funtor $\text{hom}, [-, A]$.

Agora, suponha por um momento que uma categoria \mathfrak{A} , representando algum modelo de anyons seja localmente pequena. Então, a cada tipo de anyon \mathcal{A} , equivale o funtor $[-, \mathcal{A}]$ que envia cada outro tipo de anyon, \mathcal{B} no conjunto das flechas de $\mathcal{B} \rightarrow \mathcal{A}$, $[\mathcal{B}, \mathcal{A}]$ e todas as flechas $f : \mathcal{B} \rightarrow \mathcal{C}$ em flechas de $[\mathcal{C}, \mathcal{A}] \rightarrow [\mathcal{B}, \mathcal{A}]$ pela operação $- \circ f$ de composição com f à direita. Como \mathfrak{A} é semissimples, podemos simplificar esta descrição lembrando que todo objeto $\mathcal{A} \in \mathfrak{A}$ é escrito como uma soma

$$\mathcal{A} \cong \bigoplus_{j \in \mathbf{J}} N_j S_j$$

de objetos simples, S_j . Assim, cada flecha $f : \mathcal{A} \rightarrow \mathcal{B}$ pode ser decomposta em uma família de flechas

$$\left(f_j = f \circ i_j : S_j \rightarrow \mathcal{B} \right)_{j \in \mathbf{J}}$$

indexada por \mathbf{J} , a coleção de classes de isomorfismos de objetos simples de \mathfrak{A} . Alternativamente, toda família de flechas $g_j : S_j \rightarrow \mathcal{B}$ produz um única flecha: $g : \mathcal{A} \rightarrow \mathcal{B}$ para algum $\mathcal{A} \in \mathfrak{A}$. Note que como a decomposição em objetos simples pode conter mais de uma instância de um mesmo objeto, N_j , a família de flechas pode conter mais de uma flecha de um dado objeto S_j para \mathcal{B} [43].

Este fato simplifica grandemente a categoria $\mathbf{Set}^{\mathfrak{A}^{\text{op}}}$ pois todos os funtores hom podem ser escritos a partir de funtores hom dos objetos simples $[-, S_j]$ onde $j \in \mathbf{J}$. Denotando por \mathfrak{S} — a letra S no estilo fraktur — a subcategoria dos objetos simples de \mathfrak{A} , podemos dizer que \mathfrak{A} é equivalente a $\mathbf{Set}^{\mathfrak{S}^{\text{op}}}$ [42]. Note, uma subcategoria de uma categoria \mathbf{C} é uma categoria cujos objetos são uma subcoleção dos objetos de \mathbf{C} e cujas flechas são uma subcoleção das flechas de \mathbf{C} . Além disso, note que estamos preocupados apenas com as classes de isomorfismos de objetos simples, não os objetos simples em si. Deste modo, podemos substituir a categoria \mathfrak{S} por um esqueleto dela, \mathfrak{S}_0 cuja coleção de objetos consiste em precisamente um representante de cada classe de isomorfismos de objetos simples. A condição de finitude de \mathfrak{A} , garante então que \mathfrak{S}_0 é pequena e portanto, localmente pequena, pois sua coleção de objetos forma um conjunto. Assim, mesmo se \mathfrak{A} não for localmente pequena, \mathfrak{A} é equivalente a \mathfrak{S}_0 , a simplificação via Yoneda descrita vale e podemos identificar \mathfrak{A} com $\mathbf{Set}^{\mathfrak{S}_0^{\text{op}}}$.

Devido aos axiomas que impusemos sobre \mathfrak{A} , principalmente devido à semissimplicidade, em \mathfrak{S}_0 temos que $[S_j, S_j] \cong \mathbb{C}$ e $[S_j, S_k] \cong \{0\}$, $j \neq k$. Assim, cada objeto \mathcal{A} de \mathfrak{A}

corresponde, segundo a identificação de \mathfrak{A} com \mathfrak{S}_0 e a imersão de Yoneda, a uma família finita de espaços vetoriais

$$(V_S = [S, \mathcal{A}])_{S \in \text{Ob}(\mathfrak{S}_0)}$$

indexada pelos objetos, simples, de \mathfrak{S}_0 . Além disso, cada flecha $\mathcal{A} \rightarrow \mathcal{B}$ são identificados com famílias arbitrárias de mapas lineares

$$(g_S : [S, \mathcal{A}] \rightarrow [S, \mathcal{B}])_{S \in \text{Ob}(\mathfrak{S}_0)}.$$

O motivo da arbitrariedade na escolha da família de mapas lineares é que as restrições sobre as flechas em \mathfrak{S}_0 são suficientes para garantir que qualquer tal família de mapas entre os espaços vetoriais satisfaça as condições de comutação para que sejam transformações naturais e, portanto, flechas em $\text{Set}^{\mathfrak{S}_0^{\text{op}}}$. Lembre-se, como \mathfrak{S}_0 é composta exclusivamente por objetos simples, todas as flechas $f : S_i \rightarrow S_j$ são flechas zero, já que \mathfrak{S}_0 possui exatamente um objeto simples de cada classe de isomorfismos de objetos simples de \mathfrak{A} . Esta é uma restrição bastante severa.

Por fim, vejamos o que estas identificações fazem às estruturas aditiva, \oplus e multiplicativas, \otimes de \mathfrak{A} . A estrutura aditiva é trivial sob a identificação, ela se torna a soma direta de espaços vetoriais. Isto se dá, basicamente porque uma flecha

$$f : S_j \oplus S_k \rightarrow \mathcal{A}$$

pode ser decomposta em flechas $f_j = f \circ i_j : S_j \rightarrow \mathcal{A}$, de modo que todos os elementos $f \in [S_j \oplus S_k, \mathcal{A}]$ são escritos como

$$f = f_j \oplus f_k$$

ou ainda, como pares (f_j, f_k) , onde $f_j \in [S_j, \mathcal{A}]$ e $f_k \in [S_k, \mathcal{A}]$. Esta é precisamente a definição da soma direto de espaços vetoriais de modo que

$$[S_j \oplus S_k, \mathcal{A}] = [S_j, \mathcal{A}] \oplus [S_k, \mathcal{A}].$$

Os espaços vetoriais de uma soma arbitrária de elementos de \mathfrak{S}_0 são obtidos por iteração.

A estrutura multiplicativa é mais interessante e complexa. Em particular, as flechas associadoras α e as tranças σ , apesar de serem dadas por mapas lineares sob a identificação, não necessariamente possuem uma estrutura simples. A análise da estrutura multiplicativa, também pode ser simplificada usando da semissimplicidade de \mathfrak{A} e também do fato que \otimes é distributiva sobre \oplus . Se soubermos como \otimes atua sobre objetos simples, saberemos como atua em somas destes objetos e, pela semissimplicidade, sobre todos os objetos. Além disto, por α e σ serem naturais em \mathfrak{A} elas comutam com as flechas projeção e inclusão do biproduto [42]. Assim, o comportamento destes isomorfismos em

objetos arbitrários pode ser obtido a partir de seu comportamento para objetos simples. Mais que isto, as identidades hexagonais e pentagonal valem em geral se valerem para todos os objetos simples.

Em resumo, então, as estruturas aditivas e multiplicativas de \mathfrak{A} podem ser descritas completamente fornecendo-se:

- uma lista completa de objetos simples não isomorfos, incluindo a identidade, $\mathbf{1}$;
- para cada par de objetos desta lista, seu produto \otimes expresso em termos da soma \oplus ;
- as flechas associadoras $\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}}$ para todos os objetos \mathcal{A}, \mathcal{B} e \mathcal{C} da lista e
- as tranças $\sigma_{\mathcal{A},\mathcal{B}}$ para todos os objetos \mathcal{A} e \mathcal{B} da lista

satisfazendo as identidades hexagonais e pentagonal. Um processo de simplificação similar pode ser feito com a estrutura de fita e a estrutura rígida de \mathfrak{A} , porém o essencial está nas estruturas aditivas e multiplicativas, como vimos. Notamos que os dois primeiros pontos em geral restringem grandemente os dois últimos, quando não os determinam completamente, em virtude das identidades hexagonais e pentagonal. Com isto, finalizamos a exposição sobre a descrição categorial de anyons. Na seção seguinte, usaremos o que foi exposto para obter uma especificação completa de anyons de Fibonacci que mencionamos brevemente no início deste capítulo.

Por último mas não menos importante, a conexão com espaços de Hilbert é obtida através da estrutura rígida. O functor $(-)^*$ corresponde precisamente à definição da transposição. Isto é, para cada flecha $f : \mathcal{A} \rightarrow \mathcal{B}$, este functor atribui a flecha $f^* : \mathcal{B}^* \rightarrow \mathcal{A}^*$, sua adjunta. Além disto, considere a flecha $e_{\mathcal{A}} : \mathcal{A}^* \otimes \mathcal{A} \rightarrow \mathbf{1}$. Esta flecha corresponde, em FDVec_K ao mapa avaliação $ev_V : V^* \otimes V \rightarrow K$:

$$ev_V(\alpha \otimes v) = \langle \alpha, v \rangle = \alpha(v).$$

Com este mapa, é possível definir um produto interno em V , transformando-o em um espaço de Hilbert, já que possui dimensão finita. A situação é similar em \mathfrak{A} , apesar do produto \otimes ser mais complicado que o produto tensorial de espaços vetoriais, como discutimos.

A questão da estrutura de Hilbert para os conjuntos $[S_i, S_i]$, onde $S_i \in \mathfrak{A}$ é simples, é mais complexa do que aparenta. Os trabalhos de Panangaden e Paquette e Blass e Gurevich [40, 42] não se aprofundam neste ponto, efetivamente parando ao estabelecer que $[S_i, S_i] \cong \mathbb{C}$. Por outro lado, se tentarmos fornecer a estrutura de espaço de Hilbert para estes conjuntos desde o início, chegamos a uma definição distinta de categoria tensorial modular [46] e cuja equivalência com a definição apresentada aqui não pudemos atestar. Acreditamos que esta discussão é necessária porque, apesar de termos postulado que $[S_i, S_i] \cong \mathbb{C}$ para qualquer objeto simples, S_i , e \mathbb{C} ser um espaço de Hilbert, não discutimos

em nenhum momento como o produto interno surge em $[S_i, S_i]$. Exploraremos um pouco mais desta questão no exemplo da próxima seção.

4.6 - Exemplo: Anyons de Fibonacci

Anyons de Fibonacci é um modelo de anyons bastante simples e bastante rico. Para os propósitos deste trabalho, consideramos anyons de Fibonacci como um “modelo de brinquedo” pois sua simplicidade o torna um ótimo modelo para usar como exemplo e a riqueza ainda nos permite ver como anyons podem ser usados para computação quântica. Apesar disto, anyons de Fibonacci são os, ou pelo menos um dos, anyons que mais se buscam para a construção de computadores quânticos topológicos por serem completamente capazes de sustentar computações. Veremos o que isto significa ao final da seção. Note que o fato de os tratarmos como um modelo de brinquedo não significa que não há prospectos de obtenção experimental destes anyons, suspeita-se que possam existir sob certas condições no efeito Hall quântico fracionário, especificamente com o fator de preenchimento $\nu = 12/5$ e existem algumas hamiltonianas propostas para sua obtenção [47].

Seguindo o que foi discutido na seção anterior, começamos por especificar os anyons simples do modelo. A simplicidade do modelo de anyons de Fibonacci é que possuem apenas dois tipos de anyons simples, um dos quais é o anyon trivial, o vácuo. Denotaremos a categoria representando os anyons de Fibonacci por \mathfrak{F} — a letra F em estilo fraktur — e os dois tipos de anyons simples por \mathcal{T} e $\mathbf{1}$, onde $\mathbf{1}$ é o anyon trivial. As regras de fusão do modelo são:

$$\begin{aligned}\mathcal{T} \otimes \mathcal{T} &= \mathcal{T} \oplus \mathbf{1}; \\ \mathcal{T} \otimes \mathbf{1} &= \mathbf{1} \otimes \mathcal{T} = \mathcal{T}; \\ \mathbf{1} \otimes \mathbf{1} &= \mathbf{1}.\end{aligned}$$

Estes dois tipos de anyons são seus próprios duais. A partir das regras de fusão e usando o fato que o produto \otimes é distributivo sobre \oplus pelos axiomas, podemos ver de onde surge o nome “anyons de Fibonacci”. No que segue, representaremos uma instância específica de uma anyon de tipo \mathcal{T} por τ e de $\mathbf{1}$ por 1 . Considere, então, a fusão de três anyons τ :

$$\tau \otimes (\tau \otimes \tau) = \tau \otimes (\tau \oplus 1) = (\tau \otimes \tau) \oplus (\tau \otimes 1) = (\tau \oplus 1) \oplus \tau = 2 \cdot \tau + 1.$$

Onde a notação $k \cdot \sigma$ indica a soma de k cópias de um dado anyon σ . Para quatro anyons

τ , temos:

$$\begin{aligned}\tau \otimes (\tau \otimes (\tau \otimes \tau)) &= \tau \otimes (2\tau \oplus 1) = 2 \cdot (\tau \otimes \tau) \oplus \tau = 2 \cdot (\tau \oplus 1) \oplus \tau \\ &= 2 \cdot \tau \oplus 2 \cdot 1 \oplus \tau = 3 \cdot \tau \oplus 2 \cdot 1\end{aligned}$$

Se continuarmos com isto veremos que a fusão de n anyons τ fornece:

$$\tau^{\otimes n} = f_n \cdot \tau \oplus f_{n-1} \cdot 1 \quad (4.2)$$

onde f_n são precisamente os elementos da sequência de Fibonacci, definida por

$$f_n = \begin{cases} 0 & \text{se } n = 0; \\ 1 & \text{se } n = 1; \\ f_{n-1} + f_{n-2} & \text{se } n > 1. \end{cases}$$

Como discutido anteriormente, \mathfrak{F} pode ser identificada com \mathfrak{S}_0 , seu esqueleto contendo apenas um objeto simples de cada classe de isomorfismo de objetos simples, que por sua vez pode ser identificada, pela imersão de Yoneda com $\mathbf{Set}^{\mathfrak{S}_0^{\text{op}}}$. Assim, cada objeto \mathcal{A} , não necessariamente simples, de \mathfrak{F} é identificado com uma família de espaços vetoriais

$$(V_S = [S, \mathcal{A}])_{S \in \text{Ob}(\mathfrak{S}_0)}.$$

Como \mathfrak{F} só possui duas classes de isomorfismos de objetos simples, aquelas representadas por \mathcal{T} e $\mathbf{1}$, cada anyons de tipo \mathcal{A} de \mathfrak{F} pode ser expresso como um par de espaços vetoriais

$$\mathcal{A} \mapsto (V_{\mathcal{T}} = [\mathcal{T}, \mathcal{A}], V_{\mathbf{1}} = [\mathbf{1}, \mathcal{A}]).$$

Sob esta identificação, a operação \oplus corresponde à soma direta de cada uma dos espaços vetoriais separadamente. Isto é, se temos dois tipos de anyons \mathcal{A} e \mathcal{B} , temos:

$$\mathcal{A} \oplus \mathcal{B} \mapsto ([\mathcal{T}, \mathcal{A}] \oplus [\mathcal{T}, \mathcal{B}], [\mathbf{1}, \mathcal{A}] \oplus [\mathbf{1}, \mathcal{B}])$$

Em particular, temos, devido aos axiomas, que

$$\mathcal{T} \mapsto (\mathbb{C}, 0) \quad \text{e} \quad \mathbf{1} \mapsto (0, \mathbb{C})$$

onde 0 é o espaço vetorial trivial: $0 = \{0\}$.

A estrutura aditiva de \mathfrak{F} é bastante simples e a discussão acima essencialmente a determina. A expressão

$$\tau^{\otimes n} = f_n \cdot \tau \oplus f_{n-1} \cdot 1$$

determina a estrutura multiplicativa de \mathfrak{F} no que concerne os objetos. Porém ainda devemos determinar como \otimes atua sobre flechas de \mathfrak{F} . Além disto, chamamos atenção para o fato de que \otimes não é simplesmente o produto tensorial de espaços pois se fosse, teríamos:

$$\tau \otimes \tau = (\mathbb{C} \otimes \mathbb{C}, 0) \cong (\mathbb{C}, 0) = \tau \neq \tau \oplus 1 = (\mathbb{C}, \mathbb{C}).$$

Assim como os objetos $\mathcal{A} \in \mathfrak{F}$ são identificados com pares de espaços vetoriais $(V_{\mathcal{T}}, V_1)$, as flechas de \mathfrak{F} , $f : \mathcal{A} \rightarrow \mathcal{B}$, são identificadas com pares de mapas lineares. Se $\mathcal{A} = (V_{\mathcal{T}}, V_1)$ e $\mathcal{B} = (W_{\mathcal{T}}, W_1)$, temos:

$$(f : \mathcal{A} \rightarrow \mathcal{B}) \mapsto (f_{\mathcal{T}} : V_{\mathcal{T}} \rightarrow W_{\mathcal{T}}, f_1 : V_1 \rightarrow W_1)$$

Podemos pensar estes mapas como pares de matrizes se fixarmos bases para os espaços vetoriais envolvidos. A escolha de base envolve um grau de arbitrariedade considerável. Um princípio útil, no entanto é: uma vez escolhidas bases para dois espaços usaremos a união destas bases produzir uma base para o produto direto destes espaços. Chamamos atenção, no entanto, para o fato de que um mesmo espaço vetorial pode ser obtido como a soma direta de vários espaços distintos. Assim, seguindo este princípio, obteremos várias bases distintas para os mesmos espaços. Devemos, então, encontrar um modo de relacionar estas diferentes bases.

Este princípio não nos diz nada sob bases para os anyons $\tau = (\mathbb{C}, 0)$ e $1 = (0, \mathbb{C})$. Como estes objetos são simples, não existe nem mesmo uma flecha não zero entre eles que possa nos informar como relacionar suas bases. Assim, identificamos estes dois espaços com \mathbb{C} e tomamos o número 1 como o vetor base de ambos. Notamos que a identificação com \mathbb{C} pode ser feita pois, os elementos de $(\mathbb{C}, 0)$ são $(c, 0)$ para algum número complexo c . Como 0 só possui um elemento, existe uma bijeção $(\mathbb{C}, 0) \rightarrow \mathbb{C}$. O mesmo ocorre para $(0, \mathbb{C})$. Notamos que podemos fornecer aos pares (V_{τ}, V_1) a estrutura de espaços vetoriais simplesmente interpretando-os como o produto de dois espaços. Além disso, 1 é a base de \mathbb{C} pensando \mathbb{C} como um espaço um espaço vetorial sobre \mathbb{C} , assim qualquer complexo c pode ser obtido por $c = 1c$. Se o como um espaço vetorial sobre \mathbb{R} , a base seria $\{1, i\}$.

Com esta escolha de base, automaticamente obtemos uma base para

$$\tau \otimes \tau = \tau \oplus 1 = (\mathbb{C}, \mathbb{C}).$$

Passamos então para o produto triplo:

$$\tau \otimes (\tau \otimes \tau) = 2 \cdot \tau \oplus 1 = (\mathbb{C} \oplus \mathbb{C}, \mathbb{C}) \cong (\mathbb{C}^2, \mathbb{C}).$$

Diferentemente do caso $\tau \otimes \tau$, o caso com três anyons carrega informação adicional, em

particular, que o par $(\mathbb{C}^2, \mathbb{C})$ foi obtido a partir da soma

$$(\tau \otimes \tau) \oplus (\tau \otimes 1) = (\mathbb{C}, \mathbb{C}) \oplus (\mathbb{C}, 0).$$

O nosso princípio, então, diz que deveríamos construir a base de \mathbb{C}^2 de modo a respeitar esta decomposição. Assim, um elemento da base deveria ser oriundo de $\tau \otimes \tau$ e outro de $\tau \otimes 1$. Agora, note que poderíamos escolher a outra ordem de parênteses:

$$(\tau \otimes \tau) \otimes \tau = (\tau \oplus 1) \otimes \tau = (\tau \otimes \tau) \oplus (1 \otimes \tau) = (\mathbb{C}, \mathbb{C}) \oplus (\mathbb{C}, 0).$$

Esta escolha também leva ao par $(\mathbb{C}^2, \mathbb{C})$ e sugere uma base para \mathbb{C}^2 , porém não há garantia de que a base sugerida por esta escolha de parêntese seja igual à base sugerida pela escolha anterior. Caso não esteja visível que possa haver diferença entre estas duas bases, lembre-se que em \mathfrak{F} exigimos, ao invés da igualdade estrita, que:

$$\mathcal{T} \otimes (\mathcal{T} \otimes \mathcal{T}) \cong (\mathcal{T} \otimes \mathcal{T}) \otimes \mathcal{T}$$

Note que existe um isomorfismo não trivial conectando estes dois objetos, a flecha associador $\alpha_{\mathcal{T}, \mathcal{T}, \mathcal{T}}$ e é exatamente ela que estabelece a relação entre as duas bases de $\tau^{\otimes 3}$. Aqui, a simplicidade dos anyons de Fibonacci se torna aparente. Como mencionamos na seção anterior, precisamos determinar α apenas para os objetos simples de \mathfrak{F} . Uma vez feito isto, ela estará determinada para todos os outros objetos. Além disso, em vista da identificação $\mathcal{T} \otimes \mathbf{1} \cong \mathcal{T} \cong \mathbf{1} \otimes \mathcal{T}$ devido às flechas unitoras, λ e ρ , precisamos determinar α apenas para os três índices iguais a \mathcal{T} . Se um ou mais dos índices forem $\mathbf{1}$, α torna-se alguma composição da identidade com ρ ou λ . Além disso, sob a identificação anterior, $\alpha_{\mathcal{T}, \mathcal{T}, \mathcal{T}}$ corresponde a dois mapas, um de $\mathbb{C}^2 \rightarrow \mathbb{C}^2$ e outro de $\mathbb{C} \rightarrow \mathbb{C}$. Ou seja, a uma matriz dois por dois e um número. Estes mapas estão sujeitos à identidade pentagonal que os determina quase completamente. Uma situação similar acontece com as tranças σ de \mathfrak{F} . Precisamos determinar apenas $\sigma_{\mathcal{T}, \mathcal{T}}$ e usaremos a identidade hexagonal para determiná-la.

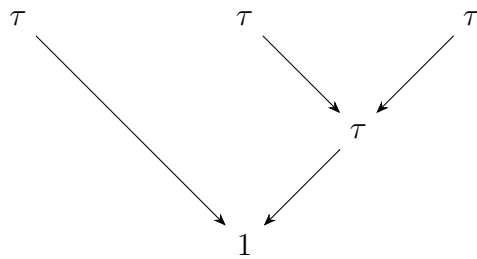
Para computar os isomorfismos $\alpha_{\mathcal{T}, \mathcal{T}, \mathcal{T}}$ e $\sigma_{\mathcal{T}, \mathcal{T}}$, os interpretaremos como matrizes usando bases adequadas dos espaços vetoriais em questão e então calcularemos as restrições impostas sobre estas matrizes pelas identidades hexagonal e pentagonal. Para tanto, é útil estabelecer uma notação mais conveniente para as bases. Ao fazer isto, discutiremos também o que estas bases representam do ponto de vista físico e elucidaremos um pouco mais sobre o significado de “ para anyons, os espaços de Hilbert representam modos distintos de fundir anyons de modo a obter uma dada configuração”.

Considere novamente a expressão

$$\tau \otimes (\tau \otimes \tau) = 2 \cdot \tau \oplus 1.$$

Olhando para esta expressão do ponto de vista físico, ela nos diz se começarmos com três anyons τ e os fundirmos progressivamente, o resultado da fusão não é único. Para ver isto, considere os parênteses. Primeiro, fundimos dois anyons τ , o resultado, pode ser outro anyon τ ou pode ser 1. Se o resultado for 1, então a fusão com o terceiro τ resulta em τ . Se o resultado for τ , então a fusão com o terceiro pode resultar em τ ou em 1. Ou seja, se realizarmos este mesmo processo, é possível que obtenhamos resultados distintos. Por outro lado, se fixarmos o resultado final, o que equivale a afirmar que a carga topológica do sistema é dada, ainda podem existir múltiplos caminhos de fusão. Por exemplo, se fixarmos o resultado com τ , ainda existem dois modos possíveis de se obter este resultado. O $2 \cdot \tau$ na expressão acima indica exatamente isto, que existem dois modos distintos de se obter τ nesta fusão, exatamente como discutimos na construção heurística inicial.

O raciocínio acima pode ser representado de modo simples por uma árvore. Por exemplo, se fixarmos o resultado final $\tau \otimes (\tau \otimes \tau) = 1$, existe apenas um caminho de fusão possível e podemos representá-lo da seguinte forma:



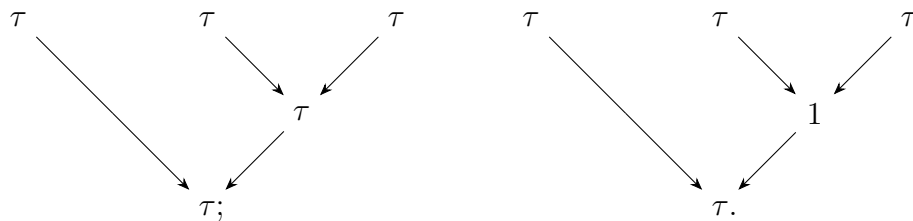
ou, em notação mais compacta:

$$(\tau \underset{1}{\cdot} (\tau \underset{\tau}{\cdot} \tau))$$

Nesta notação, os três τ 's e os parênteses descrevem o produto $\tau \otimes (\tau \otimes \tau)$ e o símbolo embaixo de \cdot indica qual dos resultados possíveis escolhemos. Assim, $\tau \underset{\tau}{\cdot} \tau$ indica que escolhemos o resultado τ dentre as possibilidades τ e 1. Os outros caminhos de fusão possíveis são

$$(\tau \underset{\tau}{\cdot} (\tau \underset{\tau}{\cdot} \tau)) \quad \text{e} \quad (\tau \underset{\tau}{\cdot} (\tau \underset{1}{\cdot} \tau)).$$

As árvores correspondendo a estas possibilidades estão mostradas abaixo. Daqui em di-



ante usaremos sistematicamente a notação compacta para representar a fusão de anyons. Em geral, para anyons quaisquer x, y , a expressão $(x \underset{y}{\cdot} y)$, por exemplo, é avaliada primeiro expandindo-se x e y em suas decomposições em termos de τ e 1 depois usamos a distributividade de \otimes e avaliamos os caminhos de fusão que resultam em 1. Note que pode

haver mais de um. Não usaremos esta notação quando ela não possui significado. Por exemplo, a expressão $(\tau \underset{1}{;})$ não possui significado pois as regras de fusão não estabelecem que $\tau \otimes 1 = \tau$. O mesmo aconteceria com $(1 \underset{\tau}{;})$, por exemplo.

Agora note que esta notação nos fornece um conjunto de símbolos muito úteis. Para a fusão de três anyons τ , temos

$$\tau \otimes (\tau \otimes \tau) = 2 \cdot \tau \oplus 1$$

e, deste processo obtemos os três símbolos:

$$(\tau \underset{1}{\cdot} (\tau \underset{\tau}{\cdot} \tau)), \quad (\tau \underset{\tau}{\cdot} (\tau \underset{\tau}{\cdot} \tau)) \quad \text{e} \quad (\tau \underset{\tau}{\cdot} (\tau \underset{1}{\cdot} \tau)).$$

Cada um destes símbolos representa um anyon τ ou 1, o resultado da fusão, indicado pelo subscrito do operador \cdot mais externo. Mas além disto, este símbolo também nos informa como este anyon foi obtido. Se olharmos para a expressão para a fusão de três anyons τ mais uma vez, mas agora explicitando as somas:

$$\tau \otimes (\tau \otimes \tau) = \tau \otimes (\tau \oplus 1) = (\tau \otimes \tau) \oplus (\tau \otimes 1) = (\tau \oplus 1) \oplus \tau$$

vemos que estes símbolos nos informar a qual fator de soma, subespaço de $\tau \otimes (\tau \otimes \tau) = (\mathbb{C}^2, \mathbb{C})$ corresponde aquele anyon. Assim, o símbolo $(\tau \underset{1}{\cdot} (\tau \underset{\tau}{\cdot} \tau))$ corresponde ao subespaço $1 = (0, \mathbb{C})$ oriundo do fator $\tau \otimes \tau$, o símbolo $(\tau \underset{\tau}{\cdot} (\tau \underset{\tau}{\cdot} \tau))$ corresponde a $\tau = (\mathbb{C}, 0)$, também do fator $\tau \otimes \tau$ e o símbolo $(\tau \underset{\tau}{\cdot} (\tau \underset{1}{\cdot} \tau))$ corresponde a τ oriundo do fator $\tau \otimes 1$. Esta é exatamente a informação que precisamos para escrever a base de $(\mathbb{C}^2, \mathbb{C})$ de modo a respeitar a decomposição da soma. Nomeando os elementos da base por este símbolo, temos que

$$\left(\left\{ (\tau \underset{\tau}{\cdot} (\tau \underset{\tau}{\cdot} \tau)), (\tau \underset{\tau}{\cdot} (\tau \underset{1}{\cdot} \tau)) \right\}, \left\{ (\tau \underset{1}{\cdot} (\tau \underset{\tau}{\cdot} \tau)) \right\} \right)$$

forma a base de $(\mathbb{C}^2, \mathbb{C})$ que precisamos, onde um elemento da base de \mathbb{C}^2 é oriundo de $\tau \otimes \tau$ e outro de $\tau \otimes 1$. A base correspondente à outra escolha de parênteses, $(\tau \otimes \tau) \otimes \tau$ é escrita de modo complementemente análogo, apenas trocando os parênteses nos símbolos acima.

Com esta notação, estamos prontos para determinar $\alpha_{\mathcal{T}, \mathcal{T}, \mathcal{T}}$ e $\sigma_{\mathcal{T}, \mathcal{T}}$. A partir de agora trocaremos \mathcal{T} por τ por simplicidade. Pois bem, as duas escolhas de parênteses $\tau \otimes (\tau \otimes \tau)$ e $(\tau \otimes \tau) \otimes \tau$ produzem bases para $(\mathbb{C}^2, \mathbb{C})$ que são escritas, respectivamente, da seguinte forma:

$$\left(\left\{ (\tau \underset{\tau}{\cdot} (\tau \underset{\tau}{\cdot} \tau)), (\tau \underset{\tau}{\cdot} (\tau \underset{1}{\cdot} \tau)) \right\}, \left\{ (\tau \underset{1}{\cdot} (\tau \underset{\tau}{\cdot} \tau)) \right\} \right);$$

$$\left(\left\{ ((\tau \underset{\tau}{\cdot} \tau) \underset{\tau}{\cdot} \tau), ((\tau \underset{1}{\cdot} \tau) \underset{\tau}{\cdot} \tau) \right\}, \left\{ ((\tau \underset{\tau}{\cdot} \tau) \underset{1}{\cdot} \tau) \right\} \right).$$

Como vimos, $\alpha_{\tau, \tau, \tau}$ corresponde a dois mapas lineares, $m_{\tau} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ e outro de $m_1 :$

$\mathbb{C} \rightarrow \mathbb{C}$, correspondendo a mudanças de base nestes dois espaços devido às diferentes escolhas de parênteses. Por definição, $\alpha_{\tau,\tau,\tau}$ é inversível, por ser um isomorfismo, então estes mapas lineares também devem ser. Vistos como matrizes, isto indica que devem ser matrizes inversíveis. Com, isto então temos que $m_1 : \mathbb{C} \rightarrow \mathbb{C}$ é dado por

$$(\tau \cdot (\tau \cdot \tau)) \mapsto p \left((\tau \cdot \tau) \cdot \tau \right)$$

onde $p \neq 0$ e $m_\tau : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ é dado por:

$$\begin{pmatrix} (\tau \cdot (\tau \cdot \tau)) \\ (\tau \cdot (\tau \cdot \tau)) \end{pmatrix} \mapsto \begin{pmatrix} q & r \\ s & t \end{pmatrix} \begin{pmatrix} ((\tau \cdot \tau) \cdot \tau) \\ ((\tau \cdot \tau) \cdot \tau) \end{pmatrix}$$

onde a matriz

$$m = \begin{pmatrix} q & r \\ s & t \end{pmatrix}$$

é não singular. Agora investigamos as restrições impostas sobre m e p pela identidade pentagonal.

4.6.1 - Associatividade

Lembramos que a identidade pentagonal trata das diferentes escolhas de parênteses para a fusão $\tau^{\otimes 4}$ e envolve termos como $\alpha_{\tau \otimes \tau, \tau, \tau}$. Como $\tau \otimes \tau = \tau \oplus 1$, esta flecha pode ser escrita como uma soma direta de $\alpha_{\tau, \tau, \tau}$ e $\alpha_{1, \tau, \tau}$. Porém, como discutimos, se qualquer um dos índices é 1, a flecha é essencialmente a identidade, devido aos isomorfismos λ e ρ . Para $\tau^{\otimes 4}$ a identidade pentagonal é:

$$\begin{array}{ccc} & \tau \otimes (\tau \otimes (\tau \otimes \tau)) & \\ \swarrow \text{id}_\tau \otimes \alpha_{\tau, \tau, \tau} & & \searrow \alpha_{\tau, \tau, \tau \otimes \tau} \\ \tau \otimes ((\tau \otimes \tau) \otimes \tau) & & (\tau \otimes \tau) \otimes (\tau \otimes \tau) \quad (4.3) \\ \downarrow \alpha_{\tau, \tau \otimes \tau, \tau} & & \downarrow \alpha_{\tau \otimes \tau, \tau, \tau} \\ (\tau \otimes (\tau \otimes \tau)) \otimes \tau & \xrightarrow{\alpha_{\tau, \tau, \tau} \otimes \text{id}_\tau} & ((\tau \otimes \tau) \otimes \tau) \otimes \tau. \end{array}$$

Usando a equação 4.2, vê-se prontamente que $\tau^{\otimes 4} = (\mathbb{C}^3, \mathbb{C}^2)$. Pela identidade pentagonal vemos que existem cinco bases distintas para este par de espaços. Precisamos, então, escrever todas as flechas presentes na identidade em termos de matrizes nas respectivas

bases distintas deste espaço e ainda escrever estas matrizes em termos de m , p e da identidade como discutimos.

Começamos então com a flecha

$$\alpha_{\tau \otimes \tau, \tau, \tau} : (\tau \otimes \tau) \otimes (\tau \otimes \tau) \rightarrow ((\tau \otimes \tau) \otimes \tau) \otimes \tau.$$

O objeto $\tau \otimes \tau$ pode ser escrito como $\tau \oplus 1$. Usando a distributividade de \otimes , podemos escrever

$$\begin{aligned} (\tau \otimes \tau) \otimes (\tau \otimes \tau) &= \tau \otimes (\tau \otimes \tau) \oplus 1 \otimes (\tau \otimes \tau) \\ ((\tau \otimes \tau) \otimes \tau) \otimes \tau &= (\tau \otimes \tau) \otimes \tau \oplus (1 \otimes \tau) \otimes \tau. \end{aligned}$$

de modo que

$$\alpha_{\tau \otimes \tau, \tau, \tau} = \alpha_{\tau, \tau, \tau} \oplus \alpha_{1, \tau \tau}.$$

Para escrever o produto acima na forma matricial, voltamos nossa atenção para os pares de espaços e suas bases. A base de $(\tau \otimes \tau) \otimes (\tau \otimes \tau)$ é:

$$\left(\left(\left(\tau \cdot \tau \right) \cdot \left(\tau \cdot \tau \right) \right), \left(\left(\tau \cdot \tau \right) \cdot \left(\tau \cdot \tau \right) \right), \left(\left(\tau \cdot \tau \right) \cdot \left(\tau \cdot \tau \right) \right), \left(\left(\tau \cdot \tau \right) \cdot \left(\tau \cdot \tau \right) \right), \left(\left(\tau \cdot \tau \right) \cdot \left(\tau \cdot \tau \right) \right) \right).$$

Note que arranjamos esta base convenientemente de modo a condizer com a decomposição que fizemos acima. Os primeiros três vetores correspondem àqueles onde a primeira fusão resulta em τ e os dois últimos, onde resulta em 1. Esta ordenação também preserva a decomposição

$$(\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^2, \mathbb{C}) \oplus (\mathbb{C}, \mathbb{C})$$

que captura a decomposição de $\alpha_{\tau, \tau, \tau}$ e $\alpha_{1, \tau, \tau}$ em pares de mapas. Apesar disto, perdemos de vista a estrutura de par, $(\mathbb{C}^3, \mathbb{C}^2)$, pois os vetores cujo resultado final é τ — indicado pela operação central — estão misturados com aqueles onde o resultado final é 1. Isto não é um problema porque, como dito, podemos pensar o par $(\mathbb{C}^3, \mathbb{C}^2)$ como um espaço vetorial propriamente dito onde as operações são definidas por componentes e a notação da base marca a qual espaço cada componente pertence. Similarmente, a base para a outra configuração de parênteses escreve-se:

$$\left(\left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right) \cdot \tau \right), \left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right) \cdot \tau \right), \left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right) \cdot \tau \right), \left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right) \cdot \tau \right), \left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right) \cdot \tau \right) \right)$$

é a base para a outra configuração de parênteses organizada de modo a respeitar a decomposição.

A utilidade desta ordenação está no fato de que $\alpha_{\tau, \tau, \tau}$ atua apenas nos primeiros três vetores e $\alpha_{1, \tau, \tau}$ atua apenas nos dois últimos, como é próprio de operadores entre espaços obtidos a partir do produto direto. Como sabemos como $\alpha_{\tau, \tau, \tau}$ se comporta e que

$\alpha_{1,\tau,\tau} = \mathbb{1}$, pois

$$\begin{aligned} 1 \otimes (\tau \otimes \tau) &= 1 \otimes (\tau \oplus 1) = \tau \oplus 1 \\ (1 \otimes \tau) \otimes \tau &= \tau \otimes \tau = \tau \oplus 1, \end{aligned}$$

podemos determinar os elementos de matriz de $\alpha_{\tau \otimes \tau, \tau, \tau}$. Primeiro, notamos que $\alpha_{1,\tau,\tau} = \mathbb{1}$ resulta em

$$\begin{aligned} ((\tau \cdot \tau) \cdot \tau) \cdot \tau &= ((\tau \cdot \tau) \cdot (\tau \cdot \tau)) \\ ((\tau \cdot \tau) \cdot \tau) \cdot \tau &= ((\tau \cdot \tau) \cdot (\tau \cdot \tau)). \end{aligned}$$

Além disso, se pensarmos o par $(\mathbb{C}^2, \mathbb{C})$ não como um par, mas como um espaço vetorial, cujas duas bases sob consideração são formadas pelos três primeiros vetores de cada, podemos escrever a matriz que representa $\alpha_{\tau,\tau,\tau}$ como

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & r & 0 \\ s & t & 0 \\ 0 & 0 & p \end{pmatrix}$$

de modo que o mapa $\alpha_{\tau \otimes \tau, \tau, \tau}$ é representado matricialmente por

$$\alpha_{\tau \otimes \tau, \tau, \tau} = \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Por fim, podemos extrair os pares de mapa $m_\tau : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ e $m_1 : \mathbb{C}^2 \rightarrow \mathbb{C}^2$, simplesmente reordenando as bases acima para respeitarem explicitamente esta decomposição. O resultado disto, é simplesmente alterar a ordem das linhas e colunas na matriz acima. Neste caso, tudo o que precisamos fazer é trocar o terceiro e quarto vetores em ambas as bases. O resultado é:

$$\alpha_{\tau \otimes \tau, \tau, \tau} = \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Caso todo o processo acima, especialmente o último passo tenha parecido um pouco nebuloso, eis uma outra explicação do processo. Se retornamos ao início, podemos enxer-

gar as decomposições que fizemos da seguinte forma:

$$\begin{aligned}(\tau \otimes \tau) \otimes (\tau \otimes \tau) &= \tau \otimes (\tau \otimes \tau) \oplus 1 \otimes (\tau \otimes \tau) \leftrightarrow (\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^2, \mathbb{C}) \oplus (\mathbb{C}, \mathbb{C}) \\ ((\tau \otimes \tau) \otimes \tau) \otimes \tau &= (\tau \otimes \tau) \otimes \tau \oplus (1 \otimes \tau) \otimes \tau \leftrightarrow (\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^2, \mathbb{C}) \oplus (\mathbb{C}, \mathbb{C}).\end{aligned}$$

Esta decomposição é útil, justamente pois já sabemos como $\alpha_{\tau, \tau, \tau} : (\mathbb{C}^2, \mathbb{C}) \rightarrow (\mathbb{C}^2, \mathbb{C})$ atua. Além disto, nesta forma, sabemos também como $\alpha_{1, \tau, \tau} : (\mathbb{C}, \mathbb{C}) \rightarrow (\mathbb{C}, \mathbb{C})$ atua, é a identidade neste espaço. Com isto, podemos determinar explicitamente como todos os vetores da base se transformam. O último passo, então, corresponde a calcular os elementos de matriz de $\alpha_{\tau \otimes \tau, \tau, \tau}$ preservando a estrutura de par $(\mathbb{C}^3, \mathbb{C}^2)$, isto é, “desfazendo-se” a decomposição no produto direto acima de modo a obter um par de mapas. Explicitamente:

$$\left((\alpha_{\tau, \tau, \tau})_{\tau}, (\alpha_{\tau, \tau, \tau})_1 \right) \oplus \left((\alpha_{1, \tau, \tau})_{\tau}, (\alpha_{1, \tau, \tau})_1 \right) = \left((\alpha_{\tau \otimes \tau, \tau, \tau})_{\tau}, (\alpha_{\tau \otimes \tau, \tau, \tau})_1 \right).$$

com $(\alpha_{\tau, \tau, \tau})_{\tau} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$, $(\alpha_{\tau, \tau, \tau})_1, (\alpha_{1, \tau, \tau})_{\tau}, (\alpha_{1, \tau, \tau})_1 : \mathbb{C} \rightarrow \mathbb{C}$, $(\alpha_{\tau \otimes \tau, \tau, \tau})_{\tau} : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ e $(\alpha_{\tau \otimes \tau, \tau, \tau})_1 : \mathbb{C}^2 \rightarrow \mathbb{C}^2$. Efetivamente, isto corresponde a agrupar as componentes τ e as componentes 1 de $\alpha_{\tau, \tau, \tau}$ e $\alpha_{1, \tau, \tau}$. É análogo a entender $(\mathbb{C}^3, \mathbb{C}^2)$ agora como a decomposição

$$(\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^3, 0) \oplus (0, \mathbb{C}^2).$$

Mas, esta decomposição corresponde justamente a trocar o terceiro e quarto vetor de ambas as bases porém mantendo em mente que o terceiro vetor da base se transforma sob $(\alpha_{1, \tau, \tau})_{\tau}$ e não $(\alpha_{\tau, \tau, \tau})_1$ e vice-versa para o quarto. É análogo, porém mais complicado, a considerar \mathbb{R}^3 como composto do plano xy mais o eixo z ou como composto do plano xz mais o eixo y . A representação matricial de uma transformação que afeta a componente z será diferente, pois em uma decomposição z pertence à segunda parcela da decomposição enquanto na outra pertence à primeira.

Em resumo, o que fizemos até agora foi fixar as bases de $(\tau \otimes \tau) \otimes (\tau \otimes \tau)$ e $((\tau \otimes \tau) \otimes \tau) \otimes \tau$ e determinar a representação matricial de $\alpha_{\tau \otimes \tau, \tau, \tau}$ nestas bases. Note que no último passo, trocamos o terceiro e quarto em ambas as bases. Isto corresponde efetivamente a uma mudança de base em ambas. Como dissemos, isto equivale a trocar as linhas e colunas da representação de $\alpha_{\tau \otimes \tau, \tau, \tau}$. Como este mapa é de $(\tau \otimes \tau) \otimes (\tau \otimes \tau) \rightarrow ((\tau \otimes \tau) \otimes \tau) \otimes \tau$, trocar o terceiro e quarto vetor de lugar na base do domínio, equivale a trocar linhas e no contradomínio, colunas. Assim, o resultado é a representação matricial mostrada anteriormente e as bases de ambos os espaços que escolhemos foram:

$$\left(\left(\begin{matrix} (\tau \cdot \tau) \\ \tau \end{matrix} \cdot \begin{matrix} (\tau \cdot \tau) \\ \tau \end{matrix} \right), \left(\begin{matrix} (\tau \cdot \tau) \\ \tau \end{matrix} \cdot \begin{matrix} (\tau \cdot \tau) \\ 1 \end{matrix} \right), \left(\begin{matrix} (\tau \cdot \tau) \\ 1 \end{matrix} \cdot \begin{matrix} (\tau \cdot \tau) \\ \tau \end{matrix} \right), \left(\begin{matrix} (\tau \cdot \tau) \\ 1 \end{matrix} \cdot \begin{matrix} (\tau \cdot \tau) \\ \tau \end{matrix} \right), \left(\begin{matrix} (\tau \cdot \tau) \\ 1 \end{matrix} \cdot \begin{matrix} (\tau \cdot \tau) \\ 1 \end{matrix} \right) \right) \quad (4.4)$$

para $(\tau \otimes \tau) \otimes (\tau \otimes \tau)$ e, para

$$\left((((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau) \right) \quad (4.5)$$

para $((\tau \otimes \tau) \otimes \tau) \otimes \tau$. Note que não são as bases que escolhemos inicialmente. Restam três outras configurações de parênteses para as quais ainda não escolhemos bases. Precisamos, agora, fixar bases para estas três configurações e determinar a representação matricial das outras quatro flechas.

Assim, considere agora a flecha

$$\alpha_{\tau, \tau, \tau \otimes \tau} : \tau \otimes (\tau \otimes (\tau \otimes \tau)) \rightarrow (\tau \otimes \tau) \otimes (\tau \otimes \tau).$$

Novamente, faremos uso da decomposição $\tau \otimes \tau = \tau \oplus 1$, mas desta vez, o produto que é decomposto é aquele à direita. Lembre-se que no caso anterior, decomparamos o produto à esquerda em $(\tau \otimes \tau) \otimes (\tau \otimes \tau)$. Isto terá consequências. A base para a primeira escolha de parênteses é escrita como:

$$\left((\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))) \right).$$

enquanto a base para a segunda, agora é:

$$\left(((\tau \cdot \tau) \cdot (\tau \cdot \tau)), ((\tau \cdot \tau) \cdot (\tau \cdot \tau)), ((\tau \cdot \tau) \cdot (\tau \cdot \tau)), ((\tau \cdot \tau) \cdot (\tau \cdot \tau)), ((\tau \cdot \tau) \cdot (\tau \cdot \tau)) \right).$$

Note que, em relação à base 4.4, o segundo, terceiro e quarto vetores estão fora de ordem. Ao final precisaremos corrigir isto. Nestas bases, no entanto, podemos proceder de modo análogo a como fizemos no primeiro caso e obter a seguinte representação matricial:

$$\alpha_{\tau, \tau, \tau \otimes \tau} = \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Como mencionamos, como a base de $(\tau \otimes \tau) \otimes (\tau \otimes \tau)$ já está fixa, precisamos transformar a base que escolhemos para determinar a representação matricial de $\alpha_{\tau, \tau, \tau \otimes \tau}$ na base 4.4. Podemos obter isto, trocando o terceiro e quarto vetores e depois o segundo e terceiro.

Isto corresponde a trocar apenas as colunas na representação matricial e obtemos:

$$\alpha_{\tau, \tau, \tau \otimes \tau} = \begin{pmatrix} q & 0 & r & 0 & 0 \\ s & 0 & t & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Por conveniência, e para deixar a base de $\tau \otimes (\tau \otimes (\tau \otimes \tau))$ de acordo com a decomposição:

$$(\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^3, 0) \oplus (0, \mathbb{C}^2)$$

trocamos também o terceiro e quarto vetores desta base, fixando-a como:

$$\left((\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))), (\tau \cdot (\tau \cdot (\tau \cdot \tau))) \right) \quad (4.6)$$

e obtendo o seguinte resultado final para $\alpha_{\tau, \tau, \tau \otimes \tau}$:

$$\alpha_{\tau, \tau, \tau \otimes \tau} = \begin{pmatrix} q & 0 & r & 0 & 0 \\ s & 0 & t & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Estas duas matrizes determinam o lado curto da identidade pentagonal:

$$\tau \otimes (\tau \otimes (\tau \otimes \tau)) \xrightarrow{\alpha_{\tau, \tau, \tau \otimes \tau}} (\tau \otimes \tau) \otimes (\tau \otimes \tau) \xrightarrow{\alpha_{\tau \otimes \tau, \tau, \tau}} ((\tau \otimes \tau) \otimes \tau) \otimes \tau.$$

A estrutura abeliana de \mathfrak{F} nos diz que a composição destes dois mapas, corresponde à multiplicação matricial, então

$$\alpha_{\tau \otimes \tau, \tau, \tau} \circ \alpha_{\tau, \tau, \tau \otimes \tau} = \begin{pmatrix} q & 0 & r & 0 & 0 \\ s & 0 & t & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} q^2 & qr & r & 0 & 0 \\ qs & rs & t & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 0 & p^2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

No lado longo, além da flecha $\alpha_{\tau, \tau \otimes \tau, \tau}$, temos as duas flechas $\text{id}_{\tau} \otimes \alpha_{\tau, \tau, \tau}$ e $\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau}$. Começamos, então, por determinar a flecha $\text{id}_{\tau} \otimes \alpha_{\tau, \tau, \tau} : \tau \otimes (\tau \otimes (\tau \otimes \tau)) \rightarrow \tau \otimes ((\tau \otimes \tau) \otimes \tau)$. Escolhemos começar com esta flecha porque já determinamos a base de $\tau \otimes (\tau \otimes (\tau \otimes \tau))$. Além disso, pela functorialidade de \otimes , a flecha $\text{id}_{\tau} \otimes \alpha_{\tau, \tau, \tau}$ nos permite ignorar o fator τ mais à esquerda pois o ele permanece inalterado enquanto o $\sigma_{\tau, \tau, \tau}$ atua normalmente

sobre o resto. Deste modo, ao invés de usarmos a decomposição de $\tau \otimes \tau$ para determinar uma base, obter a representação matricial nesta base e então transformá-la conforme necessário, construiremos uma base de $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$ a partir dos vetores linearmente independentes da imagem de 4.6 sob $\text{id}_\tau \otimes \alpha_{\tau,\tau,\tau}$ e então a transformaremos, se necessário.

Lembre-se que $\sigma_{\tau,\tau,\tau}$ relaciona as duas bases possíveis de $\tau^{\otimes 3}$, mapeando

$$\left((\tau \cdot (\tau \cdot \tau))_\tau, (\tau \cdot (\tau \cdot \tau))_{\tau_1}, (\tau \cdot (\tau \cdot \tau))_{\tau_1} \right)$$

em

$$\left(((\tau \cdot \tau) \cdot \tau)_\tau, ((\tau \cdot \tau) \cdot \tau)_{\tau_1}, ((\tau \cdot \tau) \cdot \tau)_{\tau_1} \right).$$

Usando isto, obtemos a seguinte base para $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$ a partir de 4.6:

$$\left((\tau \cdot ((\tau \cdot \tau) \cdot \tau))_\tau, (\tau \cdot ((\tau \cdot \tau) \cdot \tau))_{\tau_1}, (\tau \cdot ((\tau \cdot \tau) \cdot \tau))_{\tau_1}, (\tau \cdot ((\tau \cdot \tau) \cdot \tau))_{\tau_1}, (\tau \cdot ((\tau \cdot \tau) \cdot \tau))_{\tau_1} \right). \quad (4.7)$$

Com esta base, $\text{id}_\tau \otimes \alpha_{\tau,\tau,\tau}$ é representada matricialmente por:

$$\text{id}_\tau \otimes \alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & 0 & r & 0 & 0 \\ 0 & p & 0 & 0 & 0 \\ s & 0 & t & 0 & 0 \\ 0 & 0 & 0 & q & r \\ 0 & 0 & 0 & s & t \end{pmatrix}$$

Para ver porque esta é a representação matricial, note que o segundo vetor de 4.6 se transforma segundo $(\sigma_{\tau,\tau,\tau})_1 = [p]$, pois é dado por $\tau \otimes (\tau \cdot (\tau \cdot \tau))_{\tau_1}$. Só existe um resultado possível para este produto, já que corresponde a $\tau \otimes 1$, porém isto não é relevante. Se houvessem outras possibilidades, teríamos mais de um vetor se transformando segundo $(\sigma_{\tau,\tau,\tau})_1$. Este é o caso dos outros quatro vetores. Eles podem ser divididos em dois pares: $\tau \otimes (\tau \cdot (\tau \cdot \tau))_{\tau_1}$ e $\tau \otimes (\tau \cdot (\tau \cdot \tau))_{\tau_1}$. Estes vetores se transformam segundo $(\sigma_{\tau,\tau,\tau})_\tau = m$. Explicitamente:

$$\begin{pmatrix} (\tau \cdot (\tau \cdot \tau))_\tau \\ (\tau \cdot (\tau \cdot \tau))_{\tau_1} \end{pmatrix} = \begin{pmatrix} q & r \\ s & t \end{pmatrix} \begin{pmatrix} ((\tau \cdot \tau) \cdot \tau)_\tau \\ ((\tau \cdot \tau) \cdot \tau)_{\tau_1} \end{pmatrix}$$

Assim, o primeiro vetor de 4.6 é levado em uma combinação linear do primeiro e terceiro de 4.7. Assim, como o terceiro de 4.6. Similarmente, o quarto e o quinto são levados em combinações lineares do quarto e quinto de 4.7. Resultando na representação matricial acima. Note que a base 4.7 respeita a decomposição:

$$(\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^3, 0) \oplus (0, \mathbb{C}^2)$$

e assim, fixamos ela como a base de $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$.

Continuando, a penúltima flecha é: $\alpha_{\tau, \tau \otimes \tau, \tau} : \tau \otimes ((\tau \otimes \tau) \otimes \tau) \rightarrow (\tau \otimes (\tau \otimes \tau)) \otimes \tau$. Para esta flecha, faremos como para as duas primeiras. Iremos decompor o produto central e escrever bases para $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$ e $(\tau \otimes (\tau \otimes \tau)) \otimes \tau$ e então modificá-las conforme adequado. A base de $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$ é dada por

$$\left((\tau \cdot (\tau \cdot \tau) \cdot \tau), (\tau \cdot (\tau \cdot \tau) \cdot \tau), (\tau \cdot (\tau \cdot \tau) \cdot \tau), (\tau \cdot (\tau \cdot \tau) \cdot \tau), (\tau \cdot (\tau \cdot \tau) \cdot \tau) \right)$$

e a base de $(\tau \otimes (\tau \otimes \tau)) \otimes \tau$ é dada por:

$$\left(((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau) \right).$$

Com estas bases temos, novamente, a seguinte representação matricial:

$$\alpha_{\tau, \tau \otimes \tau, \tau} = \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Em relação à base 4.7, a base que escrevemos para $\tau \otimes ((\tau \otimes \tau) \otimes \tau)$ está com o terceiro e quarto vetores trocados. Além disso, para que a base de $(\tau \otimes (\tau \otimes \tau)) \otimes \tau$ que escrevemos respeite a decomposição

$$(\mathbb{C}^3, \mathbb{C}^2) = (\mathbb{C}^3, 0) \oplus (0, \mathbb{C}^2)$$

precisamos trocar, também, o terceiro e quarto vetores. Assim, trocamos estes vetores em ambas as bases, fixamos a base de $(\tau \otimes (\tau \otimes \tau)) \otimes \tau$ como

$$\left(((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau), ((\tau \cdot (\tau \cdot \tau)) \cdot \tau) \right) \quad (4.8)$$

e o obtemos:

$$\alpha_{\tau, \tau \otimes \tau, \tau} = \begin{pmatrix} q & r & 0 & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Por fim, chegamos à última flecha: $\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau} : (\tau \otimes (\tau \otimes \tau)) \otimes \tau \rightarrow ((\tau \otimes \tau) \otimes \tau) \otimes \tau$. As bases de ambos estes espaços já foram escolhidas. Assim como no caso de $\text{id}_{\tau} \otimes \alpha_{\tau, \tau, \tau}$ nos aproveitaremos da functorialidade de \otimes para ignorar o fator τ mais a direita. Para determinar a representação matricial de $\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau}$, então, iremos escrever a base de $((\tau \otimes \tau) \otimes \tau) \otimes \tau$ obtida a partir da imagem de 4.8 sob $\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau}$ e então transformar

esta base em 4.5. Temos então:

$$\left((((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau), (((\tau \cdot \tau) \cdot \tau) \cdot \tau) \right).$$

Com esta base, obtemos a seguinte representação matricial:

$$\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau} = \begin{pmatrix} q & 0 & r & 0 & 0 \\ 0 & p & 0 & 0 & 0 \\ s & 0 & t & 0 & 0 \\ 0 & 0 & 0 & q & r \\ 0 & 0 & 0 & s & t \end{pmatrix}.$$

Felizmente, esta base é exatamente a base 4.5 que fixamos para $((\tau \otimes \tau) \otimes \tau) \otimes \tau$, portanto não é necessário realizar mais nenhuma transformação. Com isto temos todos os resultados necessários.

O lado longo da identidade pentagonal equivale à composição:

$$(\alpha_{\tau, \tau, \tau} \otimes \text{id}_{\tau}) \circ \alpha_{\tau, \tau \otimes \tau, \tau} \circ (\text{id}_{\tau} \otimes \alpha_{\tau, \tau, \tau})$$

cuja representação matricial é:

$$\begin{pmatrix} q^3 + rs & pqr & q^2r + rt & 0 & 0 \\ pqs & p^2t & prs & 0 & 0 \\ q^2s + st & prs & qrs + t^2 & 0 & 0 \\ 0 & 0 & 0 & pq^2 + rs & pqr + rt \\ 0 & 0 & 0 & pqs + st & prs + t^2 \end{pmatrix}$$

Como a identidade pentagonal comuta, igualamos o lado curto ao lado longo e obtemos as seguintes condições sobre os números complexos p, q, r, s, t :

$$\begin{pmatrix} q^2 & qr & r & 0 & 0 \\ qs & rs & t & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 0 & p^2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} q^3 + rs & pqr & q^2r + rt & 0 & 0 \\ pqs & p^2t & prs & 0 & 0 \\ q^2s + st & prs & qrs + t^2 & 0 & 0 \\ 0 & 0 & 0 & pq^2 + rs & pqr + rt \\ 0 & 0 & 0 & pqs + st & prs + t^2 \end{pmatrix}$$

Prontamente extraímos desta igualdade as duas equações:

$$qs = pqs \quad \text{e} \quad pqr = qr$$

donde concluimos que $p = 1$. Substituindo isto em todas as outras equações, obtemos:

$$\begin{pmatrix} q^2 & qr & r & 0 & 0 \\ qs & rs & t & 0 & 0 \\ s & t & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} q^3 + rs & qr & q^2r + rt & 0 & 0 \\ qs & t & rs & 0 & 0 \\ q^2s + st & rs & qrs + t^2 & 0 & 0 \\ 0 & 0 & 0 & q^2 + rs & qr + rt \\ 0 & 0 & 0 & qs + st & rs + t^2 \end{pmatrix} \quad (4.9)$$

Agora, temos uma equação redundante

$$qs = qs$$

e três equações

$$rs = t.$$

A partir desta, obtemos

$$qrs + t^2 = 0 \implies qt + t^2 = 0 \implies t(q + t) = 0.$$

Suponha, então que $t = 0$. Neste caso, temos que $r = 0$ ou $s = 0$. Supondo $r = 0$, obtemos:

$$\begin{pmatrix} q^2 & 0 & 0 & 0 & 0 \\ qs & 0 & 0 & 0 & 0 \\ s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} q^3 & 0 & q^2r & 0 & 0 \\ qs & 0 & 0 & 0 & 0 \\ q^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q^2 & 0 \\ 0 & 0 & 0 & qs & 0 \end{pmatrix}$$

Donde obtemos que $qs = 0$ e $q^2 = s$ o que implica que $q = s = 0$. Isto contradiz a exigência de que a matriz

$$m = \begin{pmatrix} q & s \\ r & t \end{pmatrix}$$

seja não singular. Portanto, devemos ter $q = -t$.

Retornando às equações 4.9 com este resultado, temos ainda que

$$rs = t \quad \text{e} \quad q^2 = q^3 + rs.$$

Substituindo $q = -t$ obtemos

$$q^2 = q^3 - q \implies q^3 - q^2 - q = 0 \implies q(q^2 - q - 1) = 0.$$

Se $q = 0$, $t = 0$ e $rs = 0$, donde $r = 0$ ou $s = 0$. Mas, temos também que

$$s = q^2s + st \quad \text{e} \quad r = q^2r + rt$$

que implica que $r = s = 0$. Como isto contradiz a condição de que m seja invertível, devemos ter

$$q^2 - q - 1 = 0 \implies q = \frac{-1 \pm \sqrt{5}}{2}.$$

Então temos

$$q = -t = \frac{-1 \pm \sqrt{5}}{2} \quad rs = t = -q.$$

Junto com $p = 1$ isto satisfaz completamente o sistema de equações.

Para ver isto, note que existem três equações $rs = t$ e uma equação redundante, $qs = qs$. As outras equações são $q^2 = q^3 + rs$, de onde obtivemos a solução para q e

$$s = s(q^2 + t), \quad r = r(q^2 + t) \quad \text{e} \quad qrs + t^2 = 0.$$

As duas primeiras equações resultam em $q^2 - q - 1 = 0$, que é satisfeita por q e a última resulta em $-q^2 + t^2 = 0$ que é satisfeita pois $q = -t$. As últimas quatro equações são

$$qs + st = 0 \implies q = -t \quad \text{e} \quad qr + rt = 0 \implies q = -t$$

e

$$q^2 + rs = 1 \implies q^2 - q - 1 = 0 \quad \text{e} \quad rs + t^2 = 1 \implies t^2 + t - 1 = 0.$$

Como $t = -q$, ambas são iguais e o são satisfeitas. Portanto o conjunto solução da identidade pentagonal para

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & r & 0 \\ s & t & 0 \\ 0 & 0 & p \end{pmatrix}$$

é dado por

$$p = 1, \quad q = -t = \frac{-1 \pm \sqrt{5}}{2}, \quad rs = -q.$$

Podemos extrair mais informação impondo que $\alpha_{\tau,\tau,\tau}$ seja unitária. Isto equivale a dizer que todos os espaços $[\mathcal{A}, \mathcal{B}]$ sejam espaços de Hilbert [42]. Isto não é um problema pois, como havíamos discutido, estes espaços já carregam estrutura o suficiente para serem espaços de Hilbert. Além disto, lembre-se que as flechas associadoras α correspondem a trançamentos de anyons, então, para que estas flechas representem uma evolução temporal genuína dos anyons, devemos realmente impor que sejam unitárias. Como $p = 1$, a

unitariedade não nos fornece novas informações. Porém para a matriz m , obtemos

$$\begin{pmatrix} q & r \\ s & t \end{pmatrix} \begin{pmatrix} q & s^* \\ r^* & t \end{pmatrix} = \begin{pmatrix} q^2 + |r|^2 & qs^* + rt \\ qs + r^*t & t^2 + |s|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

donde

$$q^2 + |r|^2 = q^2 + |s|^2 = 1 \quad \text{e} \quad q(s - r^*) = q(s^* - r) = 0.$$

Note que usamos o fato de q ser real. Do segundo par de equações obtemos que $r = s^*$. Substituindo isto em $rs = -q$, obtemos que $-q = |r|^2 > 0$. Logo, devemos ter

$$q = \frac{-1 - \sqrt{5}}{2}.$$

e

$$s = \sqrt{-q}e^{i\theta} \quad \text{e} \quad r = \sqrt{-q}e^{-i\theta}$$

para alguma θ real. Assim, o resultado final é

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & \sqrt{-q}e^{-i\theta} & 0 \\ \sqrt{-q}e^{i\theta} & -q & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Repare que $q = -\phi$, onde ϕ é a razão áurea. A menos do parâmetro θ , determinamos completamente $\alpha_{\tau,\tau,\tau}$. Se quiséssemos, poderíamos eliminar a fase aplicando rotações adequadas nas bases que escolhemos [42]. Além disso, note que não existe uma base canônica para os cinco espaços correspondentes às diferentes escolhas de parênteses em $\tau^{\otimes 4}$. As bases “mais canônicas” são aquelas que diferem apenas por permutações das bases que escolhemos aqui. Se escolhêssemos outras bases, obteríamos expressões similares, afinal a forma de $\alpha_{\tau,\tau,\tau}$, ou melhor, de suas componentes $(\alpha_{\tau,\tau,\tau})_{\tau}$ e $(\alpha_{\tau,\tau,\tau})_1$ não deve mudar. No entanto, poderíamos obter outros valores de q como

$$q = \frac{-1 + \sqrt{5}}{2}$$

por exemplo. Este foi o resultado obtido no trabalho de Blass e Gurevich [42]. A única diferença deste problema para um problema usual de mecânica quântica é que não existe uma base natural. Porém, fixadas as bases, todas as soluções são equivalentes e conectadas por transformações de base. O próximo passo agora é repetir todo este procedimento para os isomorfismos de trançamento, σ e as identidades hexagonais. Felizmente o processo é razoavelmente mais simples.

4.6.2 - Trançamentos

A estrutura de tranças de \mathfrak{F} é dada pelos isomorfismos $\sigma_{\mathcal{A},\mathcal{B}} : \mathcal{A} \otimes \mathcal{B} \rightarrow \mathcal{B} \otimes \mathcal{A}$ para todos os objetos $\mathcal{A}, \mathcal{B} \in \mathfrak{F}$. Assim como no caso das flechas associadores, os únicos isomorfismos que precisamos nos preocupar neste caso são aqueles com ambos os índices iguais a τ . Ou seja, $\sigma_{\tau,\tau}$ e $\sigma_{\tau,\tau}^{-1}$. Se algum destes índices for 1, temos a identidade pois

$$\mathcal{A} \otimes \mathbf{1} = \mathcal{A} = \mathcal{A} \otimes \mathbf{1}.$$

Além disso, se os índices forem objetos não simples de \mathfrak{F} , então estes objetos podem ser escritos como uma soma de objetos simples e a flecha σ , conseqüentemente, pode ser escrita em termos de $\sigma_{\tau,\tau}$ e $\sigma_{\tau,\tau}^{-1}$. Por fim, note que $\sigma_{\tau,\tau}$ é uma flecha de $\tau \otimes \tau = \tau \oplus 1 \rightarrow \tau \oplus 1$. Assim, as duas componentes de $\sigma_{\tau,\tau}$ são unidimensionais. Como σ são inversíveis, as duas componentes são dadas por escalares não nulos. Escrevemos:

$$\sigma_{\tau,\tau} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Para determinar a e b , usamos as duas identidades hexagonais, dadas por:

$$\begin{array}{ccccc}
 & & (\tau \otimes \tau) \otimes \tau & \xrightarrow{\sigma_{\tau \otimes \tau, \tau}} & \tau \otimes (\tau \otimes \tau) \\
 & \nearrow \alpha_{\tau, \tau, \tau} & & & \searrow \alpha_{\tau, \tau, \tau} \\
 \tau \otimes (\tau \otimes \tau) & & & & (\tau \otimes \tau) \otimes \tau \\
 & \searrow \text{id}_{\tau} \otimes \sigma_{\tau, \tau} & & & \nearrow \sigma_{\tau, \tau} \otimes \text{id}_{\tau} \\
 & & \tau \otimes (\tau \otimes \tau) & \xrightarrow{\alpha_{\tau, \tau, \tau}} & (\tau \otimes \tau) \otimes \tau
 \end{array} \tag{4.10}$$

e o mesmo diagrama porém com $\sigma_{\tau,\tau}^{-1}$ no lugar de $\sigma_{\tau,\tau}$. O procedimento que devemos realizar é completamente análogo, porém agora temos a vantagem que, seguindo o princípio que escolhemos, existem apenas duas bases para $\tau^{\otimes 3}$, dadas por:

$$\left(\left(\tau \cdot \left(\tau \cdot \tau \right) \right), \left(\tau \cdot \left(\tau \cdot \tau \right) \right), \left(\tau \cdot \left(\tau \cdot \tau \right) \right) \right) \tag{4.11}$$

$$\left(\left(\left(\tau \cdot \tau \right) \cdot \tau \right), \left(\left(\tau \cdot \tau \right) \cdot \tau \right), \left(\left(\tau \cdot \tau \right) \cdot \tau \right) \right). \tag{4.12}$$

Fixaremos então estas bases para as duas escolhas de parênteses possíveis. Como já

sabemos que com estas escolhas $\alpha_{\tau,\tau,\tau}$ é dada por:

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & \sqrt{-q}e^{-i\theta} & 0 \\ \sqrt{-q}e^{i\theta} & -q & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

onde $q = -(1 + \sqrt{5})/2$ e $\theta \in \mathbb{R}$, precisamos calcular apenas as flechas $\text{id}_{\tau} \otimes \sigma_{\tau,\tau}$, $\sigma_{\tau,\tau} \otimes \text{id}_{\tau}$ e $\sigma_{\tau \otimes \tau, \tau}$.

As flechas com a identidade são simples de calcular, pois a identidade nos permite ignorar o produto mais externo e considerar apenas a ação de $\sigma_{\tau,\tau}$ no termo $(\tau \otimes \tau)$ interno. Como vimos, $\sigma_{\tau,\tau}$ multiplica $(\tau \dot{\vdash} \tau)$ por a e $(\tau \dot{\vdash} \tau)$ por b . Assim, obtemos:

$$\text{id}_{\tau} \otimes \sigma_{\tau,\tau} = \sigma_{\tau,\tau} \otimes \text{id}_{\tau} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & a \end{pmatrix}.$$

Para determinar a representação matricial de $\sigma_{\tau \otimes \tau, \tau}$, usamos a decomposição $\tau \otimes \tau = \tau \oplus 1$. Com esta decomposição, escrevemos a base de $(\tau \otimes \tau) \otimes \tau$ como:

$$\left(((\tau \cdot \tau) \cdot \tau)_{\tau}, ((\tau \cdot \tau) \cdot \tau)_{\dot{1}}, ((\tau \cdot \tau) \cdot \tau)_{\tau} \right).$$

Note que trocamos os dois últimos vetores em relação a 4.12. Nesta base, temos

$$\sigma_{\tau \otimes \tau, \tau} = \sigma_{\tau,\tau} \oplus \sigma_{1,\tau}.$$

Como dito, $\sigma_{1,\tau}$ é a identidade. Assim, aplicando $\sigma_{\tau \otimes \tau, \tau}$ na base acima, obtemos

$$\left((\tau \cdot (\tau \cdot \tau))_{\tau}, (\tau \cdot (\tau \cdot \tau))_{\dot{1}}, (\tau \cdot (\tau \cdot \tau))_{\tau} \right)$$

e a seguinte representação matricial:

$$\sigma_{\tau \otimes \tau, \tau} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note que o segundo e terceiro vetores na base resultante também estão trocados em relação a 4.11. Para passarmos para as bases que escolhemos, trocamos estes vetores em ambas as bases, resultando em:

$$\sigma_{\tau \otimes \tau, \tau} = \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & b \end{pmatrix}.$$

Com isto todas as flechas estão determinadas e podemos usar as identidades hexagonais para determinar a e b . Note que a identidade hexagonal com $\sigma_{\tau,\tau}^{-1}$ corresponde simplesmente a trocar a por a^{-1} e b por b^{-1} , já que a matriz $\sigma_{\tau,\tau}$ é diagonal. No que segue escreveremos

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & r & 0 \\ s & -q & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

por conveniência. O lado superior da identidade hexagonal corresponde à composta

$$\alpha_{\tau,\tau,\tau} \circ \sigma_{\tau \otimes \tau, \tau} \circ \alpha_{\tau,\tau,\tau}$$

que corresponde ao produto:

$$\begin{pmatrix} q & r & 0 \\ s & -q & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} q & r & 0 \\ s & -q & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} aq^2 + rs & aqr - qr & 0 \\ aqs - qs & ars + q^2 & 0 \\ 0 & 0 & b \end{pmatrix}.$$

Por outro lado, o lado inferior da identidade hexagonal corresponde à composta:

$$(\sigma_{\tau,\tau} \otimes \text{id}_{\tau}) \circ \alpha_{\tau,\tau,\tau} \circ (\text{id}_{\tau} \otimes \sigma_{\tau,\tau})$$

cuja representação matricial é:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & a \end{pmatrix} \begin{pmatrix} q & r & 0 \\ s & -q & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & a \end{pmatrix} = \begin{pmatrix} a^2q & abr & 0 \\ abs & -b^2q & 0 \\ 0 & 0 & a^2 \end{pmatrix}.$$

Como a identidade hexagonal comuta, igualamos as duas compostas e obtemos as seguintes restrições sobre a e b :

$$\begin{pmatrix} aq^2 + rs & aqr - qr & 0 \\ aqs - qs & ars + q^2 & 0 \\ 0 & 0 & b \end{pmatrix} = \begin{pmatrix} a^2q & abr & 0 \\ abs & -b^2q & 0 \\ 0 & 0 & a^2 \end{pmatrix}. \quad (4.13)$$

A outra identidade hexagonal produz o mesmo resultado porém com a^{-1} e b^{-1} no lugar de a e b .

Destas equações, vemos prontamente que

$$b = a^2.$$

Além disto, sabemos que $rs = -q$, substituindo estes dois resultados nas equações acima

obtemos:

$$\begin{pmatrix} aq^2 - q & aqr - qr & 0 \\ aqs - qs & q^2 - aq & 0 \\ 0 & 0 & a^2 \end{pmatrix} = \begin{pmatrix} a^2q & a^3r & 0 \\ a^3s & -a^4q & 0 \\ 0 & 0 & a^2 \end{pmatrix}.$$

Destas equações obtemos:

$$\begin{cases} aq^2 - q = a^2q \implies q(a^2 - aq + 1) = 0; \\ q^2 - aq = -a^4q \implies q(a^4 - a + q) = 0. \end{cases}$$

Como sabemos que $q \neq 0$, obtemos

$$a^2 - aq + 1 = 0 \quad \text{e} \quad a^4 - a + q = 0.$$

Além disto, temos ainda

$$\begin{cases} aqs - qs = a^3s \implies a^3 - aq + q = 0; \\ aqr - qr = a^3r \implies a^3 - aq + q = 0. \end{cases}$$

Agora, resolvendo a equação quadrática para a , obtemos:

$$a = \frac{q \pm \sqrt{q^2 - 4}}{2}.$$

como $q = -(1 + \sqrt{5})/2$, temos que $-2 < q < -1$, já que $2 < \sqrt{5} < 3$, temos que $q^2 - 4 < 0$, logo a é imaginário e possui parte real igual a $q/2$. Além disto, escrevendo

$$a = \frac{1}{2} \left[q \pm i\sqrt{4 - q^2} \right]$$

obtemos que

$$|a|^2 = \frac{1}{4} \left[q^2 + 4 - q^2 \right] = 1.$$

Assim, ambas as soluções, que são conjugadas complexas uma da outra, possuem módulo unitário e parte real $q/2$. Podemos escolher, arbitrariamente, qualquer uma das soluções. Isto corresponde, efetivamente a escolher se os geradores do grupo de tranças, $\sigma_{\tau,\tau}$ movem uma trança ao redor de outra no sentido horário ou anti-horário. Efetivamente, uma solução fornece $\sigma_{\tau,\tau}$ e a outra $\sigma_{\tau,\tau}^{-1}$.

Apesar desta arbitrariedade, existe mais que podemos dizer sobre a e b . Temos

$$a^2 - aq + 1 = 0 \implies q = \frac{a^2 + 1}{a}.$$

Substituindo em $a^3 - aq + q = 0$ e $a^4 - a + q$, obtemos

$$\begin{cases} a^3 - a^2 - 1 + \frac{a^2 + 1}{a} = 0 \implies a^4 - a^3 + a^2 - a + 1 = 0; \\ a^4 - a + \frac{a^2 + 1}{a} = 0 \implies a^5 + 1 = 0. \end{cases}$$

Ambos os resultados fornecem que $-a$ é uma raiz primitiva quántupla de unidade, de modo a é uma raiz primitiva décupla da unidade. Ou seja:

$$a = e^{2k\pi i/10} = e^{k\pi i/5}, \quad k = 1, 2, \dots, 9$$

com k tal que $a^n \neq 1$ para $n = 1, 2, \dots, 9$. Existem apenas quatro tais raízes: $e^{\pi i/5}$, $e^{3\pi i/5}$, $e^{7\pi i/5}$ e $e^{9\pi i/5}$. Note que estas raízes estão relacionadas por conjugação complexa pois

$$e^{9\pi i/5} = e^{(10-1)\pi i/5} = e^{2\pi i} e^{-\pi i/5} = e^{-\pi i/5} \quad e \quad e^{7\pi i/5} = e^{(10-3)\pi i/5} = e^{2\pi i} e^{-3\pi i/5} = e^{-3\pi i/5}.$$

Destes dois pares de raízes, apenas $e^{\pm\pi i/5}$ possui parte real positiva. Lembre-se que a parte real de a é igual a $q/2$, a menos de conjugações complexas, obtemos o resultado final:

$$a = e^{\pi i/5} \quad e \quad b = e^{2\pi i/5}.$$

Assim como ocorreu no caso das flechas α , as escolhas de base também podem alterar os valores exatos de a e b . Em particular, as escolhas de base quando calculamos α também influenciam o resultado aqui, por meio do valor de q . Escolhendo outras bases podemos ter que $a = e^{3\pi i/5}$ a menos de conjugação complexa [42]. De todo modo, todos os resultados são conectados por mudanças de base. Com isto, temos explicitamente:

$$\alpha_{\tau,\tau,\tau} = \begin{pmatrix} q & \sqrt{-q}e^{-i\theta} & 0 \\ \sqrt{-q}e^{i\theta} & -q & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad e \quad \sigma_{\tau,\tau} = \begin{pmatrix} e^{\pi i/5} & 0 \\ 0 & e^{2\pi i/5} \end{pmatrix},$$

onde $q = -(1 + \sqrt{5})/2$, $\theta \in \mathbb{R}$ e $\sigma^{\tau,\tau}$ é definida a menos de conjugações complexas. Com isto, determinamos praticamente toda a estrutura cinemática dos anyons de Fibonacci.

Os aspectos restantes a serem calculados dizem respeito à rigidez e à estrutura de fita de \mathfrak{F} . Não os calcularemos aqui, porém notamos que o processo é análogo. Basta que calculemos a torção e as flechas i e e para τ . Em particular, lembre que τ é seu próprio dual. Lembrando que as condições que devem ser usadas não são somente aquelas que definem as torções θ e as flechas i e e separadamente, mas sim a condição de que \mathfrak{F} possua estrutura de fita, isto é, as torções são compatíveis com os duais em \mathfrak{F} . De todo modo, apenas α e σ são suficientes para capturar quase completamente o comportamento de anyons. Em particular, nos permitem ver como o anyons de Fibonacci podem ser usados

para realizar computações quânticas topológica.

4.7 – Computação quântica

Neste ponto, finalizamos o conteúdo deste trabalho. Antes de passarmos às considerações finais, olhemos novamente para o que obtivemos agora que terminamos. Primeiro, como vimos, as flechas α e σ correspondem a elementos do grupo de tranças que atuam sobre os anyons. Em particular $\sigma_{\tau,\tau}$ correspondem aos geradores do grupo de tranças sobre dois anyons τ . Olhando para esta flecha, vemos que ela corresponde a uma transformação bidimensional que atua sobre os anyons. Isto é exatamente o comportamento que anunciamos sobre anyons não abelianos: que permutá-los pode gerar uma rotação sobre a função de onda e não meramente uma fase. Além disso, se considerarmos mais anyons, a dimensão destes operadores aumenta ainda mais. Estas transformações podem ser usadas para realizar computações quânticas com anyons.

Posto de modo mais simples, uma computação é uma sequência específica de operações que atuam sobre um conjunto inicial de informações, a entrada, e produz outras informações, a saída. Qual sequência de operações é realizada pode depender da entrada, além de que existem entradas válidas e não válidas para certas computações. Por exemplo, não podemos somar as quantidades a e b se houver uma definição de como fazer isto. Em computadores clássicos, a unidade básica de informação é armazenada no conjunto $\{0, 1\}$ e uma certa informação corresponde a uma palavra formada a partir deste conjunto. Isto é, uma sequência $b_0 b_1 b_2 \dots$ de elementos de $\{0, 1\}$. Por exemplo, a sequência 1000 poderia representar o número 8 ou poderia representar outra coisa, isto depende de como a informação é codificada. Outra possível interpretação de 1000 poderia ser o número -7 .

Similarmente, os detalhes de como esta informação é armazenada fisicamente variam, podendo ser na orientações de domínios magnéticos ou na presença ou não de corrente elétrica. Também, como se realiza operações sobre a informação armazenada varia. Apesar disso, podemos pensar qualquer operação como constituída a partir de algum conjunto de operações básicas. Por exemplo, denotando, $\{0, 1\}$ por \mathbb{F}_2 , poderíamos escolher como operações básicas o conjunto formado por $\wedge : \mathbb{F}_2 \times \mathbb{F}_2 \rightarrow \mathbb{F}_2$ dado por $b_1 \wedge b_2 = 1 \iff b_1 = b_2 = 1$, $\vee : \mathbb{F}_2 \times \mathbb{F}_2 \rightarrow \mathbb{F}_2$ dada por $b_1 \vee b_2 = 0 \iff b_1 = b_2 = 0$ e $\neg : \mathbb{F}_2 \rightarrow \mathbb{F}_2$ dado por $\neg b = 1 - b$. Estas operações também chamadas de portas lógicas, são suficientes para produzir qualquer transformação que desejemos em uma palavras formadas a partir de elementos de \mathbb{F}_2 [48]. Dizemos que são universais. Outros nomes para estas operações são as portas lógicas AND, OR e NOT.

Em computadores clássicos, estas são as, ou um subconjunto das operações básicas presentes. Se 1 indica a presença de corrente elétrica e 0 a ausência, então estas operações são implementadas como circuitos elétricos. A porta \wedge é um circuito que com dois fios de entrada e um de saída e que só permite a passagem de corrente na saída se houver

corrente nas duas entradas. A porta \vee faz o contrário e permite passagem de corrente em qualquer situação exceto o caso onde não há corrente em nenhum dos fios de entrada. A porta \neg possui um fio de entrada e um de saída e se não há corrente chegando, ela gera corrente e se há, impede sua passagem. A partir destes circuitos básicos, são construídos circuitos mais elaborados que realizam operações mais complexas.

A relevância disto para computação quântica é que a situação é análoga. Notamos que esta conclusão não é tão direta quanto parece [48, 49, 50]. A unidade básica de informação quântica seria um qbit, de *quantum bit*. Podemos pensar um qbit como um espaço de Hilbert bidimensional. Os vetores da base fazem o papel de 0 e 1 de \mathbb{F}_2 . Se quisermos realizar operações nestes qbits ou em coleções de qbits, precisamos implementar transformações unitárias sobre eles. A unitariedade sendo exigida pelos postulados da mecânica quântica. Se conseguirmos obter um conjunto de transformações unitárias que seja universal, assim como o conjunto $\{\wedge, \vee, \neg\}$ é para \mathbb{F}_2^* , então poderemos implementar qualquer transformação unitária sobre um dado conjunto de qbits e, conseqüentemente, poderemos realizar computações. Note \mathbb{F}_2^* é o conjunto de todas as palavras formadas a partir de \mathbb{F}_2 .

O papel de anyons nesta história, em especial anyons de Fibonacci, é que as flechas α e σ são suficientes para gerar um conjunto de transformações universal, dado que tenhamos um certo número mínimo de anyons no sistema [40]. O diferencial de anyons, em relação a outras propostas de computação quântica, no entanto, é sua natureza topológica. O trançamento de anyons e sua fusão decorre de propriedades topológicas globais do espaço, como vimos. Assim, um sistema de anyons é muito mais robusto contra perturbações externas. Em um sistema clássico, podemos detectar a presença ou ausência de corrente elétrica com uma interferência que é geralmente negligenciável. Em um sistema quântica, geralmente a interferência causada pelo processo de medição é apreciável. Por exemplo, se desejarmos medir a componente de spin de uma partícula, há uma chance razoável de que a alteremos no processo. Por outro lado, se desejarmos saber o estado de um sistema de anyons, devemos olhar para dois ou mais anyons presentes [15]. Não podemos realizar medições no sistema — acessar a informação — olhando apenas para um anyon.

Este fato confere a anyons uma robustez muito almejada. Primeiro, como o trançamento de anyons depende de propriedades topológicas apenas, se durante o processo físico de permutá-los a “trajetória” dos anyons sofrer pequenas perturbações o resultado final não será afetado. Assim, os processos de computação são, em si, relativamente robustos. Contraste isto com a necessidade de manter o spin de uma partícula sempre na mesma orientação durante uma computação. É uma tarefa relativamente difícil já que flutuações térmicas e outras flutuações externas poderiam afetar isto além da própria interação com outras partículas. Além disso, para anyons, os estados da base são a carga topológica do sistema. Isto quer dizer que fixamos um resultado de fusão para o sistema, por exemplo, fixamos $\tau^{\otimes 4} = \tau$. Os vetores da base seriam todas as possibilidades distintas de fusão

destes quatro anyons de se fundir em τ . Se desejarmos investigar em qual estado o sistema de fato está, devemos verificar o resultado da fusão de todos os anyons. Assim, é altamente improvável que uma perturbação externa realize uma medida do sistema, interferindo com a informação armazenada.

Paramos por aqui com esta discussão. Há muito o que se falar e estudar sobre a computação topológica com anyons, mesmo se assumíssemos que podemos criar e manipular anyons a vontade. Inclusive, este é o outro lado da moeda. Todas estas propriedades são bastante almeçadas, mas até agora provou-se difícil obter anyons experimentalmente, quanto mais manipulá-los para realizar computações. Além disto, sistemas de computação anyônica estão sujeitos a outros tipos de erros, como erros de vazamento [40], onde mesmo se fixarmos a carga de um sistema, o resultado da fusão ainda pode ser correspondente a outra carga, vazando para outro subespaço do espaço total. Por fim, a descrição categorial de anyons nos fornece uma ferramenta extremamente útil da perspectiva de computação. Podemos obter de modo relativamente simples as expressões para as transformações às quais estão sujeitos quando os trançamos. E obtemos esta informação basicamente a partir da lista de anyons e de suas regras de fusão. Neste sentido, a descrição categorial é uma ferramenta poderosa para investigar possíveis modelos de anyons, além da sua conexão com teoria de computação que pode ajudar a desenvolver ainda mais o aspecto computacional.

Capítulo 5

Conclusão

Como o título sugere, este trabalho trata-se de um ensaio, um gênero textual que se propõe mais como uma reflexão sobre um dado assunto sem pretender esgotá-lo. Neste sentido, como pode ser observado ao longo dos três capítulos de desenvolvimento, não há muito sobre o que se concluir e, ao mesmo, tempo o capítulo de conclusões é o mais importante. Neste capítulo revisaremos todo o desenvolvimento feito nos três capítulos visando apontar precisamente os pontos que foram deixados em aberto. Assim, trataremos de quais seriam os próximos passos de estudo e pesquisa no assunto, levantando questões sobre o que seriam perspectivas futuras e sobre o potencial da abordagem categorial para a descrição de anyons em particular.

A natureza ensaística deste trabalho também se presta bem a uma “abordagem didática” sobre o assunto, com que entendemos que pode ser considerado como um trabalho introdutório sobre o assunto e que é relativamente acessível. Em particular a apresentação sobre categorias no capítulo 3 é tida como um ponto forte já que teoria de categorias não é um assunto com que se costuma ter contato em cursos de graduação, ou mesmo pós-graduação em física. Por outro lado, é um assunto que é bastante útil para questões de computação quântica por possuir uma interface grande com ciência da computação [17] e simultaneamente por fornecer uma descrição de modelos anyônicos, particularmente própria para questões de computação, como vimos.

Para que a discussão sobre a exposição feita no desenvolvimento tenha sentido, precisamos contextualizá-la, situando-a no contexto maior da pesquisa sobre anyons e computação quântica com anyons. Depois, tratamos das omissões que foram feitas neste trabalho, já que é da natureza de qualquer trabalho acadêmico limitar seu escopo ou arriscar tornar-se infinito. Acreditamos que ao pontuar alguns tópicos especialmente relevantes e que não foram tratados aqui, fique mais nítido qual a posição que a abordagem categorial ocupa em relação ao resto da pesquisa em anyons. Após isto, chegaremos, finalmente, às considerações finais.

5.1 - Síntese

O propósito desta seção é “tirar o zoom” deste trabalho e comentar sobre como os vários elementos abordados em ao longo do desenvolvimento se encaixam e quais papéis desempenham. Como cada capítulo entra em detalhes sobre um dado assunto, pode-se perder de vista o contexto geral, especialmente se não houver familiaridade com o assunto.

O ponto de partida é o capítulo 2, Sobre anyons. O propósito deste capítulo é apresentar o problema da indistinguibilidade e introduzir anyons a partir dele. Como discutido, existem nuances na própria definição de indistinguibilidade de partículas e no formalismo matemático para lidar com tais partículas. Em particular, um tratamento ingênuo do assunto é completamente insensível à dimensionalidade do sistema de modo a ser incapaz de descrever anyons.

Uma vez tratado este problema, a ideia de anyons surge naturalmente e o resto do capítulo é dedicado à discutir alguns modos de quantizar anyons e suas propriedades mais básicas. Assim, o capítulo 2 é a fundação deste trabalho. Seu papel é estabelecer a existência de anyons e algumas maneiras de descrevê-los quanticamente. Apesar de não ser estritamente necessária, acreditamos que a discussão mais aprofundada sobre indistinguibilidade possui valor tanto histórico quanto didático.

No capítulo Sobre categorias, o foco é alterado para o básico de teoria de categorias, necessário para a descrição de anyons abordada neste trabalho. Este capítulo é essencialmente um apêndice, porém está inserido no corpo do texto devido, principalmente, ao fato de que teoria de categorias não é um assunto comum em cursos graduação ou pós-graduação em física. Como um apêndice, sua função é fornecer definições e resultados que serão utilizados posteriormente.

O capítulo Sobre anyons e categorias é o cerne do trabalho. Neste ponto estão estabelecidas a existência e algumas propriedades de anyons e resultados básicos sobre teoria de categorias. Neste capítulo, então, alavanca-se o uso de categorias para construir uma certa descrição de anyons, a chamada de descrição categorial. Nesta, anyons são descritos por meio de estruturas algébricas que são, conceitualmente, bastante simples, apesar dos detalhes técnicos poderem ser bastante complicados.

Duas coisas devem ser feitas para construir esta estrutura algébrica, primeiro, devemos delimitar seus elementos e depois as operações possíveis entre elementos. Aqui se revela a primeira vantagem de categorias. As flechas em uma categoria são, em geral, a parte mais relevante em seu estudo. Assim, os objetos podem ser bastante simples, possivelmente até sem qualquer estrutura interna enquanto toda a complexidade é carregada pelas flechas. Com isto o primeiro passo é simples: os anyons, ou melhor, tipos de anyons, são objetos de uma dada categoria \mathfrak{A} .

O segundo passo, delimitar as operações que podem ser realizadas com estes objetos, então ocupa o cerne do capítulo. Como discutido, existem duas operações que devem ser

definidas formalmente para anyons, a operação de fusão, representada por \otimes e a soma, \oplus necessária para que faça sentido falar de decomposição de anyons em tipos irreduzíveis, em analogia com o momento angular. Além disto, é necessário também introduzir na categoria informações sobre o grupo de tranças.

Cada uma destas exigências são introduzidas gradualmente. Primeiro, a fusão é introduzida por meio da exigência que a categoria \mathfrak{A} seja monoidal. Em seguida as informações sobre o grupo de tranças são inseridas exigindo-se que, além de monoidal, a categoria seja trançada. Por fim, a operação \oplus é formalizada exigindo-se que a categoria \mathfrak{A} seja, também, semissimples. Outras exigências também são formalizadas impondo-se os demais axiomas dispostos no capítulo.

Deste modo o trabalho realizado no capítulo 4 é essencialmente um de modelagem física. Por exemplo, ao tratar de cinemática básica, uma das primeiras coisas discutidas é que posição, velocidade e aceleração são descritas adequadamente por vetores. As condições sobre estas quantidades são: devem possuir módulo, direção e sentido. Todas estas coisas podem ser capturadas exigindo-se que estas quantidades sejam descritas por elementos de um espaço que satisfaça aos oito axiomas de um espaço vetorial. A situação é idêntica neste trabalho. Por fim, o último tópico do capítulo é um exemplo para ilustrar que os axiomas enunciados de fato são suficientes para descrever anyons.

A síntese e o produto deste trabalho então é esta: a construção da descrição categorial de anyons. Novamente, ao se aprofundar nos detalhes técnicos é fácil perder de vista o todo. Acredito que esta síntese seja o suficiente para explicar como todas as partes de encaixam e o porquê de estarem presente neste trabalho. Passemos agora à discussão sobre o escopo.

5.2 - Sobre escopo

Começamos por revisitar, mais uma vez, o primeiro capítulo do desenvolvimento: Sobre anyons. O primeiro ponto da exposição que fizemos que possui mais discussões do que aparenta trata da exclusão da diagonal generalizada

$$\Delta = \left\{ x \in X^n \mid x_i = x_j, \text{ para algum } i \neq j, i, j = 1, \dots, n \right\}$$

na construção do espaço de configurações \mathcal{C}_n^d . Existe alguma discussão se estes pontos devem ser removidos ou não [51]. Para nós, este não é um ponto de interesse e o consideramos como uma questão técnica do processo de quantização. Em termos clássicos, é evidente que partículas não podem ocupar a mesma posição ao mesmo tempo, mas isto não é verdade em termos quânticos. Em geral, a função de onda de duas partículas se sobrepõem e no caso de bósons, eles podem, de fato, ocupar o mesmo estado, por exemplo, posição, ao mesmo tempo.

Ainda tratando do processo de quantização, mencionamos brevemente que este processo corresponde, aproximadamente ao de quantização geométrica. Evitamos nos aprofundar em technicalidades do processo de quantização por dois motivos: o primeiro é que as desconhecemos e o segundo é que no que diz respeito aos aspectos que conhecemos, acreditamos que não acrescentaria muito em termos de entender anyons. Dito isto, um tópico particularmente interessante e que não abordamos é a quantização de anyons não abelianos. Ao final do capítulo dissemos que é possível obter estatísticas não abelianas se considerarmos representações de dimensões superiores do grupo de tranças. Estas representações podem surgir se existirem estados degenerados no sistema.

Neste aspecto, tratar das technicalidades do processo de quantização, além de escolher outros espaços, não só o euclidiano para realizá-la pode render resultados interessantes. Em particular, anyons não abelianos são descritos exclusivamente por meio de teoria de campos ou teoria de categorias. Neste sentido, uma descrição destes anyons por meio de mecânica quântica básica, se possível, seria bastante útil, no mínimo para propósitos didáticos. Anyons não abelianos são um assunto notoriamente complexo porque exigem ferramentas bastante avançadas, como a extensão deste trabalho sugere. Seria bastante útil se pudéssemos introduzi-los de maneira razoavelmente formal e técnica usando um ferramental no mesmo nível do capítulo 2, mesmo que não houvesse muita possibilidade de aplicação prática para problemas de pesquisa.

O último ponto a tratar relaciona-se à escolha que fizemos de tomar a curvatura da conexão, o tensor \mathcal{F} como identicamente nulo. Como discutido, isso corresponde, basicamente, a uma escolha de calibre e nos permitiu chegar aos resultados mostrados que capturam o comportamento estatístico de anyons. No entanto, como mencionamos posteriormente a outra escolha, não zerar \mathcal{F} , nos fornece uma perspectiva que sob certo sentido é mais prática. Em particular, elas nos permite realizar cálculos “reais” sobre anyons [32] e também se presta a uma interpretação mais conectada com a realidade de anyons — a de Wilczek —, como partículas emergentes em sistemas de matéria condensada.

Esta escolha tem sua origem na escolha de escopo deste trabalho. Ele é voltado para a descrição categorial e não a descrição por meio de teorias de campos. A descrição usando teorias de campo, tanto teorias topológicas quanto teorias de campos conformes é certamente o método dominante de estudar anyons. Basicamente porque, como eles só podem existir em sistemas de matéria condensada, em virtude de vivermos em um universo com três dimensões espaciais, esta é a abordagem mais útil. Sistemas de matéria condensada são bem descritos por teorias de campos por sua natureza e assim, usar este formalismo para tratar de anyons é o modo mais direto de estudá-los. O formalismo categorial é um tanto removido do reino da praticidade neste sentido. Suas vantagens são outras.

Passando para o segundo capítulo do desenvolvimento, Sobre categorias, a situação é mais simples. Este capítulo essencialmente desempenha o papel de um apêndice. Seu pro-

pósito é introduzir o vocabulário básico de teoria de categorias. Sendo assim, é limitado ao mínimo necessário para que seja possível acompanhar o desenvolvimento no capítulo seguinte. Escolhemos, no entanto por não torná-lo demasiadamente enxuto, devido ao nível de abstração presente em teoria de categorias. Por mais que as ideias possam ser relativamente simples de entender, é necessário experiência em outros tópicos da matemática, particularmente álgebra abstrata, para entender sua real utilidade.

Por outro lado, não abordamos outros tópicos, presentes em um curso básico de teoria de categorias como o lema de Yoneda ou a noção de adjunção, que poderiam ser úteis, como o lema de Yoneda foi, para não fugir do escopo do trabalho. Um apêndice não possui a função de abarcar todo o assunto, apenas introduzir uma espécie de mínimo necessário para o trabalho. Isto não significa que outras ideias não possam ser úteis ou auxiliar na compreensão. Por fim, este capítulo foi apresentado como um capítulo genuíno e não como um apêndice em vista da sua importância para o trabalho. Primeiro, porque este é um trabalho especificamente sobre a descrição categorial de anyons e segundo, porque teoria de categorias não é um assunto presente em currículos de graduação ou pós-graduação em física, público para o qual este trabalho é direcionado.

Por fim, temos o terceiro capítulo do desenvolvimento: Sobre anyons e categorias. Como discutimos no início dele, seu propósito é combinar os resultados do primeiro e segundo capítulos do desenvolvimento e obter uma descrição categorial de anyons. Para fazer isto, são necessários vários novos conceitos de teoria de categorias que foram introduzidos e discutidos na medida do necessário. Basicamente, isto foi feito porque um tratamento detalhado de todos estes conceitos demandaria muito mais espaço e tempo, na forma de número de páginas e tempo de leitura, especialmente dado que são assuntos complexos e o retorno não seria muito significativo.

Comparamos esta situação a como novas ideias matemáticas são usualmente introduzidas em textos de física. Geralmente elas são trabalhadas conforme a necessidade surge de modo a melhor descrever o problema físico em questão. Gradualmente, conforme se estudam mais problemas ou se aprofunda mais em um dado problema, aprofunda-se também nestes conceitos. Investigar o conceito matemático por si só, separado de sua aplicação, embora possa ser divertido, não necessariamente se traduzirá em um melhor entendimento do problema físico. É uma situação análoga a como estudar análise real, embora melhore seus conhecimentos sobre cálculo, não necessariamente ajuda os conhecimentos de física, especialmente física básica. Evidentemente, análise pode te ajudar em coisas mais avançadas, o que nos leva ao próximo ponto.

A apresentação que fizemos sobre a descrição categorial de anyons é um tanto artificial. Assim como teoria de categorias surgiu a partir de topologia algébrica como uma forma de tornar mais nítidas e organizar as ideias envolvidas, a descrição de anyons usando categorias surgiu a partir das descrições via teorias de campos. Em particular, cresceram a partir de estudos das estruturas algébricas envolvidas em teorias de campos de anyons.

Neste sentido, a exposição feita carece de contextualização de algumas ideias, em particular, do axioma de modularidade, por exemplo. Sua presença é necessária para que as categorias que descrevem anyons sejam equivalentes às teorias de campos que os descrevem. Porém, o que ele realmente significa e porque é necessário é difícil de contextualizar, diferente do que acontece com a estrutura de tranças, por exemplo, que está presente para que as transformações sobre os anyons, as flechas na categoria, equivalham a elementos do grupo de tranças. Similarmente, existem outros aspectos da descrição categorial que não abordamos.

Deste modo a principal omissão do terceiro capítulo do desenvolvimento é este contexto. Evidentemente, isto é uma consequência da delimitação de escopo do trabalho. Uma introdução à descrição de anyons por meio de teoria de campos resultaria em um trabalho tão extenso quanto este. Além disto, dado que um dos propósitos deste trabalho é ser introdutório, pouco sentido há em assumir que o leitor possua familiaridade com teorias topológicas de campos anyônicos ou teorias de campos conformes. Mencionamos isto aqui apenas para fornecer um pouco mais de contexto sobre o assunto e para apontar ao leitor os tópicos que foram deixados de fora, ou mesmo, que existem mais coisas a serem estudadas sobre um ponto em particular da descrição.

Com isto, finalizamos esta seção tratando do escopo deste trabalho e as omissões que foram feitas. Notamos que apontamos as omissões não como uma espécie de defeito deste trabalho, apenas como um indicativo de pontos onde existem mais coisas a serem estudadas e mais material disponível sobre. Como dissemos no início, um trabalho que não delimita seu escopo corre o risco de se tornar infinito, afinal, sempre há mais o que se estudar. Passamos então às considerações finais.

5.3 - Considerações finais

As considerações finais deste trabalho consistem, essencialmente em perspectivas futuras. Como dissemos, a descrição categoria de anyons foi algo que surgiu a partir da descrição via teorias de campos. Apesar disto, é uma descrição que é própria para certos aspectos de computação quântica, em particular por permitir calcular as transformações sobre anyons de maneira relativamente simples, partindo simplesmente da lista de anyons básicos do sistema e suas regras de fusão. Além disso, esta descrição torna fácil testar novos modelos de anyons. Qualquer lista de anyons e qualquer conjunto de regras de fusão, a princípio podem descrever algum modelo de anyons se satisfizerem aos axiomas que enunciamos. Assim, podemos produzir modelos de anyons e verificar sua validade, especialmente para computação quântica, de modo significativamente mais simples do que faríamos se tivéssemos que construir hamiltonianas para estes modelos. Especialmente dado que anyons são emergentes.

Em particular isto acontece porque, sendo partículas emergentes na realidade, não é

possível garantir que um dado termo em uma hamiltoniana irá dar origem a um dado tipo de anyon. O propósito inicial deste trabalho e objetivo de futuras continuações, era começar a investigar, então, o que podemos dizer a respeito da teoria de campos que descreve um dado sistema de anyons a partir de sua descrição categorial. A formulação categorial é relativamente afastada do aspecto físico do problema. Isto é uma desvantagem em se tratando de resolver problemas práticos, mas como objeto de estudo pode ser uma enorme vantagem pois ela captura apenas o essencial de anyons, suas estatísticas. Assim, é plenamente possível imaginar um dado comportamento estatístico e obter uma categoria que descreve partículas com este comportamento.

Se estas partículas são factíveis ou não, depende se satisfazem os axiomas ou não. Ao menos em princípio. Sendo assim, se for possível extrair informações físicas a partir da descrição categorial, então esta descrição pode se tornar uma forte ferramenta para classificar e organizar tipos de anyons distintos. O melhor cenário possível seria aquele onde obtemos uma “tradução” termo a termo do comportamento estatístico em uma hamiltoniana. O mais razoável de se esperar, no entanto, é que possamos falar que certo tipo de interação pode dar origem a comportamento não abeliano enquanto outros tipos o proíbem. Isto seria útil na busca por materiais que sustentam a existência de anyons, pois poderíamos descartar, de imediato, materiais que não possuem uma dada interação.

Outro tópico de interesse é se existe uma classificação de tipos de anyons. Segundo a exposição que fizemos, qualquer modelo anyônico deve satisfazer aos axiomas que enunciamos. Porém, entre as categorias tensoriais modulares que descrevem anyons, existem mais subdivisões? Por exemplo, existem “classes” distintas de regras de fusão? Por exemplo, poderíamos gerar divisões arbitrárias de tipos de regras de fusão, como: regras de fusão onde todos os produtos só possuem dois resultados possíveis. Regras onde possuem três. Regras onde um dos resultados é necessariamente o vácuo. Há questão, no entanto, é se é possível obter alguma informação significativa a partir de alguma classificação. No melhor dos casos, poderíamos dizer, este tipo de interação permite estes tipos de regras de fusão e proíbe estes outros tipos. Isto, novamente, ajuda na busca por materiais úteis, por exemplo, para a construção de computadores quânticos a base de anyons.

Aqui a natureza ensaística deste trabalho realmente se revela. Este trabalho é, essencialmente, uma plataforma a partir da qual podemos começar a investigar se as perguntas colocadas assim possuem respostas. Ele introduz a maior parte do ferramental necessário para entender o que significa a descrição categorial de anyons e como ela funciona. Como apontamos, existem tópicos mais profundos sobre esta descrição que não abordamos. Apesar disto, podemos concluir que este trabalho cumpre sua função enquanto um ensaio sobre o assunto, uma discussão ou exposição sobre o mesmo e também como uma introdução. Conhecimentos de álgebra abstrata formam uma espécie de pré-requisito suave para este trabalho, no entanto, não é exigido realmente muito mais do que o que é esperado que alguém no final da graduação ou início da pós-graduação já não saiba.

Referências

- [1] Michael G. G. Laidlaw and Cécile Morette DeWitt. Feynman functional integrals for systems of indistinguishable particles. *Phys. Rev. D*, 3:1375–1378, Mar 1971.
- [2] J. M. Leinaas and J. Myrheim. On the theory of identical particles. *Nuovo Cimento B Serie*, 37(1):1–23, January 1977.
- [3] Frank Wilczek. Quantum mechanics of fractional-spin particles. *Phys. Rev. Lett.*, 49:957–959, Oct 1982.
- [4] Frank Wilczek. Magnetic flux, angular momentum, and statistics. *Phys. Rev. Lett.*, 48:1144–1146, Apr 1982.
- [5] Yong-Shi Wu. General theory for quantum statistics in two dimensions. *Phys. Rev. Lett.*, 52:2103–2106, Jun 1984.
- [6] H. Bartolomei, M. Kumar, R. Bisognin, A. Marguerite, J.-M. Berroir, E. Bocquillon, B. Plaçais, A. Cavanna, Q. Dong, U. Gennser, Y. Jin, and G. Fève. Fractional statistics in anyon collisions. *Science*, 368(6487):173–177, 2020.
- [7] James Nakamura, Shuang Liang, Geoffrey C. Gardner, and Michael J. Manfra. Direct observation of anyonic braiding statistics at the $\nu=1/3$ fractional quantum hall state. <https://arxiv.org/abs/2006.14115>, 2020.
- [8] K. v. Klitzing, G. Dorda, and M. Pepper. New method for high-accuracy determination of the fine-structure constant based on quantized hall resistance. *Phys. Rev. Lett.*, 45:494–497, Aug 1980.
- [9] R. B. Laughlin. Anomalous quantum hall effect: An incompressible quantum fluid with fractionally charged excitations. *Phys. Rev. Lett.*, 50:1395–1398, May 1983.
- [10] H. L. Stormer and D. C. Tsui. The quantized hall effect. *Science*, 220(4603):1241–1246, 1983.
- [11] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical hamiltonian model of computers as represented by turing machines. *Journal of statistical physics*, 22(5):563–591, 1980.
- [12] Richard P Feynman. Simulating physics with computers. *Int. J. Theor. Phys*, 21(6/7), 1982.

- [13] David Deutsch. Quantum theory, the church–turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.
- [14] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.
- [15] Chetan Nayak, Steven H Simon, Ady Stern, Michael Freedman, and Sankar Das Sarma. Non-abelian anyons and topological quantum computation. *Reviews of Modern Physics*, 80(3):1083, 2008.
- [16] Samuel Eilenberg and Saunders MacLane. General theory of natural equivalences. *Transactions of the American Mathematical Society*, 58(2):231–294, 1945.
- [17] B.C. Pierce, B. C, B.C. Pierce, M.R. Garey, and A. Meyer. *Basic Category Theory for Computer Scientists*. Foundations of computing series: research reports and notes. MIT Press, 1991.
- [18] Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloe. *Quantum Mechanics, Volume 2*, volume 2. 1986.
- [19] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. 2017.
- [20] D.J. Griffiths. *Introduction to Quantum Mechanics*. Cambridge University Press, 2017.
- [21] Loring W. Tu. *An introduction to manifolds*. Universitext. Springer, New York, 2nd ed edition, 2011.
- [22] J. F. Humphreys. *A course in group theory*. Oxford science publications. Oxford University Press, Oxford ; New York, 1996.
- [23] R. Mirman. Experimental meaning of the concept of identical particles. *Il Nuovo Cimento B (1971-1996)*, 18(1):110–122, 1973.
- [24] A. Messiah. *Quantum Mechanics*. Number v. 2 in Dover books on physics. Dover Publications, 1999.
- [25] L.I. Schiff. *Quantum Mechanics*. International series in pure and applied physics. McGraw-Hill, 1955.
- [26] Alberto Lerda. *Anyons: Quantum Mechanics of Particles with Fractional Statistics*. Lecture Notes in Physics Monographs. Springer Berlin Heidelberg, 1992.
- [27] A. Khare. *Fractional Statistics and Quantum Theory*. World Scientific, 2005.
- [28] P. Feyerabend. *Contra O Método*. UNESP, 2007.
- [29] J.M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer, 2003.
- [30] Jean-Marie Souriau. Quantification géométrique. applications. *Annales de l’I.H.P. Physique théorique*, 6(4):311–341, 1967.

- [31] J.M. Souriau, R.H. Cushman, C.H.C. Vries, C. Cushman, and G.M. Tuynman. *Structure of Dynamical Systems: A Symplectic View of Physics*. Progress in Mathematics. Birkhäuser Boston, 1997.
- [32] J. Myrheim. Anyons. In A. Comtet, T. Jolicœur, S. Ouvry, and F. David, editors, *Aspects topologiques de la physique en basse dimension. Topological aspects of low dimensional systems*, volume 69, pages 265–413. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.
- [33] Lawrence Schulman. A path integral for spin. *Phys. Rev.*, 176:1558–1569, Dec 1968.
- [34] Gregory Moore and Nicholas Read. Nonabelions in the fractional quantum hall effect. *Nuclear Physics B*, 360(2-3):362–396, 1991.
- [35] Alexei Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321(1):2–111, 2006.
- [36] Tom Leinster. *Basic Category Theory*. Cambridge studies in advanced mathematics 143. Cambridge University Press, 1 edition, 2014.
- [37] Steve Awodey. *Category Theory*. Oxford Logic Guides 52. Oxford University Press, 2nd edition, 2010.
- [38] Saunders Mac Lane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics. Springer, 2nd edition, 1998.
- [39] Bojko Bakalov and Alexander Kirillov Jr. *Lectures on tensor categories and modular functors*. University Lecture Series 021. American Mathematical Society, 2001.
- [40] Prakash Panangaden and Éric O. Paquette. A categorical presentation of quantum computation with anyons. In *New structures for Physics*, pages 983–1025. Springer, 2010.
- [41] Tom Leinster. Rethinking set theory. *The American Mathematical Monthly*, 121(5):403–415, 2014.
- [42] Andreas Blass and Yuri Gurevich. On quantum computation, anyons, and categories. In *Martin Davis on Computability, Computational Logic, and Mathematical Foundations*, pages 209–241. Springer, 2016.
- [43] V.G. Turaev. *Quantum Invariants of Knots and 3-manifolds*. De Gruyter studies in mathematics. New York, 2010.
- [44] Martti Karvonen. Biproducts without pointedness. <https://arxiv.org/abs/1801.06488>, 2020.
- [45] P. Freyd. *Abelian Categories: An Introduction to the Theory of Functors*. A Harper international edition. Harper & Row, 1964.
- [46] Kerstin Beer, Dmytro Bondarenko, Alexander Hahn, Maria Kalabakov, Nicole Knust, Laura Niermann, Tobias J Osborne, Christin Schridde, Stefan Seckmeyer, Deniz E Stiegemann, et al. From categories to anyons: a travelogue. *arXiv preprint arXiv:1811.06670*, 2018.

- [47] Simon Trebst, Matthias Troyer, Zhenghan Wang, and Andreas W. W. Ludwig. A short introduction to fibonacci anyon models. *Progress of Theoretical Physics Supplement*, 176:384–407, 2008.
- [48] M. Hirvensalo. *Quantum Computing*. Natural Computing Series. Springer Berlin Heidelberg, 2013.
- [49] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
- [50] A.Y. Kitaev, A. Shen, M.N. Vyalyi, and M.N. Vyalyi. *Classical and Quantum Computation*. Graduate studies in mathematics. American Mathematical Society, 2002.
- [51] M. Bourdeau and R. D. Sorkin. When can identical particles collide? *Phys. Rev. D*, 45:687–696, Jan 1992.

Apêndice A

Sobre topologia

Este apêndice trata de conceitos básicos de topologia, como o que é um espaço topológico, como se define funções contínuas neste espaço e diferentes maneiras de se construir um espaço topológico. O conteúdo deste apêndice, com exceção das seções sobre espaços quocientes conexidade não é estritamente necessário no corpo do trabalho. No entanto está sempre presente indiretamente. Por exemplo, \mathbb{R}^n é sempre considerado em sua topologia canônica, que é a topologia da ordem em \mathbb{R} e a do produto em \mathbb{R}^n , para $n > 1$. Por este motivo, escolhemos fazer este apêndice de modo a partir das ideias básicas até alcançar os resultados utilizados no trabalho. Neste sentido ele pode servir como uma breve introdução à topologia para leitores que não estão familiarizados com topologia básica.

Note que a exposição aqui feita não é, de modo algum exaustiva, tanto no escopo geral quando no deste trabalho. Isto é, existem conceitos que se relacionam com as ideias dispostas no corpo do texto e que não são mencionados aqui por não serem estritamente necessários. Dito isto, comecemos por tratar da definição básica de um espaço topológico.

A.1 - Espaços topológicos

O conceito de espaços topológicos cresceu a partir do estudo da reta real e de espaços euclidianos e do estudo de funções contínuas nestes espaços. A definição de um espaço topológico que hoje é padrão passou por um longo tempo de gestação. Vários matemáticos como Fréchet, Hausdorff e outros propuseram diferentes definições por um período de anos durante as primeiras décadas do século XX, mas levou um bom tempo até que se chegasse a uma que fosse mais adequada. A questão é que desejava-se uma definição que fosse o mais ampla o possível, de modo a incluir vários casos especiais que eram úteis – espaços euclidianos, espaços euclidianos de dimensão infinita e espaços de funções, por exemplo – e ao mesmo tempo, que fosse estreita o suficiente para que os teoremas padrões sobre estes espaços familiares pudessem valer para espaços topológicos em geral. Este é um problema sempre que se formula um novo conceito matemático: decidir o quão geral ele deve ser. A definição finalmente se acomodou no que parece ser uma um pouco abstrata, mas conforme trabalha-se as várias maneiras de construir espaços topológicos, adquire-se uma intuição sobre o que o conceito realmente significa e sobre as ideias que abrange.

Definição A.1: Uma **topologia** em um conjunto, X , e uma coleção, \mathcal{T} , de subconjuntos de X que possui as seguintes propriedades:

- i \emptyset e X estão em \mathcal{T} ;
- ii A união de elementos de qualquer subcoleção de \mathcal{T} está em \mathcal{T} ;
- iii A intersecção de qualquer subcoleção finita de \mathcal{T} está em \mathcal{T} .

Um conjunto X junto a uma topologia \mathcal{T} é dito um **espaço topológico**.

Propriamente, um espaço topológico é um par ordenado (X, \mathcal{T}) que consiste em um conjunto X e uma topologia \mathcal{T} sobre X . Frequentemente, no entanto, omite-se a menção a \mathcal{T} se não houver ambiguidade.

Definição A.2: Se X é um espaço topológico com topologia \mathcal{T} , um subconjunto $U \in X$ é dito um **conjunto aberto**, **subconjunto aberto** ou, simplesmente, **aberto** de X se $U \in \mathcal{T}$.

Usando esta terminologia, pode-se dizer que um espaço topológico é um conjunto X junto a uma coleção de subconjuntos de X , chamados conjuntos abertos, tais que \emptyset e X são abertos e qualquer união arbitrária ou intersecção finita de conjuntos abertos é também um conjunto aberta. Notamos aqui que o uso da palavra coleção é apenas por conveniência, tendo o mesmo significado de conjunto. A utilizamos apenas para distinguir o fato de que os elementos de uma topologia são, também, conjuntos. Vejamos agora alguns exemplos de topologias:

Exemplo A.1: Se X é um conjunto arbitrário, a coleção de todos os subconjuntos de X é uma topologia em X . Esta topologia é chamada **topologia discreta**. A coleção que consiste apenas de X e \emptyset é, também, uma topologia em X , chamada **topologia indiscreta** ou **topologia trivial**.

Exemplo A.2: Seja X um conjunto arbitrário e seja \mathcal{T}_f a coleção de todos os subconjuntos U de X tais que $X - U$ é, ou finito ou todo X . Então, \mathcal{T}_f é uma topologia em X , chamada **topologia dos complementos finitos** ou **topologia cofinita**. Tem-se que ambos X e \emptyset pertencem a \mathcal{T}_f , pois $X - X$ é finito e $X - \emptyset$ é todo X . Se $\{U_\alpha\}$ é uma família de elementos não vazios de \mathcal{T}_f , então:

$$X - \bigcup U_\alpha = \bigcap (X - U_\alpha).$$

Como cada conjunto $X - U_\alpha$ é finito e a intersecção arbitrária de conjuntos finitos é

finita, então $\bigcup U_\alpha \in \mathcal{T}_f$.

Além disto, se $\{U_i\}$ é uma coleção finita de elementos não vazios de \mathcal{T}_f , então:

$$X - \bigcap_{i=1}^n U_i = \bigcup_{i=1}^n (X - U_i).$$

Como cada conjunto $X - U_i$ é finito e a união finita de conjuntos finitos é finita tem-se que $\bigcap U_i \in \mathcal{T}_f$. Logo, \mathcal{T}_f satisfaz a todas as propriedades exigidas de uma topologia. Análoga a \mathcal{T}_f , existe também a topologia \mathcal{T}_c de todos os subconjuntos de U de algum conjunto X tais que $X - U$ ou é todo X ou é contável. \mathcal{T}_c é chamada **topologia dos complementos contáveis** ou **topologia cocontável**. A demonstração para \mathcal{T}_c é análoga à de \mathcal{T}_f .

Exemplo A.3: Seja X um conjunto de três elementos, $X = \{a, b, c\}$. Existem várias topologias possíveis para em X , algumas das quais estão indicadas esquematicamente na figura a seguir. O diagrama no canto superior direito, indica uma topologia onde os conjuntos abertos são X , \emptyset , $\{a, b\}$, $\{b\}$ e $\{b, c\}$. A topologia no canto superior esquerdo contém apenas X e \emptyset e a topologia no canto inferior direito contém todos os subconjuntos de X .

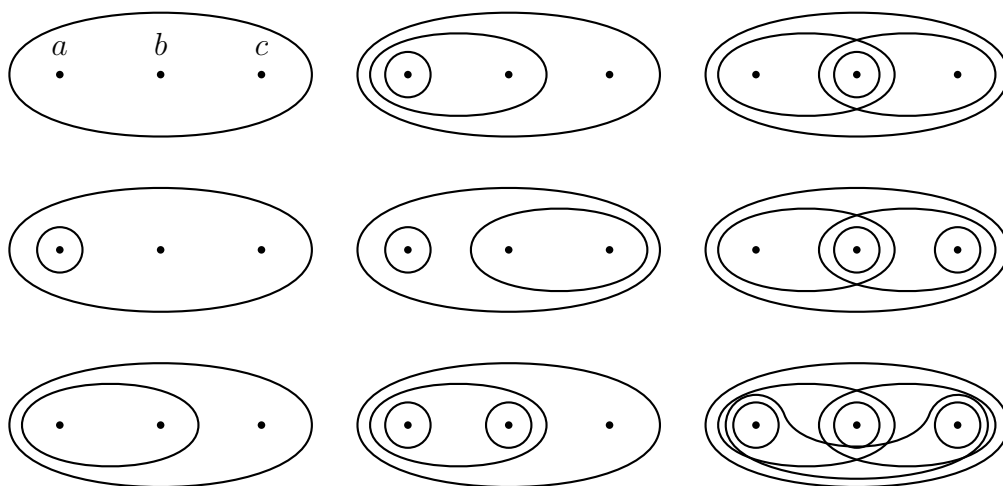


Figura A.1: Diferentes topologias para um conjunto de três elementos.

Note que um conjunto de apenas três elementos possui várias topologias possíveis, apesar disto, nem toda coleção possível de subconjuntos de X é uma topologia em X . Os dois diagramas a seguir não são topologias de X .



Figura A.2: Nenhuma destas coleções é uma topologia para este conjunto.

Bases

Nos exemplos anteriores, foi possível especificar a topologia ao descrever toda a coleção \mathcal{T} de conjuntos abertos. Geralmente, isto é muito difícil. Na maioria dos casos, especifica-se uma coleção menor de subconjuntos de X e define-se a topologia em termos destes conjuntos.

Definição A.3: Se X é um conjunto, uma **base** para uma topologia em X é uma coleção \mathcal{B} de subconjuntos de X , chamados **elementos da base** ou **elementos base**, tais que:

- i Para cada $x \in X$, existe pelo menos um elemento base, B , tal que $x \in B$;
- ii Se x pertence à interseção de dois elementos base B_1 e B_2 , então existe um elemento base B_3 contendo x tal que $B_3 \subset B_1 \cap B_2$.

Se \mathcal{B} satisfaz a estas duas condições, definimos a topologia \mathcal{T} **gerada** por \mathcal{B} da seguinte maneira: um subconjunto U de X é dito aberto em X – é um elemento de \mathcal{T} – se para cada $x \in U$, existe um elemento base $B \in \mathcal{B}$ tal que $x \in B$ e $B \subset U$. Note que cada elemento da base pertence a \mathcal{T} .

Verifiquemos se esta construção de fato gera uma topologia em X . De imediato $\emptyset \in \mathcal{T}$ por vacuidade. Do mesmo modo, $X \in \mathcal{T}$ já que para todo $x \in X$ existe pelo menos um elemento base, B , contendo x e contido em X . Agora, considere uma família de subconjuntos de X , $\{U_\alpha\}$, tais que $U_\alpha \in \mathcal{T}$, é necessário que o subconjunto $U = \bigcup U_\alpha$ pertença a \mathcal{T} . Dado $x \in U$, existe um índice α tal que $x \in U_\alpha$. Como U_α é aberto, existe um elemento base B tal que $x \in B$ e $B \subset U_\alpha$. Então, $x \in B$ e $B \subset U$, de modo que U é aberto por definição. Agora, considere dois elementos, $U_1, U_2 \in \mathcal{T}$. Devemos mostrar que $U_1 \cap U_2$ pertence a \mathcal{T} . Dado $x \in U_1 \cap U_2$, escolha um elemento base B_1 contendo x tal que $B_1 \subset U_1$ e também, um elemento base B_2 contendo x tal que $B_2 \subset U_2$. A segunda condição da definição de base garante que existe um elemento base B_3 contendo x tal que $B_3 \subset B_1 \cap B_2$. Então, $x \in B_3$ e $B_3 \subset U_1 \cap U_2$, de modo que $U_1 \cap U_2$ pertence a \mathcal{T} por definição. Este argumento está ilustrado na figura abaixo.

Por fim, suponha considere uma família finita de conjuntos abertos $\{U_i\}$ e suponha que $\bigcap_{i=1}^n U_i \in \mathcal{T}$, tem-se:

$$\bigcap_{i=1}^{n+1} U_i = \left(\bigcap_{i=1}^n U_i \right) \cup U_{n+1}.$$

Por hipótese $\bigcap_{i=1}^n U_i \in \mathcal{T}$ e, pelo resultado provado anteriormente, a interseção de dois conjuntos que pertencem a \mathcal{T} também pertence a \mathcal{T} . Pelo princípio da indução fica

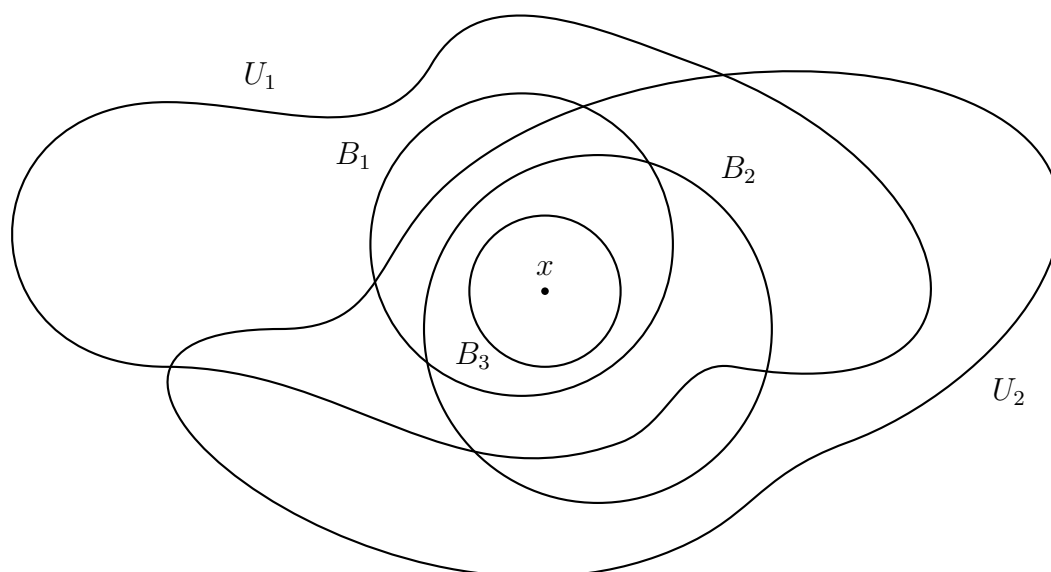


Figura A.3: $B_3 \subset B_1 \cap B_2$.

estabelecido que qualquer interseção finita de elementos de \mathcal{T} também pertence a \mathcal{T} . Deste modo, o processo descrito anteriormente gera, de fato, uma topologia em X a partir de uma base \mathcal{B} . Outra maneira de descrever a topologia gerada por uma base é descrita no seguinte lema:

Lema A.1: Seja X um conjunto e seja \mathcal{B} uma base para a topologia \mathcal{T} em X . Então, \mathcal{T} é a coleção de todas as uniões de elementos de \mathcal{B} .

Prova: Dada uma coleção de elementos de \mathcal{B} , eles também são elementos de \mathcal{T} e por \mathcal{T} ser uma topologia, qualquer união de tais elementos também está em \mathcal{T} . Por outro lado, dado $U \in \mathcal{T}$, escolhemos para cada $x \in U$ um elemento B_x de \mathcal{B} tal que $x \in B_x \subset U$. Então

$$U = \bigcup_{x \in U} B_x$$

de modo que U é uma união de elementos de \mathcal{B} .

□

Este lema afirma que todo aberto U de X pode ser expresso como uma união de elementos base. Esta expressão para U , no entanto, não é única. Assim, o uso do termo “base” em topologia difere drasticamente do uso deste termo em álgebra linear, por exemplo, onde a expressão de um vetor como combinação de vetores base é única. Considere agora alguns exemplos de bases para ilustrar o conceito.

Exemplo A.4: Se X é qualquer conjunto, a coleção de todos os subconjuntos de um elemento de X é uma base para a topologia discreta em X .

Exemplo A.5: Seja \mathcal{B} a coleção de todas as regiões circulares – interiores de círculos – no plano. Então \mathcal{B} satisfaz às duas condições de uma base. A imagem abaixo ilustra a segunda condição. Na topologia gerada por \mathcal{B} , um subconjunto U do plano é aberto se todo $x \in U$ está em alguma região circular contida em U .

Exemplo A.6: Seja \mathcal{B}' a coleção de todas as regiões retangulares – interiores de retângulos – no plano, onde os retângulos possuem lados paralelos aos eixos coordenados. Então \mathcal{B}' satisfaz às duas condições para uma base. A imagem abaixo ilustra a segunda condição. Neste caso, a condição é trivial pois a interseção de quaisquer dois elementos da base é, também, um elemento da base, ou vazia. A figura abaixo ilustra os elementos da base de \mathcal{B} e \mathcal{B}' .

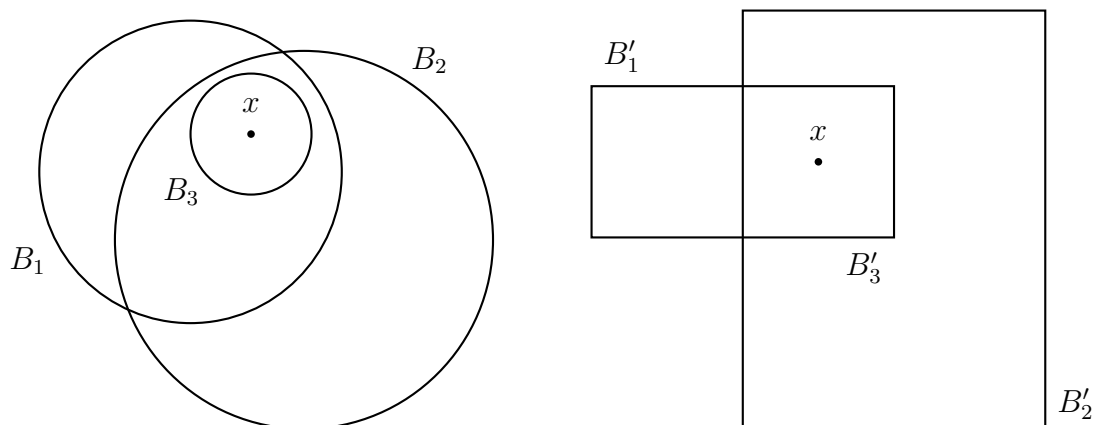


Figura A.4: Bases \mathcal{B} e \mathcal{B}' .

Até agora, foram descritos dois modos de se determinar uma topologia a partir de uma base. Algumas vezes, no entanto, é necessário realizar o processo inverso e determinar uma base a partir de uma topologia. Uma maneira frequentemente usada de fazer isto é a seguinte:

Lema A.2: Sejam X um espaço topológico e \mathcal{C} uma coleção de abertos em X tal que para cada aberto $U \subseteq X$ e para cada $x \in U$, existe um elemento $C \in \mathcal{C}$ tal que $x \in C \subset U$. Então \mathcal{C} é uma base para a topologia de X .

Prova: A primeira condição para uma base é simples, dado $x \in X$, como X é um aberto, por hipótese deve existir um elemento $C \in \mathcal{C}$ tal que $x \in C \subset X$. Para verificar a segunda condição, considere $x \in C_1 \cap C_2$, onde C_1 e C_2 são elementos de \mathcal{C} . Como C_1 e C_2 são abertos, $C_1 \cap C_2$ também o é. Portanto, existe um elemento $C_3 \in \mathcal{C}$ tal que $x \in C_3 \subset C_1 \cap C_2$. Agora, se \mathcal{T} é a topologia em X , resta mostrar

que a topologia \mathcal{T}' gerada por \mathcal{C} é igual a \mathcal{T} . Primeiro, note que se $U \in \mathcal{T}$ e $x \in U$, então, por hipótese, existe um elemento $C \in \mathcal{C}$ tal que $x \in C \subset U$ e portanto U pertence à topologia \mathcal{T}' . Por outro lado, se W pertence à topologia \mathcal{T}' , então W é uma união de elementos de \mathcal{C} , pelo lema A.1. Como cada elemento de \mathcal{C} está em \mathcal{T} e \mathcal{T} é uma topologia, W também está em \mathcal{T} .

□

Além do conceito de base, topologias também podem ser determinadas por meio de sub-bases. Como vimos, dado uma base, \mathcal{B} , uma topologia é gerada considerando-se uma coleção de uniões arbitrárias de elementos de \mathcal{B} . Por outro lado, pode-se descrever uma topologia considerando-se também as interseções finitas de elementos de uma coleção, \mathcal{S} , de conjuntos e uniões arbitrárias destas interseções. A coleção \mathcal{S} forma uma sub-base.

Definição A.4: Uma **sub-base**, \mathcal{S} , de uma topologia em X é uma coleção de subconjuntos de X cuja união é igual a X . A topologia gerada pela sub-base \mathcal{S} é definida como a coleção, \mathcal{T} , de todas as uniões de interseções finitas de elementos de \mathcal{S} .

Verifiquemos que, de fato, a coleção gerada por uniões arbitrárias de interseções finitas de elementos de uma sub-base, \mathcal{S} gera uma topologia. Para isto, basta mostrar que a coleção \mathcal{B} de todas as interseções finitas de elementos de \mathcal{S} é uma base. Deste modo a coleção \mathcal{T} de todas as uniões de elementos de \mathcal{B} é uma topologia pelo lema A.1. Considere, então $x \in X$, x pertence a algum elemento de \mathcal{S} e portanto a algum elemento de \mathcal{B} , satisfazendo à primeira condição de uma base. Além disso, sejam

$$B_1 = S_1 \cap \cdots \cap S_m \quad \text{e} \quad B_2 = S'_1 \cap \cdots \cap S'_n$$

dois elementos de \mathcal{B} . A interseção

$$B_1 \cap B_2 = (S_1 \cap \cdots \cap S_m) \cap (S'_1 \cap \cdots \cap S'_n)$$

é também uma interseção finita de elementos de \mathcal{S} pertencendo, portanto, a \mathcal{B} . Logo \mathcal{B} é uma base como queríamos.

Agora que estabelecemos o conceito de topologia, vejamos como construir topologias para diferentes espaços. Por exemplo, seria conveniente um método geral para se definir a topologia no produto $X \times Y$ a partir das topologias de X e Y , ou definir uma topologia em um subconjunto de um espaço a partir da topologia do deste espaço. Começaremos, no entanto, com um método que é bastante intuitivo e produz, por exemplo, a topologia canônica da reta real. Este método utiliza relações de ordem em um conjunto para construir uma topologia para o mesmo.

A.2 - Topologia da ordem

Definição A.5: Uma **ordem linear** ou **simples**, \prec em um conjunto X é uma relação binária que satisfaz às seguintes propriedades:

- i **Comparabilidade:** para todo a, b , tal que $a \neq b$ tem-se que ou $a \prec b$ ou $b \prec a$.
- ii **Irreflexividade:** para todo a , $a \not\prec a$.
- iii **Transitividade:** se $a \prec b$ e $b \prec c$, então $a \prec c$.

Para $a, b, c \in X$. Um par (X, \prec) , é chamado de **conjunto simplesmente ordenado** ou **linearmente ordenado**.

Definição A.6: Uma **ordem total**, \preceq em um conjunto X é uma relação binária que satisfaz às seguintes propriedades:

- i **Totalidade:** para todo a, b , $a \preceq b$ ou $b \preceq a$.
- ii **Anti-simetricidade:** se $a \preceq b$ e $b \preceq a$, então $a = b$.
- iii **Transitividade:** se $a \preceq b$ e $b \preceq c$, então $a \preceq c$.

Para $a, b, c \in X$. Um par (X, \preceq) , é chamado de **conjunto totalmente ordenado** ou de **cadeia**.

Repare que a propriedade de totalidade implica que cada elemento de X se relaciona com todos os outros elementos de X . Combinando a propriedade de totalidade com a de anti-simetria, obtem-se, que, como todo elemento deve se relacionar com qualquer outro, em particular, tomando $b = a$, $a \preceq a$, logo uma ordem total também é **reflexiva**.

Proposição A.1: Dada uma ordem simples, \prec em um conjunto X , a relação definida por:

$$\preceq = \{(a, b) \in X \times X \mid (a, b) \in \prec \vee a = b\}$$

é uma ordem total.

Prova: Deve-se mostrar que \preceq satisfaz às três propriedades de uma ordem total.

- i Primeiro observe que para todo $a \in X$, o par $(a, a) \in \preceq$. Ou seja, para todo $a \in X$, $a \preceq a$. Como \prec é uma ordem linear, para todo $a, b \in X$, $a \neq b$, tem-se

ou $a \prec b$ ou $b \prec a$. Juntando estes dois fatos, tem-se que para todo $a, b \in X$, $a \preceq b$ ou $b \preceq a$. Logo \preceq é total.

ii Em seguida, note que para todo $a, b, c \in X$ com $a \neq b \neq c$, se $a \prec b$ e $b \prec c$, então $a \prec c$ pela propriedade transitiva de \prec . Portanto, o mesmo vale para \preceq , já que $\prec \subset \preceq$. Agora, suponha que $a \preceq b$ e $b \preceq c$. Se $a = b$, então $a \preceq c$ porque $b \preceq c$ e $a = b$. Do mesmo modo se $b = c$, então $a \preceq c$ porque $a \preceq b$ e $b = c$. Por fim, se $a = c$, $a \preceq c$ pela definição de \preceq . Logo \preceq é transitiva.

iii Por fim, note que podemos escrever \preceq como

$$\preceq = \prec \cup \Delta,$$

onde Δ é dado por

$$\Delta = \{(a, b) \in X \times X \mid a = b\}.$$

Pela irreflexividade de \prec , tem-se que $\prec \cap \Delta = \emptyset$. Se supusermos que $a \preceq b$ e $b \preceq a$, devemos ter então que, ou $a \prec b$ e $b \prec a$, ou $a = b$. Pela propriedade de comparabilidade de \prec , se $a \neq b$, ou $a \prec b$, ou $b \prec a$, ambos não podem ser verdadeiros ao mesmo tempo. Portanto, devemos ter que $a = b$. Logo \preceq é anti-simétrica.

□

Esta proposição mostra que dada uma ordem linear em X , sempre podemos definir, a partir dela, uma ordem total em X . O contrário também é verdade, dada uma ordem total, podemos sempre obter uma ordem linear. Apresentamos esta afirmação sem prova, mas basta realizar o processo inverso ao executado na prova da proposição anterior. Isto é, basta remover o conjunto Δ da ordem total para obter a ordem simples. Por este motivo, também, não se costuma fazer distinção entre ordens lineares e ordens totais, sendo usual falar apenas de ordens totais.

Agora, dado um conjunto simplesmente ordenado, X , existe uma topologia canônica em X , definida usando a relação de ordem. Esta topologia é chamado **topologia da ordem**. Denotando por $<$ a relação de ordem simples em X , dados dois elementos a, b em X tais que $a < b$, podemos definir quatro subconjuntos de X , chamados **intervalos**,

determinados por a e b . São eles:

$$(a, b) = \{x \mid a < x < b\};$$

$$(a, b] = \{x \mid a < x \leq b\};$$

$$[a, b) = \{x \mid a \leq x < b\};$$

$$[a, b] = \{x \mid a \leq x \leq b\};$$

Onde \leq é a ordem total definida a partir de $<$ segundo a proposição A.1.

Se $X = \mathbb{R}$, a noção apresentada aqui fornece os intervalos usuais da reta real. Apresentada desta maneira, no entanto, os intervalos são definidos em um conjunto simplesmente ordenado arbitrário. Um conjunto do primeiro tipo é dito um **intervalo aberto** em X . Um conjunto do último tipo é dito um **intervalo fechado** em X . E os dois conjuntos intermediários são ditos **intervalos semi-abertos** ou **semi-fechados**. O termo “aberto” sugere uma conexão na qual intervalos abertos em X são conjuntos abertos de X . Como veremos, isto é verdade.

Definição A.7: Seja X um conjunto simplesmente ordenado e que possui mais de um elemento e seja \mathcal{B} a coleção de todos os conjuntos dos seguintes tipos:

- i Todos os intervalos abertos (a, b) em X .
- ii Todos os intervalos da forma $[a_0, b)$, onde a_0 é o menor elemento de X , se existir.
- iii Todos os intervalos da forma $(a, b_0]$, onde b_0 é o maior elemento de X , se existir.

A coleção \mathcal{B} forma uma base de uma topologia em X . Esta topologia é chamada **topologia da ordem** em X .

Se X não possui um menor elemento, então não existem conjuntos de tipo ii em \mathcal{B} , e se X não possui um maior elemento, não existem conjuntos de tipo iii. Para $X = \mathbb{R}$, por exemplo, existem apenas conjuntos do tipo i.

Verifiquemos se \mathcal{B} é, de fato, uma base. Primeiro, note que todo elemento $x \in X$ pertence a pelo menos um elemento de \mathcal{B} . Se x é o menor elemento de X então x pertence a todos os conjuntos de tipo ii e se x for o maior elemento de X então x pertence a todos os elementos de tipo iii. Se X não possuir nem maior nem menor elemento, então x pertence a algum conjunto de tipo i. Segundo, note que a interseção de quaisquer dois conjuntos de qualquer um dos tipos especificados é um conjunto de um dos três tipos. Segue imediatamente da definição que a interseção de um intervalo aberto com outro

intervalo aberto ou com um intervalo semi-aberto é um intervalo aberto. A interseção de conjuntos do tipo ii com outro do tipo ii é também do tipo ii e a interseção de dois conjuntos do tipo iii também é do tipo iii. Por fim, a interseção de um conjunto de tipo ii com um de tipo iii é um intervalo aberto, ou seja, um conjunto de tipo i. Deste modo, fica estabelecido que \mathcal{B} é de fato uma base. Vejamos alguns exemplos.

Exemplo A.7: A topologia canônica em \mathbb{R} é simplesmente a topologia da ordem de \mathbb{R} sob a relação de ordem usual $<$. Ou seja, a topologia canônica de \mathbb{R} é a coleção de todos os intervalos abertos (a, b) , $a, b \in \mathbb{R}$.

Exemplo A.8: Os inteiros positivos, \mathbb{Z}^+ formam um conjunto ordenado que possui um elemento mínimo. A topologia da ordem em \mathbb{Z}^+ é a topologia discreta em \mathbb{Z}^+ . Para todo conjunto de um único elemento, se $n > 1$, $\{n\} = (n - 1, n + 1)$ é um elemento da base. Se $n = 1$, o conjunto $\{1\} = [1, 2)$ é um elemento da base.

Exemplo A.9: Considere o conjunto $X = \{1, 2\} \times \mathbb{Z}^+$. Denotaremos os elementos deste conjunto por $a \times b$, onde $a \in \{1, 2\}$ e $b \in \mathbb{Z}^+$, por conveniência. Equipando X com a ordem lexicográfica, dada por

$$a \times b < c \times d \iff a < c \vee (a = c \wedge b < d)$$

Vemos que X possui um menor elemento, 1×1 . Diferente do exemplo anterior, no entanto, a topologia da ordem em X não é a topologia discreta pois existe um conjunto de um único elemento, $\{2 \times 1\}$ que não é aberto pois todo intervalo contendo 2×1 é da forma

$$(1 \times n, 2 \times m)$$

onde $m > 1$. Este intervalo sempre contém um elemento do tipo $1 \times n$ pois \mathbb{Z}^+ não possui um maior elemento.

Exemplo A.10: Considere o conjunto $\mathbb{R} \times \mathbb{R}$ com a ordem lexicográfica definida no exemplo anterior. Este conjunto não possui nem um menor elemento nem um maior elemento, de modo que todos os abertos são intervalos da forma $(a \times b, c \times d)$. As figuras abaixo mostram intervalos onde $a < c$ e onde $a = c$ e $b < d$. A subcoleção contendo apenas conjuntos do segundo tipo é, também, uma base para esta topologia.

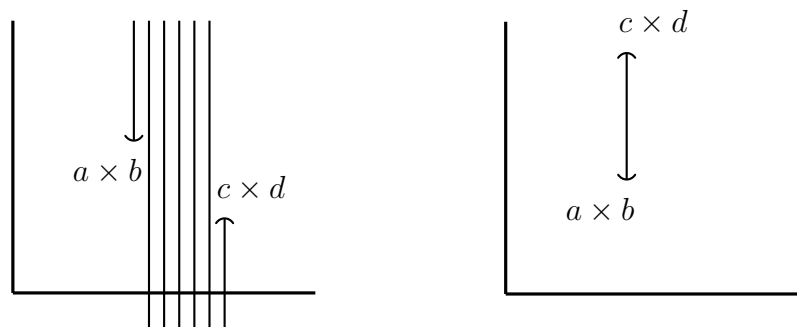


Figura A.5: Exemplos de conjuntos abertos em $\mathbb{R} \times \mathbb{R}$ na topologia da ordem lexicográfica.

Definição A.8: Se X é um conjunto simplesmente ordenado e a é um elemento de X , existem quatro subconjuntos de X , chamados **raios**, determinados por a . São eles:

$$(a, +\infty) = \{x \mid x > a\};$$

$$(-\infty, a) = \{x \mid x < a\};$$

$$[a, +\infty) = \{x \mid x \geq a\};$$

$$[-\infty, a) = \{x \mid x \leq a\};$$

Conjuntos dos dois primeiros tipos são chamados **raios abertos** e conjuntos dos dois últimos são chamados **raios fechados**. Novamente, o termo “aberto” sugere que raios abertos são conjuntos abertos de X . Isto é verdade. Considere por exemplo o raio $(a, +\infty)$. Se X possui um maior elemento, b_0 , então $(a, +\infty)$ coincide com o elemento base $(a, b_0]$. Se X não possui um maior elemento, $(a, +\infty)$ é igual à união de todos os elementos base da forma (a, x) , para $x > a$. Em ambos os casos, $(a, +\infty)$ é aberto. Um argumento análogo vale para $(-\infty, a)$.

Os raios abertos, inclusive, formam uma sub-base da topologia da ordem em X . Cada raio aberto é aberto na topologia da ordem, de modo que a topologia gerada por eles está contida na topologia da ordem. Por outro lado, cada elemento base da topologia da ordem é igual a uma interseção finita de raios abertos. O intervalo (a, b) é igual a interseção de $(-\infty, b)$ e $(a, +\infty)$, enquanto $[a_0, b)$ e $(a, b_0]$, se existirem, são raios abertos. Assim, a topologia gerada pelos raios abertos contém a da ordem.

A.3 - Topologia produto

Se X e Y são espaços topológicos, existe um modo canônico de definir uma topologia no produto cartesiano $X \times Y$.

Definição A.9: Sejam X e Y espaços topológicos. A **topologia produto** em $X \times Y$ é a topologia cuja base é a coleção \mathcal{B} de todos os conjuntos da forma $U \times V$ onde U é um subconjunto aberto de X , e V , de Y .

Verifiquemos se \mathcal{B} é uma base. A primeira condição é direta dado que todo o conjunto $X \times Y$ é um elemento base. A segunda condição também segue sem complicações. Dados dois elementos base $U_1 \times V_1$ e $U_2 \times V_2$, tem-se $(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2)$. Logo, a interseção de dois elementos da base é também um elemento da base, já que $U_1 \cap U_2$ e $V_1 \cap V_2$ são abertos, em X e Y respectivamente. Este argumento está ilustrado na figura a seguir.

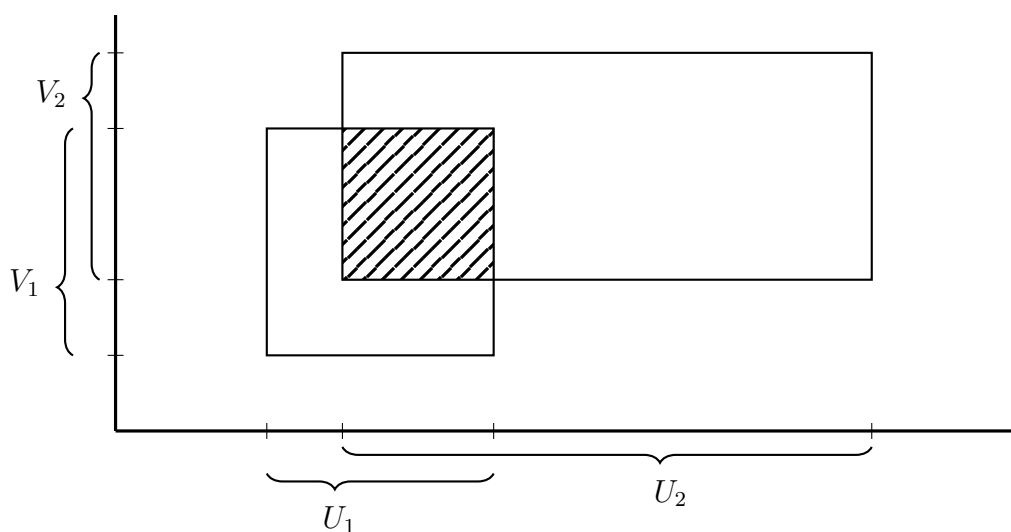


Figura A.6: Conjunto aberto na topologia produto.

Note que a coleção \mathcal{B} não é uma topologia em $X \times Y$. A união dos dois retângulos na figura acima, por exemplo, não é o produto de dois conjuntos, de modo que não pode pertencer a \mathcal{B} , no entanto, é aberto em $X \times Y$. Como usual ao introduzir um novo conceito, é útil tentar relacioná-lo com conceitos já conhecidos. Sendo assim, começamos por investigar o que pode-se dizer a respeito da topologia produto em $X \times Y$, se as topologias de X e Y são dadas em termos de bases.

Teorema A.1: Sejam \mathcal{B} uma base para a topologia de X e \mathcal{C} uma base para a topologia de Y . A coleção

$$\mathcal{D} = \{B \times C \mid B \in \mathcal{B} \wedge C \in \mathcal{C}\}$$

é uma base para a topologia de $X \times Y$.

Prova: Usamos o lema A.2. Dado um aberto $W \subset X \times Y$ e um ponto $x \times y \in W$, por definição da topologia produto, existe um elemento base $U \times V$ tal que $x \times y \in U \times V \subset W$. Como \mathcal{B} e \mathcal{C} são bases para X e Y , respectivamente, podemos escolher elementos $B \in \mathcal{B}$ e $C \in \mathcal{C}$ tais que $x \in B \subset U$ e $y \in C \subset V$. Deste modo, $x \times y \in B \times C \subset W$. Portanto, a coleção \mathcal{D} é, segundo o lema A.2, uma base para a topologia produto em $X \times Y$.

□

Exemplo A.11: A topologia canônica de \mathbb{R} é a topologia da ordem de \mathbb{R} . O produto desta topologia com si mesma, é a topologia canônica de $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. Ela tem como base a coleção de todos os produtos de conjuntos abertos de \mathbb{R} . No entanto, como o teorema anterior provou, a coleção muito menor de todos os produtos $(a, b) \times (c, d)$ de intervalos abertos em \mathbb{R} também é uma base para esta topologia. Cada conjunto destes pode ser visualizado como o interior de um retângulo em \mathbb{R}^2 . Assim, a topologia canônica de \mathbb{R}^2 é aquela considerada no exemplo A.6.

Por vezes é útil expressar a topologia produto por meio de uma sub-base. Para tal, definimos as seguintes funções:

Definição A.10: Seja $\pi_1 : X \times Y \rightarrow X$ definida por

$$\pi_1(x, y) = x$$

e seja $\pi_2 : X \times Y \rightarrow Y$ definida por

$$\pi_2(x, y) = y.$$

Os mapas π_1 e π_2 são chamados **projeções canônicas** ou simplesmente **projeções** de $X \times Y$ em seu primeiro e segundo fatores, respectivamente. Ambas as projeções são sobrejetivas.

Se U é um aberto de X , então o conjunto $\pi_1^{-1}(U)$ é precisamente o conjunto $U \times Y$ que é aberto em $X \times Y$. Similarmente, se V é um aberto de Y , o conjunto $\pi_2^{-1}(V)$ é o conjunto $X \times V$, que é aberto em $X \times Y$. A interseção destes dois conjuntos é o conjunto $U \times V$ como indicado na figura abaixo.

Formalmente, temos o seguinte resultado:

Teorema A.2: Sejam \mathcal{T}_X e \mathcal{T}_Y as topologias de X e Y , respectivamente. A coleção

$$\mathcal{S} = \{\pi_1^{-1}(U) \mid U \in \mathcal{T}_X\} \cup \{\pi_2^{-1}(V) \mid V \in \mathcal{T}_Y\}$$

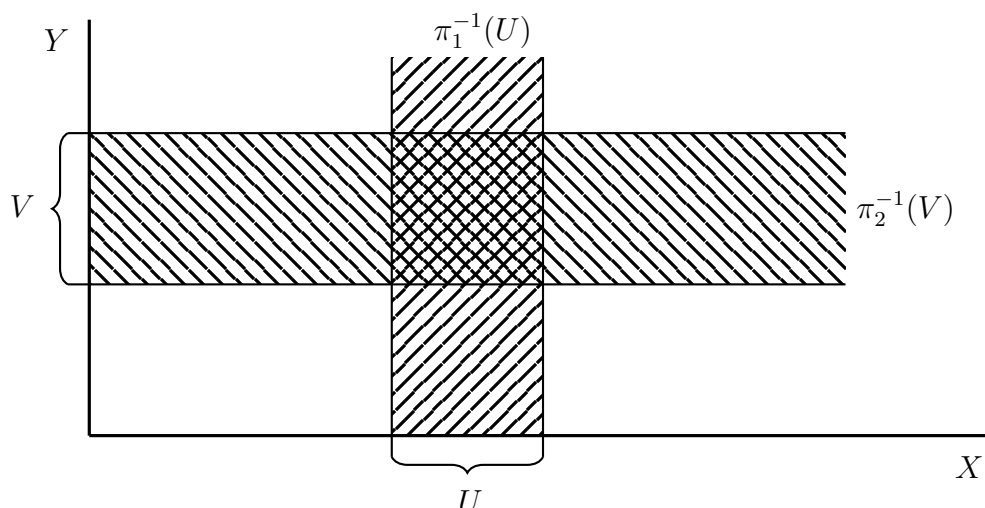


Figura A.7: Conjunto aberto na topologia produto determinado a partir das funções projeção.

é uma sub-base para a topologia produto em $X \times Y$.

Prova: Sejam \mathcal{T} a topologia produto em $X \times Y$ e \mathcal{T}' a topologia gerada por \mathcal{S} . Como cada elemento de \mathcal{S} pertence a \mathcal{T} , uniões arbitrárias e interseções finitas de elementos de \mathcal{S} também pertencem. Logo, $\mathcal{T}' \subset \mathcal{T}$. Por outro lado, cada elemento base, $U \times V$ da topologia \mathcal{T} é uma interseção finita de elementos de \mathcal{S} , pois:

$$U \times V = \pi_1^{-1}(U) \cap \pi_2^{-1}(V).$$

Portanto, $U \times V$ pertence a \mathcal{T}' e então $\mathcal{T} \subset \mathcal{T}'$.

□

A.4 - Topologia do subespaço

Definição A.11: Seja X um espaço topológico com topologia, \mathcal{T} . Se Y é um subconjunto de X , a coleção

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\}$$

é uma topologia em Y , chamada **topologia do subespaço**. Com esta topologia, Y é chamado de **subespaço** de X . Os abertos de Y são todas as interseções de abertos de X com Y .

A verificação de que \mathcal{T}_Y é uma topologia é simples. Tem-se que \emptyset e Y pertencem a \mathcal{T}_Y pois

$$\emptyset = Y \cap \emptyset \quad \text{e} \quad Y = Y \cap X$$

onde $\emptyset, Y \in \mathcal{T}$. O fato de ser fechado sob interseções finitas e uniões arbitrárias segue das seguintes propriedades das operações de união e interseção de conjuntos:

$$(U_1 \cap Y) \cap \cdots \cap (U_n \cap Y) = (U_1 \cap \cdots \cap U_n) \cap Y$$

e

$$\bigcup_{\alpha} (U_{\alpha} \cap Y) = \left(\bigcup_{\alpha} U_{\alpha} \right) \cap Y.$$

Lema A.3: Se \mathcal{B} é uma base para a topologia de X , a coleção

$$\mathcal{B}_Y = \{B \cap Y \mid B \in \mathcal{B}\}$$

é uma base para a topologia do subespaço em Y .

Prova: Dado um aberto U em X e dado $y \in U \cap Y$, podemos escolher um elemento base B tal que $y \in B \subset U$. Assim, $y \in B \cap Y \subset U \cap Y$. Segue, do lema A.2 que \mathcal{B}_Y é uma base para a topologia do subespaço em Y .

□

Ao lidar com um espaço, X e um subespaço, Y , é necessário ter cuidado quando se usa o termo “conjunto aberto”, pois ele pode significar tanto um elemento da topologia de X quanto um da topologia de Y . Convencionalmente, um conjunto U é **aberto em** Y se ele pertence à topologia de Y . Em particular, isto implica que U é um subconjunto de Y . Do mesmo modo, diremos que U é aberto em X se pertence à topologia de X .

Lema A.4: Seja Y um subespaço de X . Se U é aberto em Y e Y é aberto em X , então U é aberto em X .

Prova: Como U é aberto em Y , $U = Y \cap V$ para algum V aberto em X . Como Y e V são, ambos abertos em X , $Y \cap V$ também é aberto em X .

□

Agora, exploremos a relação entre a topologia de subespaço e as topologias da ordem e do produto.

Teorema A.3: Se A é um subespaço de X e B um subespaço de Y , então a topologia produto em $A \times B$ é a mesma topologia que $A \times B$ herda como um subespaço de $X \times Y$.

Prova: O conjunto $U \times V$ é um elemento base genérico de $X \times Y$, onde U é aberto em X e V é aberto em Y . Assim, $(U \times V) \cap (A \times B)$ é um elemento base genérico da topologia do subespaço $A \times B$. Note que

$$(U \times V) \cap (A \times B) = (U \cap A) \times (V \cap B).$$

Como $U \cap A$ e $V \cap B$ são abertos na topologia do subespaço em A e B , respectivamente, o conjunto $(U \cap A) \times (V \cap B)$ é um elemento base da topologia do produto em $A \times B$. Conclui-se, portanto, que as bases para a topologia do produto e topologia do subespaço são a mesma de modo que as duas topologias também são iguais. \square

Agora, considere um conjunto ordenado X com a topologia da ordem e seja Y um subconjunto de X . A relação de ordem em X , quando restrita a Y torna Y um conjunto ordenado. A topologia da ordem em Y , no entanto, não necessariamente é a mesma topologia que Y herda como subespaço de X . Vejamos alguns exemplos.

Exemplo A.12: Considere o subconjunto $Y = [0, 1]$ da linha real, \mathbb{R} , com a topologia do subespaço. A topologia do subespaço tem como base todos os conjuntos da forma $(a, b) \cap Y$, onde (a, b) são intervalos abertos de \mathbb{R} . Este conjunto é de um dos seguintes tipos:

$$(a, b) \cap Y = \begin{cases} (a, b) & \text{se } a, b \in Y; \\ [0, b) & \text{se apenas } b \in Y; \\ (a, 1] & \text{se apenas } a \in Y; \\ Y \text{ ou } \emptyset & \text{se } a, b \notin Y. \end{cases}$$

Por definição todos estes conjuntos são abertos em Y , no entanto, conjuntos do segundo e terceiro tipos não são abertos em \mathbb{R} . Agora, note que estes conjuntos, por definição, formam uma base para a topologia da ordem em Y . Assim, neste caso, a topologia do subespaço e a da ordem são iguais.

Exemplo A.13: Seja $Y \subset \mathbb{R}$ o conjunto dado por $Y = [0, 1) \cup \{2\}$. Na topologia do subespaço de Y , o conjunto $\{2\}$ é aberto, porque é a interseção do aberto $(3/2, 5/2)$ com Y . Mas, na topologia da ordem em Y , o conjunto $\{2\}$ não é aberto em Y . Qualquer elemento base da topologia da ordem de Y que contém 2 é da forma

$$\{x \mid x \in Y \wedge a < x \leq 2\}$$

para algum $a \in Y$. Tal conjunto necessariamente contém pontos de Y menores que 2.

Exemplo A.14: Considere $I = [0, 1]$. A ordem lexicográfica em $I \times I$ é simplesmente a restrição a $I \times I$ da ordem lexicográfica no plano $\mathbb{R} \times \mathbb{R}$. No entanto, a topologia da ordem em $I \times I$ não é igual à topologia do subespaço em $I \times I$ obtida a partir da topologia da ordem lexicográfica em $\mathbb{R} \times \mathbb{R}$. Por exemplo, o conjunto $\{1/2\} \times (1/2, 1]$ é aberto em $I \times I$ na topologia do subespaço, mas não o é na topologia da ordem. Em particular, não existe um elemento base na topologia da ordem de $I \times I$ que contenha o ponto $1/2$ e esteja inteiramente contido em $\{1/2\} \times (1/2, 1]$. O argumento está ilustrado na figura abaixo.

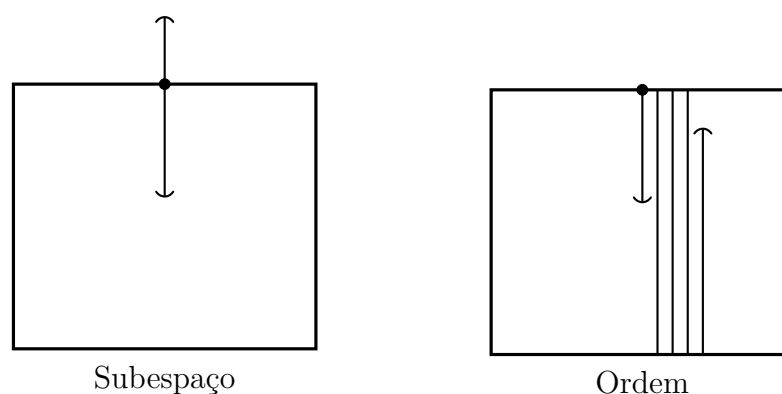


Figura A.8: Comparação das topologias do subespaço e da ordem em um subespaço.

O conjunto $I \times I$ com a topologia da ordem lexicográfica é chamado de **quadrado ordenado**, denotado por I_o^2 .

A anomalia que ocorre nos segundo e terceiro exemplos não ocorre para intervalos ou raios em um conjunto ordenado, X . Dado um conjunto ordenado X , dizemos que um subconjunto Y de X é **convexo** em X se para todo par de pontos $a < b$ em Y , todo o intervalo (a, b) de pontos de X está contido em Y . Note que intervalos e raios em X são convexos em X .

Teorema A.4: Sejam X um conjunto ordenado e com a topologia da ordem e Y um subconjunto de X convexo. Então, a topologia da ordem em Y é a mesma topologia que Y herda como subespaço de X .

Prova: Considere o raio $(a, +\infty)$ em X . Se $a \in Y$, tem-se que

$$(a, +\infty) \cap Y = \{x \mid x \in Y \wedge x > a\}.$$

Este conjunto é um raio aberto no conjunto ordenado Y . Se $a \notin Y$, então a é ou uma limite inferior em Y ou um limite superior em Y , já que Y é convexo. No primeiro

caso, o conjunto $(a, +\infty) \cap Y$ é igual a todo Y e no segundo caso, é o conjunto vazio. Um argumento análogo mostra que a interseção do raio $(-\infty, a)$ com Y é ou um raio aberto de Y ou o próprio Y ou vazio. Como os conjuntos $(a, +\infty) \cap Y$ e $(-\infty, a) \cap Y$ formam uma sub-base para a topologia do subespaço em Y e são abertos na topologia da ordem, a topologia da ordem contém a topologia do subespaço.

Para demonstrar a volta, note que qualquer raio aberto de Y é igual à interseção de um raio aberto de X com Y , de modo que é aberto na topologia do subespaço em Y . Como cada raio aberto de Y é uma sub-base da topologia da ordem de Y , esta topologia está contida na topologia do subespaço.

□

Usualmente, quando se considerar um conjunto ordenado X com a topologia da ordem e um subconjunto Y de X , assume-se que a Y é dada a topologia do subespaço a menos que especificado de outro modo. Se Y for convexo em X , então ambas as topologias são iguais. Agora, faremos um breve interlúdio para tratar de mais alguns conceitos básicos de espaços topológicos antes de tratarmos de um último método de construção de espaços topológicos.

A.5 - Conjuntos fechados e pontos aderentes

Além dos conjuntos abertos, existem outros conceitos básicos associados a espaços topológicos. Nesta seção falaremos brevemente sobre conjuntos fechados, o fecho de um conjunto e pontos aderentes. Todos estes conceitos são úteis para estudar aspectos de um espaço topológico. Inclusive, é comum que certas propriedades formuladas apenas em termos de abertos de um espaço possam ser reformuladas de modo equivalente usando um destes conceitos. Isto é útil pois fornece perspectivas distintas sobre uma mesma propriedade. Começamos com conjuntos fechados.

Definição A.12: Se X é um espaço topológico, um subconjunto A de X é dito **fechado** se $X - A$ é aberto.

Exemplo A.15: Os intervalos fechados em \mathbb{R} , $[a, b]$ são conjuntos fechados, pois:

$$\mathbb{R} - [a, b] = (-\infty, a) \cup (b, +\infty)$$

é aberto. Similarmente, $[a, +\infty)$ é fechado pois seu complemento, $(-\infty, a)$ é aberto. Estes resultados justificam a nomenclatura: intervalo fechado e raio fechado. De fato, a definição de conjunto fechado vem justamente de modo a capturar as propriedades destes subconjuntos de \mathbb{R} . Note que $[a, b) \subset \mathbb{R}$ não é nem aberto nem fechado.

Como o exemplo anterior ilustra, assim como portas podem estar abertas ou fechadas, subconjuntos de um espaço topológico, podem ser abertos ou fechados. Diferentemente de portas, por outro lado, eles também podem ser simultaneamente abertos e fechados ou nem abertos, nem fechados. A coleção de conjuntos fechados de um espaço possui propriedades muito similares à coleção de abertos, isto é, à topologia do espaço.

Teorema A.5: Se X é um espaço topológico, então as seguintes condições valem:

- i \emptyset e X são fechados.
- ii Interseções arbitrárias de subconjuntos fechados são fechadas.
- iii Uniões finitas de subconjuntos fechados são fechados.

Prova:

- i \emptyset e X são fechados pois seus complementos são respectivamente X e \emptyset , que são abertos por definição.
- ii Dada uma coleção de subconjuntos fechados, $\{A_\alpha\}$, tem-se que

$$X - \bigcap_{\alpha} A_{\alpha} = \bigcup_{\alpha} (X - A_{\alpha}).$$

Como os conjuntos $X - A_{\alpha}$ são abertos, por definição, e uniões arbitrárias de abertos são abertas, também por definição, então interseções arbitrárias de fechados são fechadas.

- iii Dada uma coleção finita de subconjuntos fechados $\{A_i\}$, tem-se que

$$X - \bigcup_i A_i = \bigcap_i (X - A_i).$$

Por definição $X - A_i$ é aberto e interseções finitas de aberto são abertas. Portanto, uniões finitas de subconjuntos fechados são fechadas.

□

Note que seria perfeitamente possível usar conjuntos fechados para definir a topologia de um conjunto. Isto é, a topologia seria uma coleção de subconjuntos, ditos fechados satisfazendo ao teorema acima. Os conjuntos abertos, então, seriam os complementos destes conjuntos fechados. Este procedimento não possui nenhuma vantagem particular sobre o procedimento usual de modo que convencionou-se definir topologias em termos de abertos. Naturalmente, subconjuntos fechados possuem outros usos no estudo de topo-

logia. Não os trataremos aqui. Mas, notamos que muitas definições de topologia podem ser reformuladas em termos de produtos fechados. Isto fornece um segundo modo de obter certos resultados. Para alguns espaços, pode ser mais interessante usar subconjuntos fechados.

Deve-se ter um cuidado especial ao falar de conjuntos fechados quando um subespaço está envolvido. Se Y é um subespaço de X , dizemos que um conjunto A é fechado em relação a Y se A é um subconjunto de Y e A é fechado na topologia do subespaço de Y , isto é, $Y - A$ é aberto em Y . Tem-se, então o seguinte resultado:

Teorema A.6: Se Y é um subespaço de X , então o conjunto A é fechado em Y se, e somente se é a interseção de um conjunto fechado X com Y .

Prova: Suponha que $A = C \cap Y$ onde C é fechado em X . Então $X - C$ é aberto em X . Deste modo, $(X - C) \cup Y$ é aberto em Y pela definição da topologia do subespaço. Mas, $(X - C) \cap Y = Y - A$, logo $Y - A$ é aberto em Y de modo que A é fechado em Y . Este argumento está ilustrado na figura à esquerda abaixo. A área hachurada indica que se trata de um conjunto aberto.

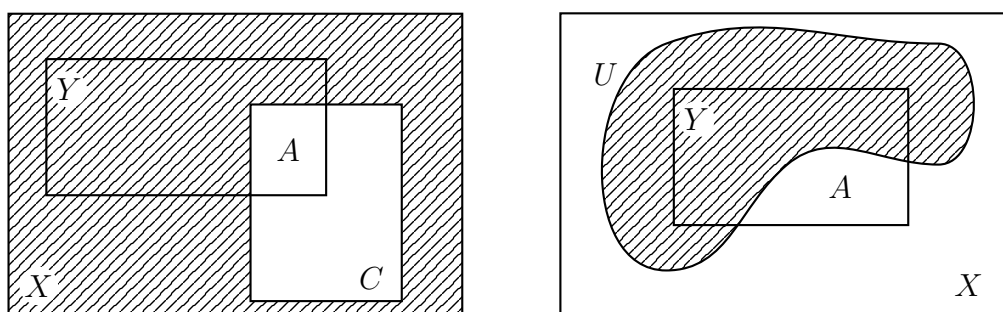


Figura A.9: Digramas ilustrando A e sua relação com X e Y .

Por outro lado, se A é fechado em Y isto significa que $Y - A$ é aberto em Y . Ou seja, existe um aberto U de X tal que $U \cap Y = Y - A$. O conjunto $X - U$ então é fechado em X . Mas, $Y \cap (X - U) = A$, logo A é a interseção de um fechado de X com Y . Este argumento está ilustrado na figura à direita acima onde a área não hachurada indica que o conjunto é fechado.

□

Este resultado será útil mais adiante. Passemos agora ao conceito de fecho de um conjunto.

Definição A.13: Dado um subconjunto A de um espaço topológico X , o **interior** de A , denotado $\text{Int } A$, é definido como a união de todos os abertos contidos em A . O **fecho** de A , denotado por \bar{A} é definido como a interseção de todos os fechados contendo A .

Vemos então que $\text{Int } A$ é o maior aberto contido em A e \bar{A} é o menor fechado que contém A , ou seja

$$\text{Int } A \subseteq A \subseteq \bar{A}.$$

Se A é aberto, então $A = \text{Int } A$ e se A é fechado, $A = \bar{A}$. O interior de um subconjunto não será muito usado, mas o fecho será. Note que ao tratar de um espaço X e um subespaço $Y \subset X$, é necessário cautela ao falar do fecho de conjuntos. Se $A \subseteq Y$, o fecho de A em Y e o fecho de A em X em geral serão diferentes. A notação \bar{A} geralmente fica reservada para o fecho de A em X . O fecho de A em Y pode ser expresso em termos de \bar{A} do seguinte modo:

Teorema A.7: Seja Y um subespaço de X e A um subconjunto de Y . O fecho de A em Y é dado por $\bar{A} \cap Y$.

Prova: Denotemos por B o fecho de A em Y . O conjunto \bar{A} é fechado em X de modo que $\bar{A} \cap Y$ é fechado em Y , segundo o teorema A.6. Como $\bar{A} \cap Y$ contém A e, por definição, B é a interseção de todos os fechados de Y contendo A , $B \subseteq (\bar{A} \cap Y)$. Por outro lado, como B é fechado, novamente pelo teorema A.6, $B = C \cap Y$ para algum C fechado em X . Deste modo, C é um fechado em Y contendo A , pois \bar{A} é a interseção de todos tais fechados de X . Ou seja, $\bar{A} \subseteq C$. Então $(\bar{A} \cap Y) \subseteq (C \cap Y) = B$.

□

A definição de fecho de um conjunto, no entanto, não fornece um método conveniente de determinar, de fato, o fecho, pois a coleção de todos os subconjuntos fechados de X é, em geral, muito grande para se trabalhar, assim como ocorre com a coleção de todos os abertos. Existe, no entanto, uma forma mais conveniente de se descrever o fecho de um conjunto envolvendo apenas a base da topologia de X . No que se segue, diremos que um conjunto A intersecciona outro conjunto B se $A \cap B \neq \emptyset$. Deste modo tem-se:

Teorema A.8: Se A é um subconjunto de X , então

- i $x \in \bar{A}$ se, e somente se, todo aberto U contendo x intersecciona A .
- ii supondo que a topologia de X seja dada por uma base, $x \in \bar{A}$ se, e somente se, todo elemento base B contendo x intersecciona A .

Prova:

- i Note que podemos escrever i como: $x \notin \bar{A}$ se, e somente se, existe um aberto U de X que não intersecciona A . Nesta forma, a prova é simples. Se $x \notin \bar{A}$, então, o conjunto $U = X - \bar{A}$ é um aberto contendo x , que não intersecciona A . Por outro lado, se existe um aberto U contendo x que não intersecciona

A , então $X - U$ é fechado e contém A . Por definição do fecho, \bar{A} , o conjunto $X - U$ deve conter \bar{A} e, portanto, x não pode pertencer a \bar{A} .

- ii Se todo aberto contendo x intersecciona A , também o fazem todos os elementos base B contendo x , pois os elementos base são, também, abertos. Por outro lado, se todo elemento base contendo x intersecciona A , também o fazem todos os abertos U contendo x , pois U contém pelo menos um elemento base contendo x . Mas, se todo aberto contendo x intersecciona A , por i, $x \in \bar{A}$.

□

Existe uma terminologia específica para dizer que U é um aberto contendo x . Diz-se que U é uma **vizinhança** de x . Alguns matemáticos definem “vizinhança” de modo diferente. Dizem que A é uma vizinhança de x se A contém algum aberto contendo x . É necessário atenção a qual definição está sendo usada, usaremos apenas a primeira. Por último introduzimos o conceito de **ponto aderente**.

Definição A.14: Se A é um subconjunto de um espaço X e x um ponto de X , dizemos que x é um **ponto aderente** de A se toda vizinhança de x intersecta A em algum ponto diferente de x .

Em outras palavras, x é um ponto aderente de A se pertence ao fecho de $A - \{x\}$. Note que não faz diferença se o ponto x pode pertencer ou não a A . Vejamos um exemplo simples.

Exemplo A.16: Seja $A = (0, 1] \subset \mathbb{R}$. Tem-se que 0 é um ponto aderente de A . Note que A não contém 0, de modo que $A - \{0\} = A$. Agora, o fecho de A é dado por $\bar{A} = [0, 1]$, pois toda vizinhança de 0, união de intervalos abertos, pelo menos um dos quais contém 0, intersecciona A . Apenas para números x fora do intervalo $[0, 1]$ existem abertos contendo x que não interseccionam A . Assim, como $0 \in \bar{A}$, 0 é um ponto aderente de A .

Como pode-se ver, existe uma relação entre o fecho de um conjunto e os pontos aderentes deste conjunto. Esta relação é explicitada no resultado a seguir.

Teorema A.9: Seja A é um subconjunto de um espaço X e A' o conjunto de todos os pontos aderentes de A . Então

$$\bar{A} = A \cup A'.$$

Prova: Se $x \in A'$, toda vizinhança de x intersecta A , em um ponto diferente de x . Pelo teorema A.8, então, $x \in \bar{A}$. Portanto, $A' \subset \bar{A}$. Como, por definição $A \subseteq \bar{A}$, tem-se que $A \cup A' \subset \bar{A}$. Por outro lado, considere $x \in \bar{A}$. Se $x \in A$ então $x \in A \cup A'$. Se $x \notin A$, como $x \in \bar{A}$, tem-se que toda vizinhança U de x intersecta A . Como $x \notin A$, U deve intersectar A em um ponto diferente de x . Então $x \in A'$ pela definição de ponto aderente. Logo $x \in A \cup A'$ ou seja, em ambos os casos tem-se que $\bar{A} \subset A \cup A'$. \square

Corolário A.9.1: Um subconjunto A de um espaço X é fechado se, e somente se, contém todos os seus pontos aderentes.

A.6 - Funções contínuas

O conceito de função contínua é fundamental para boa parte da matemática. Funções reais contínuas aparecem no início de qualquer livro de cálculo, seguidas de funções contínuas no plano e no espaço. Conforme se avança na matemática tipos cada vez mais gerais de funções contínuas aparecem. Nesta seção trataremos da definição e algumas propriedades de funções contínuas em espaços topológicos arbitrários.

Definição A.15: Sejam X e Y espaços topológicos. Uma função $f : X \rightarrow Y$ é **contínua** se para todo aberto $V \subseteq Y$, a imagem inversa de V sob f , isto é, o conjunto $f^{-1}(V)$ é aberto em X .

A continuidade de uma função depende, não somente de própria função, como também das topologias do domínio e contradomínio. Case deseje-se enfatizar este fato, diz-se que f é contínua relativa às topologias de X e Y .

Note que, se a topologia do contradomínio é especificada por uma base \mathcal{B} , então para provar a continuidade de f basta provar que a pré-imagem de qualquer elemento base é aberta. Isto é, se $B \in \mathcal{B}$ é um elemento base, devemos ter que $f^{-1}(B) \subseteq X$ é aberto em X . Esta condição é suficiente pois todo aberto $V \subseteq Y$ pode ser escrito como:

$$V = \bigcup_{\alpha} B_{\alpha}$$

de modo que

$$f^{-1}(V) = \bigcup_{\alpha} f^{-1}(B_{\alpha})$$

é aberto em X se cada conjunto $f^{-1}(B_{\alpha})$ o for.

Se a topologia de Y é especificada por uma sub-base, \mathcal{S} , para provar a continuidade de uma função f , bastará, então, provar que a pré-imagem de cada elemento da sub-base

é aberto. Como cada elemento base B de topologia de Y é escrito como uma interseção finita de elementos S da sub-base, então o fato da pré-imagem de cada elemento base ser aberta segue da seguinte propriedade da operação de interseção

$$f^{-1}(B) = \bigcap_{i=1}^n f^{-1}(S_i).$$

Usualmente a definição de continuidade é apresentada usando-se a ideia de “ $\epsilon - \delta$ ”. Vejamos como esta definição se relaciona com a definição de continuidade apresentada.

Exemplo A.17: Considere uma função real $f : \mathbb{R} \rightarrow \mathbb{R}$. A definição de continuidade usando ϵ 's e δ 's a partir da definição A.15 da seguinte forma. Suponha f contínua segundo a definição A.15. Dado $x_0 \in \mathbb{R}$ e $\epsilon > 0$, o intervalo

$$V = (f(x_0) - \epsilon, f(x_0) + \epsilon)$$

é aberto no contradomínio, \mathbb{R} , de modo que $f^{-1}(V)$ é aberto no domínio. Como $f^{-1}(V)$ contém o ponto x_0 , contém, também, um elemento base (a, b) tal que $x_0 \in (a, b)$. Escolhendo δ como sendo o menor entre os números $x_0 - a$ e $b - x_0$, então se $|x - x_0| < \delta$, o ponto x pertence a (a, b) de modo que $f(x) \in V$ e

$$|f(x) - f(x_0)| < \epsilon$$

como desejado. Por outro lado, se f é contínua segundo a definição $\epsilon - \delta$, então, dado $\epsilon > 0$, existe δ tal que

$$|f(x) - f(x_0)| < \epsilon \implies |x - x_0| < \delta.$$

Ou seja, qualquer aberto $V \subseteq Y$ que contém $f(x)$, pode ser escrito como $V = (f(x_0) - \epsilon, f(x_0) + \epsilon)$ tal que

$$f^{-1}(V) = (x_0 - \delta, x_0 + \delta).$$

Como $f^{-1}(V)$ é um intervalo aberto, $f^{-1}(V)$ é aberto no domínio e f é contínua segundo A.15.

O conceito de continuidade nos permite definir um outro conceito, central para a topologia:

Definição A.16: Sejam X e Y espaços topológicos e $f : X \rightarrow Y$ uma bijeção. Se tanto

f quanto sua inversa:

$$f^{-1} : Y \rightarrow X$$

forem contínuas, f é chamado de **homeomorfismo**.

A condição de que f^{-1} seja contínua significa que para cada aberto $U \subseteq X$, a pré-imagem de U sob f^{-1} é aberta em Y . No entanto, a pré-imagem de U sob f^{-1} é precisamente a imagem de U sob f . A imagem abaixo ilustra esta relação. Deste modo, pode-se definir um homeomorfismo como uma bijeção $f : X \rightarrow Y$ tal que $f(U)$ é aberto se, somente se U o é. Esta condição é bastante útil.

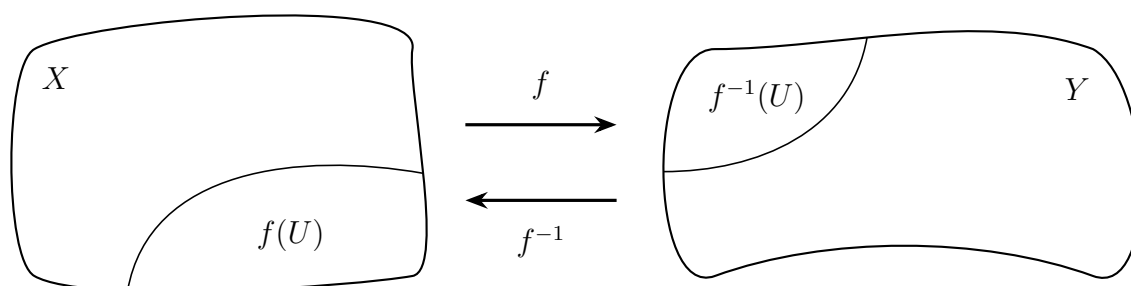


Figura A.10: Diagrama ilustrando a relação entre os abertos de X e Y sob o homeomorfismo f .

A utilidade de um homeomorfismo entre dois espaços X e Y é que ele nos dá uma correspondência unívoca, não só entre os elementos de X e Y como também entre as coleções de abertos de X e Y ou seja, entre as topologias destes dois espaços. Como consequência, propriedades que são expressas inteiramente em termos da topologia de X são mapeadas nas propriedades correspondentes de Y . Tais propriedades são ditas **propriedades topológicas** ou **invariantes topológicas** e são de grande utilidade para estudar e caracterizar espaços topológicos.

Exemplo A.18: Um exemplo simples de homeomorfismo é a função $f : (0, 1) \rightarrow (1, \infty)$ dada por

$$f(x) = \frac{1}{x}$$

A inversa de f é, simplesmente $f^{-1} : Y \rightarrow X$, tal que $f^{-1}(y) = 1/y$. De fato

$$f(f^{-1}(y)) = f\left(\frac{1}{y}\right) = y,$$

$$f^{-1}(f(x)) = f^{-1}\left(\frac{1}{x}\right) = x.$$

Se tratarmos ambos os intervalos $(0, 1)$ e $(1, \infty)$ na topologia do subespaço de \mathbb{R} ,

então uma base para a topologia de ambos são os intervalos abertos de cada espaço. Além disso, tem-se que:

$$f((a, b)) = \left(\frac{1}{b}, \frac{1}{a} \right)$$

Como f e f^{-1} mapeiam intervalos abertos em intervalos abertos, ambas são contínuas e portanto, são homeomorfismos. Isto nos diz, então, que o intervalo $(1, \infty)$ possui exatamente as mesmas propriedades topológicas que o intervalo $(0, 1)$. Isto é interessante quando consideramos o fato de que um deles é infinito e o outro não.

No cálculo existem alguns resultados úteis para construir funções contínuas, por exemplo, o produto de funções contínuas é uma função contínua, a soma de duas funções contínuas é contínua, a função constante é contínua e assim por diante. Vejamos alguns resultados similares na linguagem própria da topologia.

Teorema A.10: Sejam X, Y e Z espaços topológicos, tem-se:

- i **Função constante:** se $f : X \rightarrow Y$ mapeia todo X em um único ponto $y_0 \in Y$, então f é contínua.
- ii **Inclusão:** se A é um subespaço de X , a inclusão $j : A \rightarrow X$ é contínua.
- iii **Compostas:** Se $f : X \rightarrow Y$ e $g : Y \rightarrow Z$ são contínuas, então $g \circ f : X \rightarrow Z$ também o é.
- iv **Restrição:** Se $f : X \rightarrow Y$ é contínua e A é um subespaço de X , então a restrição $f|_A : A \rightarrow Y$ é contínua.
- v **Restrição ou expansão do contradomínio:** Seja $f : X \rightarrow Y$ uma função contínua. Se Z é um subespaço de Y contendo o conjunto imagem $f(X)$, então a função $g : X \rightarrow Z$ obtida ao restringir o contradomínio de f é contínua. Se Z é um espaço que possui Y como subespaço, então a função $h : X \rightarrow Z$ obtida ao expandir o contradomínio de f é contínua.
- vi **Continuidade local:** Uma função $f : X \rightarrow Y$ é contínua se X pode ser escrito como a união de abertos U_α tais que restrição de f a um aberto, $f|_{U_\alpha}$ é contínua para cada α .

Prova:

- i Seja $f(x) = y_0$ para todo $x \in X$. Seja V aberto em Y . O conjunto $f^{-1}(V)$ é igual a X ou a \emptyset , a depender se V contém ou não y_0 . Em qualquer caso, é aberto, logo f é contínua.

ii Se U é aberto em X , então $j^{-1}(U) = U \cap A$ que é aberto em A pela definição da topologia do subespaço.

iii Se U é aberto em Z , então $g^{-1}(U)$ é aberto em Y e $f^{-1}(g^{-1}(U))$ é aberto em X . Mas

$$f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U)$$

segundo a teoria básica de conjuntos, logo $g \circ f$ é contínua.

iv A restrição de f a A , $f|_A$ é igual à composição do mapa inclusão $j : A \rightarrow X$ e o mapa $f : X \rightarrow Y$, ambos contínuos, portanto, pelo item iii, $f|_A$ é contínua.

v Seja $f : X \rightarrow Y$ uma função contínua e seja $g : X \rightarrow Z$ a função obtida a partir de f restringindo seu contradomínio a Z com $f(X) \subseteq Z \subseteq Y$. Então se B é um aberto de Z , $B = Z \cup U$ para algum aberto $U \subseteq Y$. Como Z contém toda a imagem de f tem-se

$$f^{-1}(U) = g^{-1}(B).$$

Como $f^{-1}(U)$ é aberto, $g^{-1}(B)$ também o é. Por outro lado, se $h : X \rightarrow Z$ é a função obtida expandindo-se o contradomínio de f com $f(X) \subseteq Y \subseteq Z$, $h = f \circ j$, onde $j : Y \rightarrow Z$ é o mapa inclusão. Assim, h é contínua.

vi Por hipótese, podemos escrever X como a união de abertos U_α tais que $f|_{U_\alpha}$ é contínua para todo α . Seja V um aberto de Y . Então

$$f^{-1}(V) \cap U_\alpha = (f|_{U_\alpha})^{-1}(V)$$

pois ambas as expressões representam o conjunto de pontos $x \in U_\alpha$ tais que $f(x) \in V$. Como $f|_{U_\alpha}$ é contínua, este conjunto é aberto em U_α e, portanto, aberto em X . Mas

$$f^{-1}(V) = \bigcup_{\alpha} (f^{-1}(V) \cap U_\alpha)$$

de modo que $f^{-1}(V)$ também é aberta em X . Logo f é contínua.

□

Além destes resultados existe mais um, chamado de lema da colagem que é bastante útil, especialmente no estudo de homotopias no apêndice B. Antes de enunciá-lo, no entanto precisamos reformular o conceito de continuidade em termos de conjuntos fechados. A formulação é a seguinte:

Teorema A.11: Sejam X e Y espaços topológicos e $f : X \rightarrow Y$ um mapa entre estes espaços. Este mapa é contínuo se, e somente se, para todo subconjunto fechado $B \subseteq Y$, $f^{-1}(B)$ é fechado em X .

Prova: Primeiro note que decorre da teoria básica de conjuntos que

$$f^{-1}(Y - B) = X - f^{-1}(B)$$

Como B é fechado, $Y - B$ é aberto. Supondo f contínua, então, tem-se que $X - f^{-1}(B)$ é aberto de modo que $f^{-1}(B)$ é fechado. Por outro lado, como B é fechado, $f^{-1}(B)$ também o é, por hipótese. Como todo aberto de Y pode ser escrito como $Y - B$ para algum fechado $B \subseteq Y$, tem-se, pela identidade acima, que f mapeia abertos $Y - B$ em abertos $X - f^{-1}(B)$. Portanto, f é contínua. □

Com isto, podemos provar o seguinte resultado:

Teorema A.12 – Lema da colagem: Seja $X = A \cup B$, onde A e B são fechados em X e sejam $f : A \rightarrow Y$ e $g : B \rightarrow Y$ funções contínuas. Se $f(x) = g(x)$ para todo $x \in A \cap B$ então f e g se combinam para formar uma função contínua $h : X \rightarrow Y$ definida por

$$h(x) = \begin{cases} f(x) & \text{se } x \in A, \\ g(x) & \text{se } x \in B. \end{cases}$$

Prova: Seja C um subconjunto fechado de Y . Tem-se

$$h^{-1}(C) = f^{-1}(C) \cup g^{-1}(C).$$

Como f é contínua, $f^{-1}(C)$ é fechado em A e, conseqüentemente, fechado em X . Do mesmo modo $g^{-1}(C)$ é fechado em B e, conseqüentemente, fechado em X . A união de ambos, portanto, é fechado em X . □

Note que se A e B forem abertos, este resultado é apenas um caso particular da continuidade local com descrita em A.10, de modo que o lema da colagem em termos de fechados de X é um resultado mais geral. Por último, provamos o seguinte resultado sobre a continuidade de funções cujo contradomínio é o produto cartesiano de dois espaços.

Teorema A.13: Seja $f : A \rightarrow X \times Y$ dada pela equação

$$f(a) = f_1(a) \times f_2(a).$$

Esta função é contínua se, somente se as funções

$$f_1 : A \rightarrow X \quad \text{e} \quad f_2 : A \rightarrow Y$$

forem contínuas. Estes mapas são chamados **funções coordenadas** de f .

Prova: Sejam $\pi_1 : X \times Y \rightarrow X$ e $\pi_2 : X \times Y \rightarrow Y$ as projeções canônicas em X e Y respectivamente. Estes mapas são contínuos pois $\pi_1^{-1}(U) = U \times Y$ e $\pi_2^{-1}(V) = X \times V$ e estes conjuntos são abertos na topologia do produto se U e V são abertos em X e Y respectivamente. Agora, note que, para todo $a \in A$

$$f_1(a) = \pi_1(f(a)) \quad \text{e} \quad f_2(a) = \pi_2(f(a)).$$

Se f é contínua, então f_1 e f_2 o são pois são uma composição de funções contínuas. Por outro lado, suponhamos que f_1 e f_2 sejam contínuas. A base de da topologia de $X \times Y$ consistem em subconjuntos $U \times V$, onde U é um aberto de X e V , um de Y . Um ponto a pertence à pré-imagem $f^{-1}(U \times V)$ se, somente se $f(a) \in U \times V$, isto é, se, somente se $f_1(a) \in U$ e $f_2(a) \in V$. Portanto

$$f^{-1}(U \times V) = f_1^{-1}(U) \cap f_2^{-1}(V).$$

Como ambos os conjuntos $f_1^{-1}(U)$ e $f_2^{-1}(V)$ são abertos, sua interseção também o é. Logo f é contínua. □

Infelizmente não existem um resultado análogo para o caso onde o domínio é um produto de espaços. Pode-se conjecturar que que um mapa $f : A \times B \rightarrow X$ é contínuo se for contínuo em cada variável separadamente mas isto não é verdade. De todo modo, finalizamos aqui a discussão sobre continuidade e passamos a mais um modo de construir espaços topológicos a partir de outros e em seguida trataremos do conceito de conexidade de um espaço.

A.7 – Topologia quocientes

Dado um espaço topológico, (X, \mathcal{T}) , existem várias maneiras de gerar outros espaços topológicos a partir dele. Pode-se formar um espaço produto de X com outro espaço

topológico Y , como é o caso de \mathbb{R}^n , por exemplo. Pode-se considerar subespaços de interesse, pode-se gerar uma topologia a partir de uma métrica em um conjunto Z . Uma forma particularmente interessante de formar novos espaços topológicos é formando o espaço quociente. Uma das principais motivações para o conceito de espaços quocientes vem da geometria. Frequentemente geômetras desejam formar novos objetos a partir de operações de cortar e colar. Exemplos incluem a construção de um toro, que pode ser obtido colando-se os lados de um retângulo, como indicado na figura a seguir:

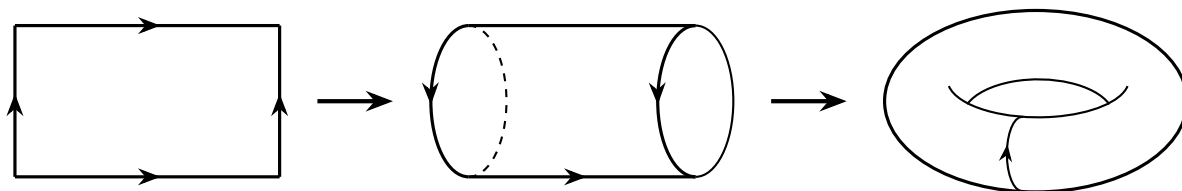


Figura A.11: Processo de construção de um toro a partir de um retângulo.

E o de uma casca esférica, obtida colapsando-se a fronteira de um disco em um único ponto:

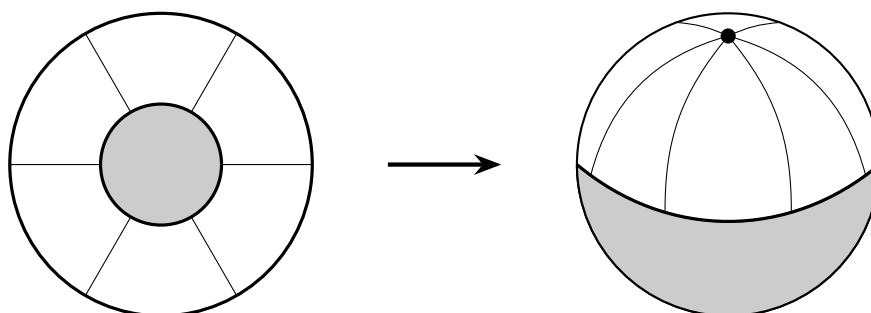


Figura A.12: Processo de formação de uma casca esférica a partir de um disco.

A formalização destas construções requer o conceito de **topologia quociente**. Definimos:

Definição A.17 – Mapa quociente: Sejam X e Y espaços topológicos e $p : X \rightarrow Y$ um mapa sobrejetivo. Ele é dito um **mapa quociente** se um subconjunto $U \subseteq Y$ for aberto em Y se, e somente se, $f^{-1}(U)$ for aberto em X .

Esta condição é mais forte que simplesmente exigir que p seja contínua e por vezes é chamada de **continuidade forte**. Equivalentemente, esta condição pode ser expressa em termos de conjuntos fechados como $U \subseteq Y$ é fechado em Y se, e somente se, $p^{-1}(U)$ for fechado em X . A equivalência das duas definições é garantida pela identidade:

$$f^{-1}(Y - B) = X - f^{-1}(B)$$

onde $f : X \rightarrow Y$ é um mapa sobrejetivo e $B \subseteq Y$.

Um exemplo de mapa quociente é a projeção $\pi_1 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ que projeta sobre a primeira coordenada, isto é $\pi_1(x, y) = x$. π_1 contínuo e sobrejetivo e, além disso, é um mapa aberto. Isto é para cada aberto $W \subseteq \mathbb{R} \times \mathbb{R}$, $\pi_1(W)$ é aberto. Note que aqui subentende-se que estamos usando as topologias canônicas de $\mathbb{R} \times \mathbb{R}$ e \mathbb{R} . Neste caso, tem-se que $W = U \times V$ onde U e V são abertos de \mathbb{R} de modo que $\pi_1(U \times V) = U$. Isto mostra que π_1 é um mapa quociente pois a imagem inversa de um subconjunto $U \subseteq \mathbb{R}$ é aberta em $\mathbb{R} \times \mathbb{R}$ se, e somente se, U é aberto em \mathbb{R} .

Um mapa quociente pode ser usado para construir uma topologia em um conjunto.

Definição A.18 – Topologia quociente: Se X é um espaço topológico e A um conjunto e se $p : X \rightarrow A$ é um mapa sobrejetivo existe exatamente uma topologia \mathcal{T} em A segundo a qual p é um mapa quociente. Esta topologia é chamada **topologia quociente** induzida por p .

A topologia quociente \mathcal{T} é definida de modo a consistir exatamente dos subconjuntos $U \subseteq A$ tais que $p^{-1}(U)$ é aberto em X . A verificação é direta. Tem-se $\emptyset \in \mathcal{T}$ e $A \in \mathcal{T}$ pois: $p^{-1}(\emptyset) = \emptyset$ e $p^{-1}(A) = X$, ambos abertos em X por definição. As outras duas condições necessárias são satisfeitas devido às igualdades:

$$p^{-1}\left(\bigcup_{\alpha} U_{\alpha}\right) = \bigcup_{\alpha} p^{-1}(U_{\alpha}) \quad \text{e} \quad p^{-1}\left(\bigcap_{i=1}^n U_i\right) = \bigcap_{i=1}^n p^{-1}(U_i).$$

Um exemplo de topologia quociente é a seguinte. Considere a reta real e o conjunto de três elementos $A = \{a, b, c\}$ e defina o mapa p por

$$p(x) = \begin{cases} a & \text{se } x > 0; \\ b & \text{se } x < 0; \\ c & \text{se } x = 0. \end{cases}$$

Tem-se que $p^{-1}(\{a\}) = (0, +\infty)$ e $p^{-1}(\{b\}) = (-\infty, 0)$, ambos são abertos em \mathbb{R} . Por outro lado, $p^{-1}(\{c\}) = \{0\}$ que não é aberto em \mathbb{R} . A topologia quociente de A está mostrada na figura abaixo:

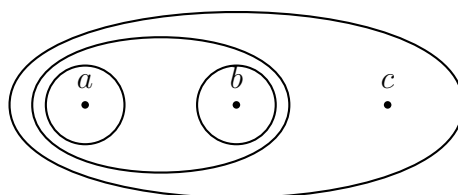


Figura A.13: Representação da topologia quociente de A .

Em essência o que o mapa p faz é colapsar todos os reais positivos em um ponto, a , os

negativos em b e mapear 0 em c . Poderíamos, por outro lado, “dobrar” e colapsar a linha real em torno do ponto zero se mapeássemos tanto os reais positivos quando os negativos em um mesmo elemento, a .

Outra situação onde a topologia produto ocorre com bastante frequência e de especial interesse para este trabalho é o caso de **espaços quocientes** de um espaço topológico X .

Definição A.19 – Espaço quociente: Seja X um espaço topológico e X^* uma partição de X em subconjuntos disjuntos. Seja $p : X \rightarrow X^*$ o mapa que leva cada elemento $x \in X$ à elemento $x^* \in X^*$ que o contem. Na topologia quociente induzida por p , o espaço X^* é chamado de **espaço quociente** de X .

Dado X^* existe uma relação de equivalência de modo que cada elemento de X^* corresponde a uma classe de equivalência. Pode-se pensar em X^* como obtido a partir da identificação de pares de pontos equivalentes. Por este motivo, o espaço quociente X^* também é chamado **espaço de identificação** ou **espaço de decomposição** de X .

Naturalmente, dado uma relação de equivalência o inverso também vale. Se \sim é uma relação de equivalência em X , o espaço X^* é simplesmente o conjunto de todas as classes de equivalência, usualmente denotado por X/\sim e a topologia quociente é aquela induzida pela sobrejeção canônica, $p : X \rightarrow X/\sim$ que leva cada $x \in X$ à sua classe de equivalência, usualmente denotada por $[x]$.

A.8 – Espaços conexos

Existem várias propriedades que podem ser usadas para caracterizar espaços topológicos. Estas propriedades são ditas topológicas e, como discutido, espaços homeomorfos, isto é, espaços para os quais existe um homeomorfismo entre eles, possuem propriedades topológicas iguais. Propriedades comuns incluem compacticidade, que trata do conceito de um espaço ser “limitado” no sentido de todos os seus pontos estarem a uma certa uns dos outros. O intervalo fechado $[a, b] \subset \mathbb{R}$ é um exemplo de espaço compacto. Metrizabilidade, um espaço (X, \mathcal{T}) é metrizável se existe uma métrica em X que induz a topologia \mathcal{T} . Um espaço é dito Hausdorff se para quaisquer dois pontos, $x_1, x_2 \in X$ existem abertos U_1 e U_2 contendo x_1, x_2 respectivamente tais que $U_1 \cap U_2 = \emptyset$. A lista de propriedades continua. Nesta seção final do apêndice, trataremos da propriedade de conexidade e conexidade por caminhos, pois são extremamente importantes para tratar de homotopia, assunto do apêndice B e essencial para este trabalho.

A definição de conexidade de um espaço topológico é bastante intuitiva. Dizemos que um espaço pode ser “separado” se pode ser dividido em pedaços, isto é, subconjuntos abertos disjuntos. Se isto não for possível, então o espaço é conexo. Formalmente:

Definição A.20 – Conexidade: Seja X um espaço topológico. Uma **separação** de X é um par U, V de subconjuntos abertos, disjuntos e não vazios de X cuja união é X . O espaço X é dito **conexo** se não existe separação de X .

Por ser definida em termos apenas da coleção de abertos de X , isto é, da topologia de X , conexidade é uma propriedade topológica. Isto implica, em particular, que se um espaço X é conexo, qualquer espaço Y homeomorfo a X também o é. Podemos, ainda, reformular a definição como segue:

Teorema A.14: Um espaço X é conexo se, e somente, se os únicos subconjuntos de X que são simultaneamente abertos e fechados em X são o conjunto vazio, \emptyset e X .

Prova: A demonstração é simples. Se A é um subconjunto próprio de X que é não vazio e simultaneamente aberto e fechado, então, os conjuntos $U = A$ e $V = X - A$ constituem uma separação de X , pois ambos são abertos, disjuntos, não vazios e $U \cup V = X$. Alternativamente, se U e V são uma separação de X , então U é não vazio, diferente de X e é simultaneamente aberto e fechado pois $X - U = V$ é aberto. \square

Dado um subespaço Y de X , existe outra formulação útil de conexidade:

Lema A.5: Se Y é um subespaço de X , a separação de Y é um par de conjuntos A, B , não vazios, disjuntos, cuja união é Y e tal que nenhum deles contém pontos aderentes do outro. O espaço Y é conexo se não possui separação.

Prova: Suponha primeiro que A e B são uma separação de Y . Então A é tanto aberto quanto fechado em Y . O fecho de A em Y é o conjunto $\bar{A} \cap Y$, onde \bar{A} é o fecho de A em X . Como A é fechado em Y , $A = \bar{A} \cap Y$, isto é, A é igual ao seu fecho. Alternativamente, $\bar{A} \cap B = \emptyset$. Como \bar{A} é a união de A com os seus pontos aderentes, B não contém nenhum ponto aderente de A . Um argumento análogo mostra que A não contém nenhum ponto aderente de B . Por outro lado, supondo que A e B são conjunto não vazio, disjuntos cuja união é Y e que não contém nenhum ponto aderente um do outro, então $\bar{A} \cap B = \emptyset$ e $A \cap \bar{B} = \emptyset$, de modo que $\bar{A} \cap Y = A$ e $\bar{B} \cap Y = B$. Assim, ambos, A e B , são fechados em Y e como $A = Y - B$ e $B = Y - A$, são, também, abertos. \square

Vejamos alguns exemplos simples

Exemplo A.19: Seja $Y = [-1, 0) \cup (0, 1] \subset \mathbb{R}$. Os conjuntos $[-1, 0)$ e $(0, 1]$ são abertos em Y . Note que eles não são abertos em \mathbb{R} no entanto. Evidentemente são, também, não vazios e sua união é Y , por definição. Deste modo, formam uma separação de Y . Alternativamente note que nenhum deles contém um ponto aderente do outro, apesar de possuir 0 como um ponto aderente em comum.

Exemplo A.20: Seja $X = [-1, 1] \subset \mathbb{R}$. Os conjuntos $[-1, 0]$ e $(0, 1]$ são disjuntos e não vazios, mas não formam uma separação de X pois o primeiro não é aberto em X . Alternativamente, note que o primeiro contém o ponto aderente, 0 , que é um ponto aderente do segundo. Provaremos mais adiante o resultado extremamente útil de que não existe separação de $[-1, 1]$.

Exemplo A.21: O conjunto dos racionais, \mathbb{Q} não é conexo. De fato, os únicos subespaços conexos de \mathbb{Q} são os conjuntos de um único ponto. Se Y é um subespaço de \mathbb{Q} contendo dois pontos, p e q , pode-se escolher um irracional a entre p e q e escrever Y como a união dos abertos

$$Y \cap (-\infty, a) \quad \text{e} \quad Y \cap (a, +\infty).$$

Vejam agora como pode-se obter novos espaços conexos a partir de outros espaço, também conexos. Estes resultados sobre a construção de espaços conexos podem ser usados para demonstrar a conexidade de alguns espaços. Em particular, mostraremos que a reta real, \mathbb{R} , é conexa.

Lema A.6: Se C e D são uma separação de X e Y é um subespaço conexo de X , então Y está contido inteiramente ou em C ou em D .

Prova: Como C e D são ambos abertos em X , os conjuntos $C \cap Y$ e $D \cap Y$ são abertos em Y . Estes dois conjuntos são disjuntos e sua união é Y . Se ambos forem não vazios, constituiriam uma separação de Y , porém, como Y é conexo, tal separação não pode existir e devemos ter ou $C \cap Y = \emptyset$ ou $D \cap Y = \emptyset$.

□

Teorema A.15: A união de uma coleção de subespaços conexos de X que possuem um ponto em comum é conexa.

Prova: Seja $\{A_\alpha\}$ uma coleção de subespaços conexos de X , p um ponto de $\bigcap A_\alpha$ e $Y = \bigcup A_\alpha$. Suponha que $Y = C \cup D$ é uma separação de Y . O ponto p , então, pertence a um dos conjuntos C ou D . Suponha que $p \in C$. Como A_α é conexo, deve estar contido inteiramente em C ou em D pelo lema A.6. Mas, não pode estar contido em

D pois $p \in A_\alpha$ e $p \in C$. Assim, $A_\alpha \subset C$ para todo α de modo que $\bigcup A_\alpha = Y \subset C$, contradizendo a hipótese de que D é não vazio. Portanto, não pode haver separação de Y , ou seja, Y é conexo.

□

Teorema A.16: A imagem de um espaço conexo sob um mapa contínuo é conexa.

Prova: Seja $f : X \rightarrow Y$ um mapa contínuo e X um espaço conexo. Pelo teorema A.10, o mapa obtido pela restrição do contradomínio de f a $Z = f(X)$ é contínuo. Assim, consideramos, um mapa sobrejetivo e contínuo

$$g : X \rightarrow Z.$$

Suponha que $Z = A \cup B$ é uma separação de Z . Então, $g^{-1}(A)$ e $g^{-1}(B)$ são conjuntos disjuntos cuja união é X . Além disso, pela continuidade de g , ambos são abertos e pela sobrejetividade são não vazios. Assim, formam uma separação de X , contradizendo a hipótese de que X é conexo.

□

Existem mais resultados acerca da obtenção de novos espaços conexos a partir de outros. No entanto estes dois resultados são suficientes para a discussão deste trabalho. Um resultado digno de nota mas que não provaremos aqui, no entanto é que o produto cartesiano finito de espaços conexos também é conexo. Agora, usaremos estes resultados para mostrar que a reta real é conexa. A conexidade da reta real e de seus intervalos e raios. Este resultado é importante pois caminhos em um espaço topológico X são definidos como mapas contínuos de um intervalo $[a, b] \subset \mathbb{R}$, geralmente $[0, 1]$ em X . Em vista do teorema A.16, isto indica que podemos usar caminhos para definir um critério de conexidade. De fato, tal critério é chamado de conexidade por caminhos e é central para o estudo de homotopia que, por sua vez, é central para este trabalho. Provemos então que \mathbb{R} e seus intervalos são conexos.

Definição A.21: Um conjunto simplesmente ordenado L com mais de um elemento é dito um **contínuo linear** se

- i todo subconjunto L_0 de L que é limitado superiormente possui um supremo.
- ii Se $x < y$, existe z tal que $x < z < y$.

Teorema A.17: Se L é um contínuo linear com a topologia da ordem, então L é conexo e também o são seus intervalos e raios.

Prova: Lembre-se que um subespaço $Y \subset L$ é dito convexo se para qualquer par de pontos $a, b \in Y$, com $a < b$, todo o intervalo $[a, b]$ de pontos de L está contido em Y . Suponha então que Y é um subespaço convexo de L .

Supomos, então que Y é a união de conjuntos não vazios e disjuntos, A e B , abertos em Y e escolhamos dois pontos, $a \in A$ e $b \in B$. Por conveniência, supomos que $a < b$. O intervalo $[a, b]$ de pontos de L está contido em Y , já que Y , por hipótese, é convexo. Assim, $[a, b]$ é a união de conjuntos disjuntos

$$A_0 = A \cap [a, b] \quad \text{e} \quad B_0 = B \cap [a, b]$$

cada um dos quais é aberto em $[a, b]$ na topologia do subespaço, que coincide com a da ordem pois $[a, b]$ é, também, convexo. Os conjuntos A_0 e B_0 são não vazios pois $a \in A_0$ e $b \in B_0$. Assim, A_0 e B_0 constituem uma separação de $[a, b]$. Seja $c = \sup A_0$, o supremo de A_0 . Mostraremos que c não pertence nem a A_0 nem a B_0 , contradizendo a hipótese de que $[a, b]$ é a união de A_0 , e B_0 . Note, visto como subconjuntos de $[a, b]$ que herda a propriedade do supremo de L , o supremo de qualquer subconjunto de $[a, b]$ que é limitado superiormente deve pertencer a $[a, b]$.

- Caso 1: Suponha que $c \in B_0$, então $c \neq a$ de modo que ou $c = b$ ou $a < c < b$. Em qualquer caso, segue do fato que B_0 é aberto em $[a, b]$ que existe algum intervalo da forma $(d, c]$ contido em B_0 . Se $c = b$ tem-se uma contradição pois d é um limite superior de A_0 menor que c . Se $c < b$, notamos que $(c, b]$ não intersecta A_0 , já que c é um limite superior de A_0 . Então

$$(d, b] = (d, c] \cup (c, b]$$

não intersecta A_0 . Novamente, d é um limite superior de A_0 menor que c , contrariando o fato de c ser o supremo de A_0 . Este argumento está ilustrado no lado esquerdo da figura abaixo.

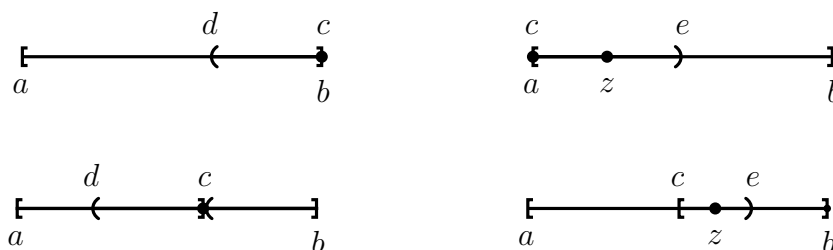


Figura A.14: Ilustração dos conjuntos que compõe $[a, b]$.

- Caso 2: Suponha que $c \in A_0$, então $c \neq b$ e ou $c = a$ ou $a < c < b$. Como A_0 é aberto em $[a, b]$, existe intervalo da forma $[c, e)$ contido em A_0 , como

ilustrado no lado direito da figura acima. Pela propriedade ii de uma ordem linear, podemos escolher um ponto $z \in L$ tal que $c < z < e$. Então, $z \in A_0$, contrariando o fato de c ser o supremo de A_0 .

Assim, A_0 e B_0 não podem ser uma separação de $[a, b]$, ou seja, não é possível existir separação de $[a, b]$ de modo que $[a, b]$ é conexo. Por sua vez, como $a \in A$ e $b \in B$, o intervalo $[a, b]$ intersecta A e B . No entanto, isto contradiz o fato de que A e B são uma separação de Y , pois, segundo o lema A.6, se A e B são uma separação, $[a, b]$ deveria estar contido inteiramente ou em A ou em B . Deste modo, não pode haver separação de Y , Y ser conexo implica que Y é convexo. Como intervalos e raios de L , são todos convexos, eles são convexos. E, como os intervalos formam uma base para a topologia da ordem, e por definição, a união de todos os abertos é L , tem-se que L é a união de conjuntos conexos e portanto é conexo pelo teorema A.15. \square

Corolário A.17.1: A reta real, \mathbb{R} e todos os seus intervalos e raios são conexos.

Por fim, podemos estabelecer um critério bastante útil para determinar se um espaço X é conexo.

Definição A.22: Dado pontos $x, y \in X$, um **caminho** em X de x a y é um mapa contínuo $f : [a, b] \rightarrow X$ onde $[a, b] \subset \mathbb{R}$, tal que $f(a) = x$ e $f(b) = y$. O espaço X é dito **conexo por caminhos** se todo par de pontos de X pode ser ligado por um caminho em X .

Para ver que conexidade por caminhos implica conexidade, suponha que $X = A \cup B$ seja uma separação de um espaço X conexo por caminhos e seja $f : [a, b] \rightarrow X$ um caminho em X . O conjunto $f([a, b])$ é conexo pois é a imagem, sob um mapa contínuo, de um conjunto conexo. Deste modo, $f([a, b])$ está inteiramente contido ou em A ou em B . De modo que não existe caminho em X ligando pontos de A a pontos de B , o que contradiz a hipótese de que X é conexo por caminhos. O inverso não necessariamente vale, no entanto. Isto é, um caminho que é conexo não necessariamente é conexo por caminhos. Vejamos alguns exemplos de espaços conexos por caminhos.

Exemplo A.22: Defina a **bola unitária**, B^n em \mathbb{R}^n por

$$B^n = \{ x \mid \|x\| \leq 1 \}$$

onde

$$\|x\| = \|(x_1, \dots, x_n)\| = (x_1^2 + \dots + x_n^2)^{1/2}.$$

A bola unitária é conexa por caminhos, dado dois pontos $x, y \in B^n$, o caminho reto, $f : [0, 1] \rightarrow \mathbb{R}^n$ definido por

$$f(t) = (1 - t)x + ty$$

pertence a B^n , pois, se $x, y \in B^n$ e $t \in [0, 1]$

$$\|f(t)\| \leq (1 - t)\|x\| + t\|y\| \leq (1 - t) + t \leq 1.$$

Exemplo A.23: Definimos o **espaço euclidiano perfurado** como o espaço $\mathbb{R}^n - \{0\}$, onde $\{0\}$ é a origem de \mathbb{R}^n . Se $n > 1$, este espaço é conexo por caminhos. Dado x e y diferentes de 0, pode-se juntá-los por um caminho reto entre eles se este caminho não passar pela origem. Caso contrário, pode-se escolher um ponto z que não pertence à reta ligando x e y e tomar a poligonal, formada pelas retas de x a z e de z a y .

Apêndice B

Sobre homotopia

Este apêndice trata do conceito de grupos de homotopia, em particular o primeiro grupo de homotopia, também conhecido como grupo fundamente e requer muitos dos conceitos estabelecidos no apêndice A. Caso não esteja familiarizado com conceitos básicos de topologia é recomendado que leia-o. Estes grupos surgem como um modo de provar que dois espaços não são homeomorfos. Para mostrar que dois espaços X e Y não são homeomorfos é necessário mostrar que pelo menos uma propriedade topológica — conexidade, compacticidade, entre outras — difere entre os dois espaços. Isto é suficiente para mostrar que o intervalo fechado $[0, 1]$ não é homeomorfo ao intervalo aberto $(0, 1)$ pois o primeiro é compacto enquanto o segundo não. Similarmente, a reta real \mathbb{R} não pode ser homeomorfa ao plano real \mathbb{R}^2 pois a deleção de um ponto de \mathbb{R}^2 resulta em um espaço conectado enquanto a deleção de um ponto de \mathbb{R} , não. Existem vários outros exemplos deste tipo.

No entanto, as propriedades básicas de um espaço topológico não são suficientes para mostrar se certos espaços são não homeomorfos. Por exemplo, propriedades como compacticidade, conexidade, conexidade local, metrizabilidade não são suficientes para mostrar se \mathbb{R}^3 e \mathbb{R}^2 são não homeomorfos. Do mesmo modo, não são suficientes para determinar se a superfície de uma 2-esfera, S_2 , a de um toro, T e a de um toro duplo $T\#T$ são não homeomorfos. Uma propriedade que pode ser usada para determinar se estas superfícies não são homeomorfas é a de **conexidade simples**. Basicamente um espaço é simplesmente conexo se todo caminho fechado pode ser contraído a um ponto. Por exemplo, remover um ponto de \mathbb{R}^2 resulta em um espaço que não é simplesmente conexo, diferente do que ocorre para \mathbb{R}^3 que permanece simplesmente conexo. Do mesmo modo, S_2 é simplesmente conexa enquanto T e $T\#T$ não o são. Esta propriedade não é capaz, no entanto, de diferenciar T de $T\#T$.

Como veremos, a conexidade simples é um caso particular de um outro conceito, o de grupo fundamental. Um espaço é simplesmente conexo quando seu grupo de homotopia é trivial. Além disso, dois espaços homeomorfos tem grupos fundamentais isomorfos. O grupo fundamental é capaz de distinguir T e $T\#T$, por exemplo. O grupo fundamental de T é abeliano enquanto o de $T\#T$ não o é. Deste modo não pode haver isomorfismo entre os dois. Como mencionamos, o grupo fundamental é o primeiro grupo de homotopia

de um espaço. Assim, é um caso particular de outro conceito, o de grupos de homotopia. Não trataremos dos grupos de homotopia superiores. Apesar disso, notamos que eles mantêm a função de servirem como ferramentas para mostra que duas superfícies não são homeomorfas. Iniciamos com o conceito de **homotopia**.

B.1 – Homotopia de caminhos

A homotopia de caminhos é uma relação de equivalência entre caminhos em um espaço topológico X . Algumas operações interessantes podem ser definidas no conjunto das classes de equivalência desta relação e estas operações fornecem a este conjunto a estrutura algébrica de um grupoide.

Definição B.1 – Homotopia: Se f e g são mapas contínuos de um espaço topológico X em outro, Y , diz-se que f é **homotópica** a g se existe um mapa contínuo $F : X \times I \rightarrow Y$ tal que:

$$F(x, 0) = f(x) \quad \text{e} \quad F(x, 1) = g(x)$$

para todo $x \in X$ onde I é o intervalo fechado, $I = [0, 1]$. O mapa F é chamado de **homotopia** entre f e g . Denotamos o fato de f ser homotópico a g por $f \simeq g$. Se $f \simeq g$ e g é o mapa constante, dizemos que f é nul-homotópica ou nulamente homotópica.

Uma homotopia é usualmente pensada como uma família de mapas de X em Y parametrizada por um parâmetro contínuo. Se imaginarmos o parâmetro t como o tempo, a homotopia F representa uma deformação contínua do mapa f para o mapa g conforme t vai de 0 a 1.

Agora, considere o caso especial onde f é um caminho em X . Isto é, f é um mapa contínuo de $I \rightarrow X$ tal que $f(0) = x_0$ e $f(1) = x_1$. Os pontos x_0 e x_1 são os pontos inicial e final respectivamente e diz-se que f é um caminho de x_0 a x_1 . Por conveniência usaremos o intervalo I como o domínio de todos os caminhos.

Se f e g são caminhos em X , existe uma relação mais forte que uma homotopia simples, chamada homotopia de caminhos definida como segue.

Definição B.2 – Homotopia de caminhos: Dois **caminhos** $f, g : I \rightarrow X$ são ditos **homotópicos** se eles possuem os mesmos pontos inicial e final e se existe um mapa contínuo $F : I \times I \rightarrow X$ tal que:

$$\begin{aligned} F(s, 0) = f(s) \quad \text{e} \quad F(s, 1) = g(s); \\ F(0, t) = x_0 \quad \text{e} \quad F(1, t) = x_1; \end{aligned}$$

para todo $s \in I$ e todo $t \in I$. F é chamada **homotopia de caminhos** entre f e g e representamos a relação entre f e g por $f \simeq_p g$. O subscrito p indica *path*, a palavra inglesa para caminho.

A primeira condição nos informa que F é uma homotopia entre f e g , ou seja, uma deformação contínua de f em g . A segunda condição indica que todos os caminhos intermediários devem ter como ponto inicial x_0 e ponto final x_1 . Uma representação da homotopia de caminhos entre f e g está mostrada na figura a seguir.

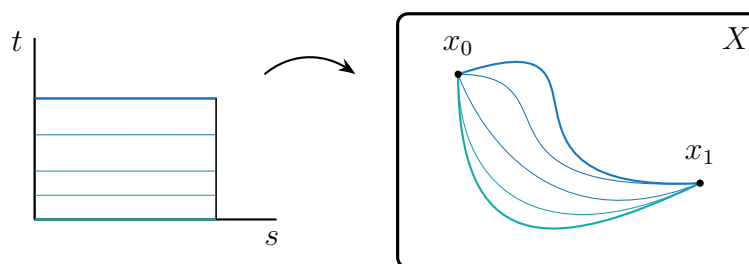


Figura B.1: Diagrama representando uma homotopia de caminhos.

Teorema B.1: As relações \simeq e \simeq_p são relações de equivalência.

Prova: Verifiquemos as propriedades de uma relação de equivalência.

- **Reflexividade:** Se $f : X \rightarrow Y$ é um mapa contínuo entre espaços topológicos. Tem-se que $f \simeq f$. O mapa $F : X \times I \rightarrow Y$, dado por $F(x, t) = f(x)$ é uma homotopia entre f e f . Se f for um caminho, F é uma homotopia de caminhos, logo $f \simeq_p f$.
- **Simetricidade:** Se $f \simeq g$ então existe uma homotopia F entre f e g . Assim, a homotopia $G(x, t) = F(x, 1 - t)$ é uma homotopia entre g e f , logo $g \simeq f$. Se f e g são caminhos, então F é uma homotopia de caminhos e, portanto, G também o é.
- **Transitividade:** Sejam $f, g, h : X \rightarrow Y$ mapas contínuos entre espaços topológicos. Suponha que $f \simeq g$ e $g \simeq h$. Para mostrar que $f \simeq h$, construímos uma homotopia entre f e h do seguinte modo. Sejam F e G as homotopias entre f e g e g e h respectivamente. Definimos o mapa $H : X \times I \rightarrow Y$ por:

$$H(x, t) = \begin{cases} F(x, 2t) & \text{para } t \in [0, 1/2], \\ G(x, 2t - 1) & \text{para } t \in [1/2, 1]. \end{cases}$$

O mapa H é bem definido pois para $t = 1/2$, tem-se: $F(x, 2t) = g(x) =$

$G(x, 2t - 1)$. Além disso, como H é contínuo nos subconjuntos fechados $X \times [0, 1/2]$ e $X \times [1/2, 1]$, é contínuo em todo o conjunto $X \times I$ pelo lema da colagem, A.12. Assim H é a homotopia requerida e $f \simeq h$.

Se F e G são homotopias de caminhos, então H também o é. Neste caso tem-se $f, g, h : X \rightarrow X$ com e $H : I \times I \rightarrow X$. Como $F(0, t) = G(0, t) = x_0$ e $F(1, t) = G(1, t) = x_1$ para todo $t \in I$, tem-se que $H(0, t) = x_0$ e $H(1, t) = x_1$ para todos os valores de t . Este resultado junto ao já estabelecido anteriormente fazem de H uma homotopia de caminhos. Portanto, $f \simeq_p h$. Um diagrama ilustrando o processo de construção de H para homotopias de caminho é mostrado na imagem a seguir.

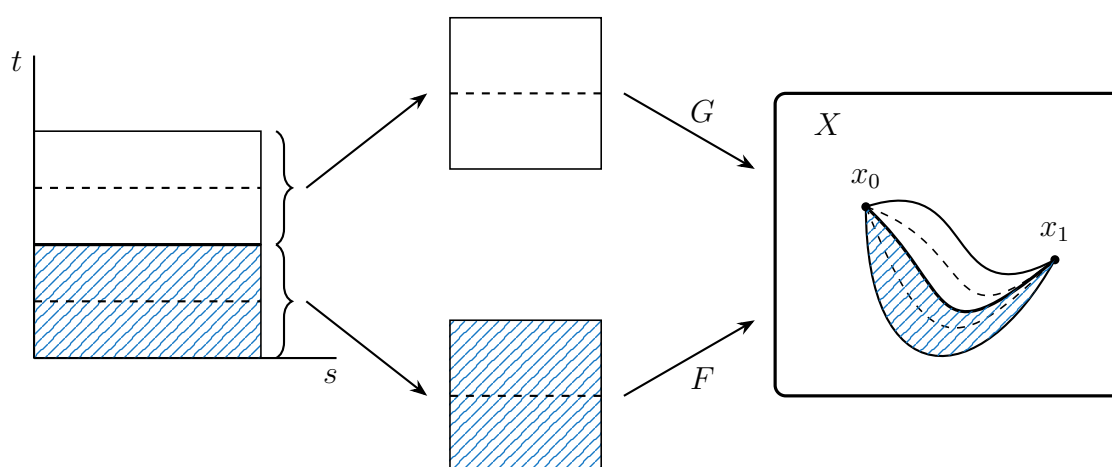


Figura B.2: Processo de construção de uma homotopia de caminho a partir duas outras, F e G .

□

Uma vez estabelecido que as relações de homotopia e homotopia de caminhos são relações de equivalência, o próximo passo é definir operações no conjunto das classes de equivalência de homotopia de caminho. Note que definiremos operações apenas no conjunto das classes de equivalência de homotopia de caminhos. Estas operações também podem ser definidas para homotopia simples com os ajustes adequados, no entanto isto não nos interessa aqui.

Definição B.3: Se f é um caminho de x_0 a x_1 e g , de x_1 a x_2 em um espaço topológico X , definimos o **produto** $f * g$ de f e g como o caminho h dado por:

$$h(s) = \begin{cases} f(2s) & \text{para } s \in [0, 1/2]; \\ g(2s - 1) & \text{para } s \in [1/2, 1]. \end{cases}$$

O mapa h é bem definido e contínuo pelo lema da colagem, A.12. É um caminho de x_0 a x_2 em X . Este caminho pode ser pensado simplesmente como uma colagem de f e g , a primeira metade do caminho é f e a segunda, g .

O produto $*$ induz uma operação bem definida no conjunto das classes de equivalência de caminhos em X . Se $f, g : I \rightarrow X$, esta operação é definida por:.

$$[f] * [g] = [f * g]$$

onde $[f], [g]$ são representantes da classe de equivalência de f e g sob homotopia de caminhos. Para verificar esta afirmação, considere os caminhos em X , f, f', g e g' tais que $f \simeq_p f'$ e $g \simeq_p g'$. Se F e G são as homotopias de caminho entre f e f' e g e g' , definimos

$$H(s, t) = \begin{cases} F(2s, t) & \text{para } s \in [0, 1/2]; \\ G(2s - 1, t) & \text{para } s \in [1/2, 1]. \end{cases}$$

Como $F(1, t) = x_1 = G(0, t)$ para todo t , o mapa H é bem definido. Além disso, ele é contínuo pelo lema da colagem, A.12. A verificação de que H é uma homotopia entre $f * g$ e $f' * g'$ é direta. A figura abaixo ilustra a definição de H .

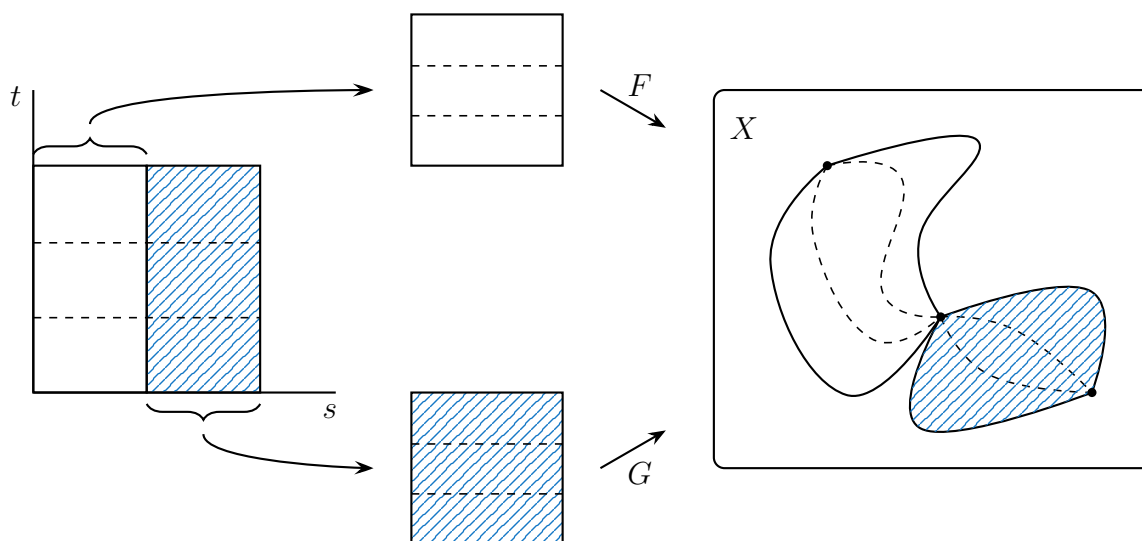


Figura B.3: Diagrama mostrando a construção da homotopia H entre $f * g$ e $f' * g'$.

Agora, note que

$$H(s, 0) = \begin{cases} F(2s, 0) = f(2s) & \text{para } s \in [0, 1/2]; \\ G(2s - 1, 0) = g(2s - 1) & \text{para } s \in [1/2, 1]; \end{cases}$$

Ou seja, $H(s, 0) = (f * g)(s)$. Do mesmo modo $H(s, 1) = (f' * g')(s)$. Assim, tem-se, que:

$$\begin{aligned} H(s, 0) &= (f * g)(s) & \text{e} & & H(s, 1) &= (f' * g')(s); \\ H(0, t) &= x_0 & \text{e} & & H(1, t) &= x_2; \end{aligned}$$

e H de fato é uma homotopia entre $f * g$ e $f' * g'$. Deste modo, fica estabelecido que se $f \simeq_p f'$ e $g \simeq_p g'$, então $f * g \simeq_p f' * g'$ e a operação $*$ é bem definida no conjunto de todas as classes de equivalência de caminhos em X . Esta operação nas classes homotopias de caminho satisfaz a propriedades muito similares às propriedades de grupo. Elas são chamadas **propriedades grupoide** de $*$ e diferem das propriedades de grupo, pois $[f] * [g]$ é definida somente para as classes de caminhos tais que $f(1) = g(0)$. Antes de estabelecermos as propriedades de grupoide de $*$, no entanto, são necessários alguns resultados preliminares.

Lema B.1 – Mapa linear positivo: Dados dois intervalos $[a, b], [c, d] \in \mathbb{R}$, existe um único mapa $p : [a, b] \rightarrow [c, d]$ da forma

$$p(x) = mx + k$$

tal que

$$p(a) = c \quad \text{e} \quad p(b) = d.$$

Este mapa é chamado **mapa linear positivo** de $[a, b]$ até $[c, d]$.

Prova: Primeiro, provamos a unicidade. Suponha que existam dois mapas $p, q : [a, b] \rightarrow [c, d]$:

$$\begin{aligned} p(x) &= mx + k, \\ q(x) &= nx + l \end{aligned}$$

tais que

$$\begin{aligned} p(a) &= c \quad \text{e} \quad p(b) = d, \\ q(a) &= c \quad \text{e} \quad q(b) = d. \end{aligned}$$

Então, podemos obter os seguinte sistema de equações:

$$\begin{aligned} a(m - n) + (k - l) &= 0, \\ b(m - n) + (k - l) &= 0. \end{aligned}$$

Subtraindo-se uma da outra, obtém-se $(a - b)(m - n) = 0$. Como $a \neq b$, devemos ter que $m = n$. Substituindo-se $m = n$ em qualquer uma das equações resulta em $k - l = 0$, de modo que $k = l$. Assim, o mapa p é único.

Quanto à existência, afirmamos que o mapa p com

$$m = \frac{d - c}{b - a} \quad \text{e} \quad k = \frac{ad - bc}{b - a}$$

satisfaz às condições. De fato, substituindo-se estes valores tem-se:

$$p(x) = \left(\frac{d - c}{b - a} \right) x - \frac{ad - bc}{b - a}$$

e fazendo-se $x = a$ e $x = b$, obtém-se, respectivamente

$$\begin{aligned} p(a) &= \left(\frac{d - c}{b - a} \right) a - \frac{ad - bc}{b - a} & p(b) &= \left(\frac{d - c}{b - a} \right) b - \frac{ad - bc}{b - a} \\ &= \frac{(ad - ac) - (ad - bc)}{b - a} & &= \frac{(bd - bc) - (ad - bc)}{b - a} \\ &= \frac{c(b - a)}{b - a} = c. & &= \frac{d(b - a)}{b - a} = d. \end{aligned}$$

□

Lema B.2: Sejam X e Y espaços topológicos, $k : X \rightarrow Y$ um mapa contínuo e F uma homotopia entre caminhos f e f' em X . Então, $k \circ F$ é uma homotopia entre $k \circ f$ e $k \circ f'$ em Y .

Prova:

A prova é direta. Note que se f e f' são caminhos em X , então tem-se são mapas de $I \rightarrow X$, assim, $k \circ f$ e $k \circ f'$ são mapas de $I \rightarrow Y$. Como k, f e f' são contínuas, as compostas $k \circ f$ e $k \circ f'$ também o são, e portanto, são caminhos em Y . Além disso, tem-se que $(k \circ F)(s, t) = k(F(s, t))$, logo:

$$\begin{aligned} (k \circ F)(s, 0) &= k(F(s, 0)) = k(f(s)) & \text{e} & \quad (k \circ F)(s, 1) = k(F(s, 1)) = (k(f'(s))); \\ (k \circ F)(0, t) &= k(F(0, t)) = k(x_0) & \text{e} & \quad (k \circ F)(1, t) = k(F(1, t)) = k(x_1); \end{aligned}$$

onde $x_0, x_1 \in X$ são os pontos inicial e final de f e f' . Como k e F são contínuas, sua composta também é contínua, de modo que $k \circ F$ é uma homotopia entre os caminhos $k \circ f$ e $k \circ f'$ de Y . A prova está ilustrada na figura abaixo.

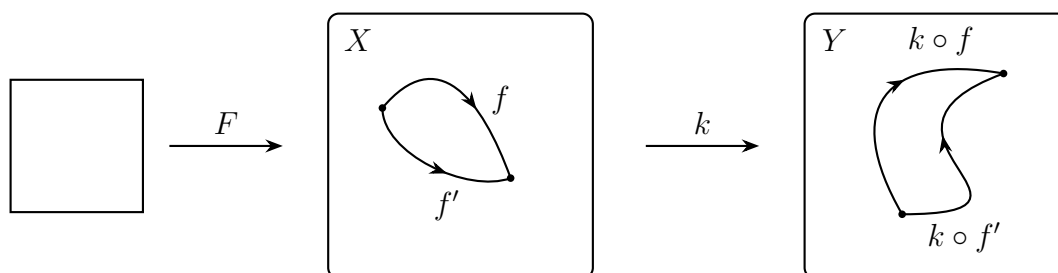


Figura B.4: Mapeamento de uma homotopia de caminhos de X em uma de caminhos de Y .

□

Lema B.3: Sejam X e Y espaços topológicos e $k : X \rightarrow Y$ um mapa contínuo. Se $f, g : I \rightarrow X$ são caminhos em X , com $f(1) = g(0)$, então:

$$k \circ (f * g) = (k \circ f) * (k \circ g).$$

Prova: Primeiro note que por k, f e g serem contínuos, as compostas $k \circ f$ e $k \circ g$ também o são, de modo que $k \circ f$ e $k \circ g$ são caminhos em Y . Além disso, $f(1) = g(0)$ garante que $(k \circ f)(1) = (k \circ g)(0)$. A prova então decorre da definição do produto $*$. O caminho $f * g$ é dado por:

$$(f * g)(s) = \begin{cases} f(2s) & \text{para } s \in [0, 1/2]; \\ g(2s - 1) & \text{para } s \in [1/2, 1]. \end{cases}$$

Portanto a composta $k \circ (f * g)$ é dada por

$$[k \circ (f * g)](s) = \begin{cases} (k \circ f)(2s) & \text{para } s \in [0, 1/2]; \\ (k \circ g)(2s - 1) & \text{para } s \in [1/2, 1]. \end{cases}$$

Mas, como $k \circ f$ e $k \circ g$ são caminhos, esta é precisamente a definição de $(k \circ f) * (k \circ g)$.

□

Teorema B.2: A operação $*$ possui as seguintes propriedades:

i **Associatividade:** Se $[f] * ([g] * [h])$ está definido, então $([f] * [g]) * [h]$ também está e ambos os produtos são iguais.

ii **Identidades:** Dado $x \in X$, seja e_x o caminho constante $e_x : I \rightarrow X$, dado por

$e_x(s) = x$. Se f é um caminho em X de x_0 a x_1 então

$$[f] * [e_{x_1}] = [f] \quad \text{e} \quad [e_{x_0}] * [f] = [f].$$

iii **Inversa**: Dado um caminho f em X de x_0 a x_1 , defina \bar{f} como o caminho em X tal que $\bar{f}(s) = f(1 - s)$. Este caminho é chamado de **caminho inverso** a f e tem-se

$$[f] * [\bar{f}] = [e_{x_0}] \quad \text{e} \quad [\bar{f}] * [f] = [e_{x_1}].$$

Prova:

Para provar este teorema usaremos os lemas B.2 e B.3. Em particular, provaremos estas propriedades para caminhos em um espaço específico, I e então usaremos os lemas para transportá-las para um espaço arbitrário X , interpretando um caminho $f : I \rightarrow X$ explicitamente como uma mapa de I para X . Começamos por ii e iii.

Para verificar ii, denotaremos por e_0 o caminho constante em I no ponto 0, isto é, $e_0 : I \rightarrow I$, tal que $e_0(s) = 0$. E denotamos por $i : I \rightarrow I$ o mapa identidade, que é um caminho em I de 0 a 1. Assim, o $e_0 * i$ também é um caminho em I de 0 a 1, como mostrado na imagem abaixo:

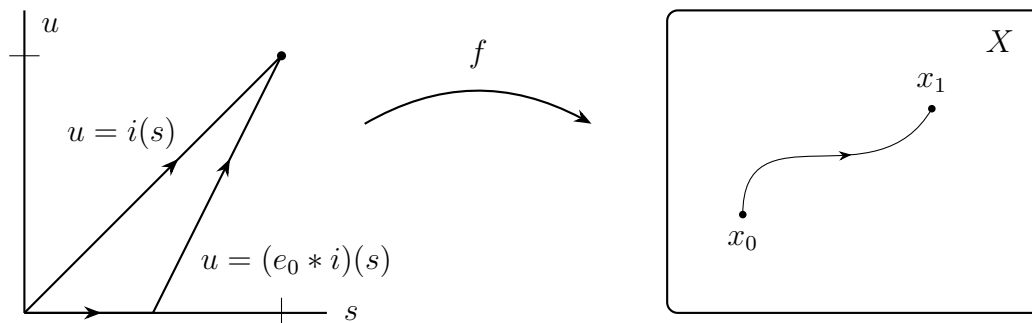


Figura B.5: Caminhos em I mapeados em um único caminho em X por f .

Agora, note que os caminhos i e $e_0 * i$ são homotópicos em I . Uma homotopia entre estes caminhos é dada por $G = (1 - t)i(s) + t(e_0 * i)(s)$. De fato, tem-se:

$$G(s, 0) = i(s) \quad \text{e} \quad G(s, 1) = (e_0 * i)(s)$$

e

$$G(0, t) = (1 - t)i(0) + t(e_0 * i)(0) = 0;$$

$$G(1, t) = (1 - t)i(1) + t(e_0 * i)(1) = 1 - t + t = 1.$$

Esta homotopia é chamada **homotopia linear** e é bastante útil. Agora note que

$f \circ i = f$ e, pelo lema B.3, $f \circ (e_0 * i) = (f \circ e_0) * (f \circ i) = e_{x_0} * f$, são caminhos em X , já que f é contínua. Além disto, $f \circ G$ é uma homotopia entre estes dois caminhos, pelo lema B.2, de modo que:

$$[e_{x_0} * f] = [e_{x_0}] * [f] = [f].$$

A prova de que $[f] * [e_{x_1}] = [f]$ é inteiramente análoga. Basta definir o caminho constante $e_1 : I \rightarrow I$, dado por $e_1(s) = 1$ e realizar o mesmo processo para $i * e_1$.

Para provar iii, note que o caminho inverso ao caminho identidade, i em I é dado por $\bar{i}(s) = i(1 - s) = 1 - s$. Assim, o caminho $i * \bar{i}$ é tal que começa e termina em 0, assim como o caminho constante e_0 . A imagem abaixo ilustra estes dois caminhos.

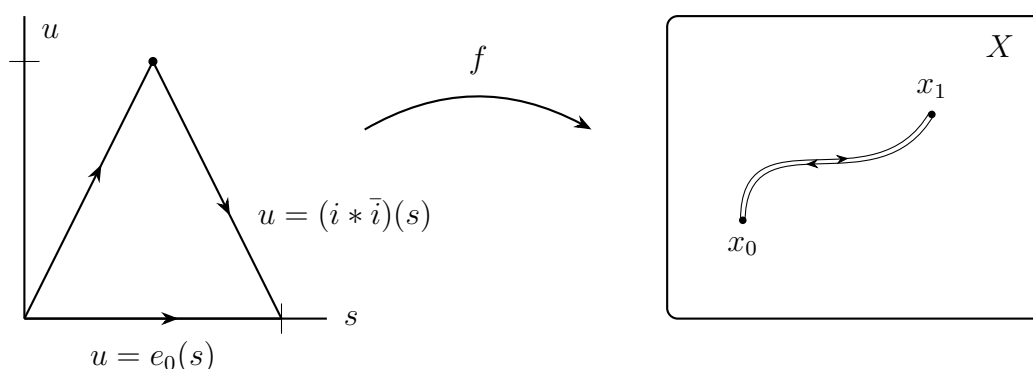


Figura B.6: Caminhos em I mapeados no caminho constante, e_{x_0} e $f * \bar{f}$.

Novamente, estabelecemos uma homotopia linear H entre e_0 e $i * \bar{i}$ em I , de modo que $f \circ H$ é uma homotopia entre os caminhos $f \circ e_0 = e_{x_0}$ e

$$f \circ (i * \bar{i}) = (f \circ i) * (f \circ \bar{i}) = f * \bar{f}.$$

De modo que

$$[f * \bar{f}] = [f] * [\bar{f}] = [e_{x_0}].$$

A prova de que $[\bar{f}] * [f] = [e_{x_1}]$ é completamente análoga, basta usar o fato de que $\bar{i} * i$ é homotópico em I a e_1 .

Para provar i, começamos por usar o lema B.1 para definir um produto triplo de caminhos $f, g, h : I \rightarrow X$. Primeiro, note que o produto $*$ de f e g pode ser definido usando mapas lineares positivos da seguinte forma:

$$(f * g) = \begin{cases} (f \circ p_{[0,1/2]})(s) & \text{se } s \in [0, 1/2]; \\ (g \circ p_{[1/2,1]})(s) & \text{se } s \in [1/2, 1]; \end{cases}$$

onde $p_{[a,b]} : [a, b] \rightarrow I$ denota o mapa linear positivo do intervalo $[a, b] \subset \mathbb{R}$ para

I . Para mostrar a associatividade, então, supomos que $f(1) = g(0)$ e $g(1) = h(0)$ e definimos caminhos $k_{ab} : I \rightarrow X$, onde $0 < a < b < 1$, do seguinte modo:

$$k_{ab} = \begin{cases} (f \circ p_{[0,a]})(s) & \text{se } s \in [0, a]; \\ (g \circ p_{[a,b]})(s) & \text{se } s \in [a, b]; \\ (h \circ p_{[b,1]})(s) & \text{se } s \in [b, 1]. \end{cases}$$

Pela definição de $*$ é direto que $f * (g * h)$ corresponde a $a = 1/2$ e $b = 3/4$ e o caso $(f * g) * h$ corresponde a $a = 1/4$ e $b = 1/2$. O caminho k_{ab} depende da escolha de a e b , no entanto, para quaisquer $a, b \in I$, os caminhos k_{ab} pertencem à mesma classe de homotopia. Para ver isto, considere o mapa $p : I \rightarrow I$ dado por:

$$p = \begin{cases} p_{[0,a]}^{[0,c]}(s) & \text{se } s \in [0, a]; \\ p_{[a,b]}^{[c,d]}(s) & \text{se } s \in [a, b]; \\ p_{[b,1]}^{[d,1]}(s) & \text{se } s \in [b, 1]; \end{cases}$$

onde $p_{[a,b]}^{[c,d]}$ é o mapa linear positivo de $[a, b] \rightarrow [c, d]$. Este mapa está ilustrado na figura abaixo.

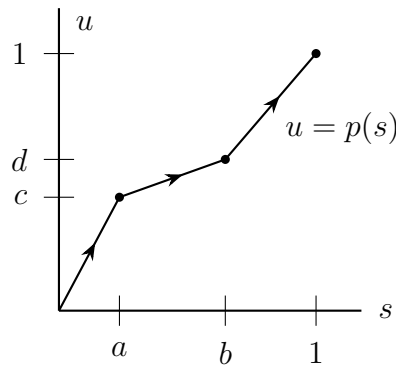


Figura B.7: Caminho $p : I \rightarrow I$.

Como $p([0, a]) = [0, c]$, $p([a, b]) = [c, d]$ e $p([b, 1]) = [d, 1]$, tem-se que:

$$k_{cd} \circ p = k_{ab}.$$

Além disso, o caminho p vai de 0 a 1 em I , assim como o caminho identidade, i . Se P é a homotopia linear entre estes dois caminhos, $k_{cd} \circ P$ é uma homotopia entre os caminhos k_{cd} e k_{ab} em X , de modo que:

$$[k_{ab}] = [k_{cd}].$$

Escolhendo $a = 1/2, b = 3/4, c = 1/4$ e $d = 1/2$, obtém-se, por fim, que:

$$[f * (g * h)] = [(f * g) * h].$$

□

Com este teorema temos o suficiente para definir e estudar o grupo fundamental ou primeiro grupo de homotopia. Antes disso, no entanto, abordaremos o conceito de espaço contrátil, uma propriedade bastante útil para entender intuitivamente o grupo fundamental de um espaço.

B.2 - Grupo fundamental

O teorema B.2 estabelece que a operação $*$ satisfaz todas os axiomas de grupo exceto pelo fato de não estar definida para todos as classes de equivalência $[f]$ e $[g]$, somente aquelas para as quais $f(1) = g(0)$. Assim, se escolhermos um ponto $x \in X$ e considerarmos apenas loops, isto é, os caminhos que começam e terminam em x . então $*$ sempre estará definida tanto para os caminhos quanto para suas classes de equivalência. Assim tem-se a seguinte definição:

Definição B.4: Seja X um espaço e $x \in X$ um ponto. Um caminho em X que começa e termina em x é um loop sitiado em x . O conjunto de todas as classes de homotopia de caminhos de loops sitiados em x com a operação $*$ forma o **grupo fundamental** de X relativo ao ponto base x . O denotamos por $\pi_1(X, x)$.

Notamos que um espaço X com um ponto $x \in X$ selecionado é chamado de **espaço pontuado** e escrevemos (X, x) para denotar este espaço. Neste sentido, o grupo fundamental é definido para cada espaço pontuado obtido a partir de X , isto é, para cada escolha de ponto base, x . Podemos nos perguntar então, em que medida o grupo fundamental depende da escolha do ponto base. Para estudar isto definimos um mapa $\hat{\alpha}$ para cada caminho α de x_0 a x_1 por:

$$\begin{aligned} \hat{\alpha} : \pi_1(X, x_0) &\rightarrow \pi_1(X, x_1) \\ [f] &\mapsto [\bar{\alpha}] * [f] * [\alpha]. \end{aligned}$$

Note que como f é um loop sitiado em x_0 , $\bar{\alpha} * f * \alpha$ é um loop sitiado em x_1 como indicado a seguir.

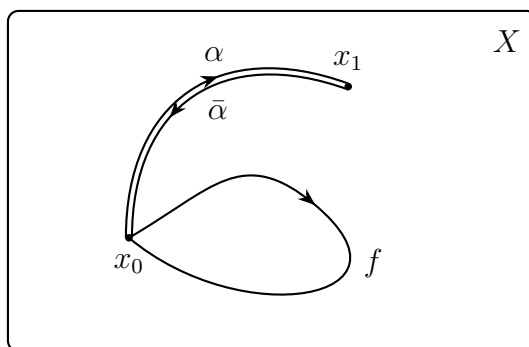


Figura B.8: Loop gerado pelo mapa $\hat{\alpha}$.

Teorema B.3: O mapa $\hat{\alpha}$ é um isomorfismo de grupos.

Prova: Primeiro, mostramos que $\hat{\alpha}$ é um homomorfismo.

$$\begin{aligned}\hat{\alpha}([f]) * \hat{\alpha}([g]) &= ([\bar{\alpha}] * [f] * [\alpha]) * ([\bar{\alpha}] * [g] * [\alpha]) \\ &= [\bar{\alpha}] * [f] * [g] * [\alpha] = \hat{\alpha}([f] * [g]).\end{aligned}$$

Em seguida, para mostrar que $\hat{\alpha}$ é um isomorfismo, considere o mapa $\hat{\beta}$ onde $\beta = \bar{\alpha}$, o caminho inverso a α . Tem-se:

$$\hat{\beta}([h]) = [\bar{\beta}] * [h] * [\beta] = [\alpha] * [h] * [\bar{\alpha}].$$

De modo que

$$\begin{aligned}\hat{\alpha}(\hat{\beta}([h])) &= [\bar{\alpha}] * ([\alpha] * [h] * [\bar{\alpha}]) * [\alpha] = [h] \\ \hat{\beta}(\hat{\alpha}([h])) &= [\alpha] * ([\bar{\alpha}] * [h] * [\alpha]) * [\bar{\alpha}] = [h].\end{aligned}$$

Ou seja, $\hat{\alpha}$ e $\hat{\beta}$ são mapas inversos e, portanto, isomorfismos de grupos. □

Vemos então que se existe um caminho conectando os pontos x_0 e x_1 os grupos fundamentais sitiados nestes dois pontos são isomorfos e escrevemos: $\pi_1(X, x_0) \cong \pi_1(X, x_1)$. Como corolário deste teorema tem-se:

Corolário B.3.1: Se X é conexo por caminhos e $x_0, x_1 \in X$, então, $\pi_1(X, x_0) \cong \pi_1(X, x_1)$.

Considere, então um espaço X e seja C uma componente de caminhos de X contendo x_0 . Tem-se que $\pi_1(C, x_0) = \pi_1(X, x_0)$ já que todos os loops sitiados em x_0 devem, obrigatoriamente, estar contidos no subespaço C . Assim, o grupo fundamental sitiado em x_0 depende apenas da componente de caminhos de X que contem x_0 . Ele não fornece

nenhuma informação sobre o resto de X . Por este motivo quando se estuda o grupo fundamental é comum fazê-lo apenas em espaços conexos por caminhos. Por sua vez, se X é um espaço conexo por caminhos todos os grupos $\pi_1(X, x)$ são isomorfos e é tentador identificá-los e falar apenas de $\pi_1(X)$, sem referência ao ponto base. O problema é que não existe uma maneira natural de identificar $\pi_1(X, x_0)$ com $\pi_1(X, x_1)$ pois caminhos diferentes α e β de x_0 a x_1 podem gerar isomorfismos diferentes. Por este motivo, omitir o ponto base pode levar a erros. Existe, porém, uma condição sobre a qual o isomorfismo entre $\pi_1(X, x_0)$ e $\pi_1(X, x_1)$ é independente do caminho escolhido:

Teorema B.4: Seja X um espaço conexo por caminhos e $x_0, x_1 \in X$, então $\hat{\alpha} = \hat{\beta}$ para cada par de caminhos α e β de x_0 a x_1 se, e somente se, $\pi_1(X, x_0)$ for abeliano.

Prova: Começamos por supor que $\pi_1(X, x_0)$ é abeliano. Neste caso, dado dois isomorfismos $\hat{\alpha}$ e $\hat{\beta}$ podemos escrever, para qualquer loop f sitiado em x_0 :

$$\begin{aligned}\hat{\alpha}([f]) &= ([\bar{\alpha}] * [f] * [\beta]) * ([\bar{\beta}] * [\alpha]) \\ &= ([\bar{\beta}] * [\alpha]) * ([\bar{\alpha}] * [f] * [\beta]) = \hat{\beta}([f]).\end{aligned}$$

Note que $[\bar{\alpha}] * [f] * [\beta]$ e $[\bar{\beta}] * [\alpha]$ são loops sitiados em x_1 . A passagem para a segunda linha é feita baseada no fato de que se $\pi_1(X, x_0)$ é abeliano, $\pi_1(X, x_1)$ também o é já que são isomorfos por X ser conexo por caminhos. Por outro lado, se $\hat{\alpha}$ e $\hat{\beta}$ são iguais para quaisquer α e β , podemos escrever, para qualquer caminho $[f]$ que:

$$[\bar{\alpha}] * [f] * [\alpha] = [\bar{\beta}] * [f] * [\beta]$$

de modo que:

$$([\bar{\alpha}] * [f] * [\beta]) * ([\bar{\beta}] * [\alpha]) = ([\bar{\beta}] * [\alpha]) * ([\bar{\alpha}] * [f] * [\beta]).$$

Como isto vale para qualquer $[f]$ e quaisquer pares de caminho α e β , vale para todos os elementos de $\pi_1(X, x_1)$, logo $\pi_1(X, x_1)$ é abeliano e, portanto, $\pi_1(X, x_0)$ também o é. □

O grupo fundamental nos permite definir a seguinte propriedade de um espaço conexo por caminhos:

Definição B.5: Um espaço X é dito **simplesmente conexo** se é conexo por caminhos e $\pi_1(X, x_0)$ é o grupo trivial, o grupo com um único elemento e escrevemos $\pi_1(X, x_0) = \mathbf{0}$.

Note que devido à exigência de X ser conexo por caminhos, se $\pi_1(X, x_0) = \mathbf{0}$ para algum x_0 , então $\pi_1(X, x) \cong \pi_1(X, x_0) = \mathbf{0}$ para todo $x \in X$. Em particular, como $\mathbf{0}$ é abeliano, podemos simplesmente escrever $\pi_1(X) = 0$ em vista do teorema B.4 se X for simplesmente conexo.

Proposição B.1: Se X é simplesmente conexo, então quaisquer dois caminhos possuindo os mesmos pontos inicial e final são homotópicos.

Prova: Sejam α e β dois caminhos de x_0 a x_1 . O caminho $\alpha * \bar{\beta}$ está definido e é um loop sitiado em x_0 . Como X é simplesmente conexo, este loop é homotópico ao caminho constante em x_0 , e_{x_0} . Então:

$$[\alpha * \bar{\beta}] * [\beta] = [e_{x_0}] * [\beta]$$

donde concluímos que $[\alpha] = [\beta]$.

□

Intuitivamente, é simples de ver que o grupo fundamental é um invariante topológico de um espaço. Para formalizarmos este resultado, introduzimos o conceito de **homomorfismo induzido por um mapa contínuo**. Considere dois espaços pontuados (X, x_0) e (Y, y_0) . Se $h : X \rightarrow Y$ é um mapa contínuo tal que $h(x_0) = y_0$, escrevemos

$$h : (X, x_0) \rightarrow (Y, y_0)$$

Se f é um loop sitiado em x_0 então a composta: $h \circ f : I \rightarrow Y$ é um loop em Y sitiado em y_0 . A correspondência $f \rightarrow h \circ f$, portanto, origina um mapa de $\pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$.

Definição B.6: Seja $h : (X, x_0) \rightarrow (Y, y_0)$ um mapa contínuo. Definimos o mapa

$$h_* : \pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$$

pela equação

$$h_*([f]) = [h \circ f].$$

O mapa h_* é chamado de **homomorfismo induzido por h** relativo ao ponto base x_0 .

O mapa h_* é bem definido pois, se F é uma homotopia de caminhos entre caminhos f e g em X , então $h \circ F$ é uma homotopia entre os caminhos $h \circ f$ e $h \circ g$ em Y , como mostrado pelo lema B.2. Além disso h_* é um homomorfismo devido ao lema B.3 isto é:

$$(h \circ f) * (h \circ g) = h \circ (f * g)$$

O homomorfismo h_* depende não somente do mapa $h : X \rightarrow Y$ como também do ponto base, x_0 . Note que uma vez que escolhamos tanto h quanto x_0, y_0 fica determinado. Mesmo que X seja conexo por caminhos, de modo que os grupos fundamentais sitiados em dois pontos distintos, x_0 e x_1 sejam isomorfos, eles ainda são grupos distintos de modo que $h_* : \pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$ é distinto de $h_* : \pi_1(X, x_1) \rightarrow \pi_1(Y, y_1)$. Por fim, podemos usar homomorfismo induzidos para mostrar as seguintes propriedades bastante úteis:

Teorema B.5: Se $h : (X, x_0) \rightarrow (Y, y_0)$ e $k : (Y, y_0) \rightarrow (Z, z_0)$ são contínuos, então $(k \circ g)_* = k_* \circ h_*$. Se $i : (X, x_0) \rightarrow (X, x_0)$ é o mapa identidade, então i_* é o homomorfismo identidade.

Prova: A demonstração segue da definição. Tem-se:

$$(k \circ h)_*([f]) = [(k \circ h) \circ f]$$

enquanto

$$(k_* \circ h_*)([f]) = k_*(h_*([f])) = k_*([h \circ f]) = [k \circ (h \circ f)].$$

Do mesmo modo: $i_*([f]) = [i \circ f] = [f]$.

□

As duas propriedades acima são ditas propriedades **functoriais**, essencialmente porque elas transportam a composição de mapas contínuos entre espaços pontuados em composição de homomorfismos induzidos por estes mapas, além de levar o mapa identidade no homomorfismo identidade. Ou seja essencialmente carregam a composição de mapas contínuos para a composição de homomorfismos induzidos. Deste teorema segue que:

Corolário B.5.1: Se $h : (X, x_0) \rightarrow (Y, y_0)$ é um homeomorfismo entre X e Y , então h_* é um isomorfismo entre $\pi_1(X, x_0)$ e $\pi_1(Y, y_0)$.

Prova: A prova é direta. Sejam i_* e j_* os homomorfismos identidade de $\pi_1(X, x_0)$ e $\pi_1(Y, y_0)$ respectivamente. Então:

$$\begin{aligned} h_*^{-1} \circ h_* &= (h^{-1} \circ h)_* = i_* \\ h_* \circ h_*^{-1} &= (h \circ h^{-1})_* = j_* \end{aligned}$$

Ou seja, h_* e h_*^{-1} são um par de isomorfismos inversos.

□

Este corolário estabelece o grupo fundamental como um invariante topológico. Espaços conectados por um homeomorfismo possuem grupos fundamentais isomorfos. Note que o isomorfismo é entre os grupos fundamentais sitiados em x_0 e $y_0 = h(x_0)$, onde $h : X \rightarrow Y$

é o homeomorfismo em questão. Se ambos X e Y forem conexos por caminhos, então os grupos fundamentais para quaisquer pontos x_0, y_0 são isomorfos.

Por último, estabelecemos a seguintes resultados sobre o grupo fundamental quando tratamos de espaços que são obtidos a partir do produto cartesiano de outros. O lema abaixo ilustra que caminhos definidos no produto de espaços podem ser tratados essencialmente, como o produto de caminhos em cada um dos espaços.

Lema B.4: Sejam X e Y espaços e $f, g : I \rightarrow X \times Y$ caminhos em $X \times Y$ de $x_0 \times y_0$ a $x_1 \times y_1$. Um mapa $F : I \times I \rightarrow X \times Y$ é uma homotopia de caminhos entre f e g se, e somente se, $\pi^1 \circ F$ for uma homotopia de caminhos entre $\pi^1 \circ f$ e $\pi^1 \circ g$ e $\pi^2 \circ F$ for uma homotopia de caminhos entre $\pi^2 \circ f$ e $\pi^2 \circ g$ respectivamente. Os mapas $\pi^1 : X \times Y \rightarrow X$ e $\pi^2 : X \times Y \rightarrow Y$ são as projeções canônicas.

Prova: Primeiro suponha que F seja uma homotopia de caminhos entre f e g . Como as projeções canônicas são contínuas, o lema B.2 estabelece que $\pi^1 \circ F$ é uma homotopia de caminhos entre $\pi^1 \circ f$ e $\pi^1 \circ g$ e que $\pi^2 \circ F$ é uma homotopia de caminhos entre $\pi^2 \circ f$ e $\pi^2 \circ g$. Por outro lado, supondo que $\pi^1 \circ F$ e $\pi^2 \circ F$ são homotopias de caminhos, então:

$$F(s, 0) = (\pi^1 \circ F)(s, 0) \times (\pi^2 \circ F)(s, 0) = (\pi^1 \circ f)(s) \times (\pi^2 \circ f)(s);$$

$$F(s, 1) = (\pi^1 \circ F)(s, 1) \times (\pi^2 \circ F)(s, 1) = (\pi^1 \circ g)(s) \times (\pi^2 \circ g)(s);$$

$$F(0, t) = (\pi^1 \circ F)(0, t) \times (\pi^2 \circ F)(0, t) = x_0 \times y_0;$$

$$F(1, t) = (\pi^1 \circ F)(1, t) \times (\pi^2 \circ F)(1, t) = x_1 \times y_1.$$

Ou seja, F é uma homotopia de caminhos entre f e g . □

Proposição B.2: Se X e Y espaços então $\pi_1(X \times Y, x_0 \times y_0)$ é isomorfo a $\pi_1(X, x_0) \times \pi_1(Y, y_0)$.

Prova: Sejam $\pi^1 : X \times Y \rightarrow X$ e $\pi^2 : X \times Y \rightarrow Y$ as projeções canônicas. Considere o mapa $\phi : X \times Y \rightarrow X \times Y$ dado por $\phi = \pi^1 \times \pi^2$. Este mapa é contínuo pois as projeções são contínuas, de modo que ϕ induz um homomorfismo:

$$\begin{aligned} \phi_* : \pi_1(X \times Y, x_0 \times y_0) &\rightarrow \pi_1(X, x_0) \times \pi_1(Y, y_0) \\ [f] &\mapsto \pi_*^1([f]) \times \pi_*^2([f]). \end{aligned}$$

Verifiquemos que ϕ_* de fato é um homomorfismo. Sejam f e g dois loops em $X \times Y$,

então:

$$\begin{aligned}
 \phi_*([f] * [g]) &= \pi_*^1([f] * [g]) \times \pi_*^2([f] * [g]) \\
 &= [\pi^1 \circ (f * g)] \times [\pi^2 \circ (f * g)] \\
 &= [(\pi^1 \circ f) * (\pi^1 \circ g)] \times [(\pi^2 \circ f) * (\pi^2 \circ g)] \\
 &= [\pi^1 \circ f] * [\pi^1 \circ g] \times [\pi^2 \circ f] * [\pi^2 \circ g] \\
 &= [\pi^1 \circ f] \times [\pi^2 \circ f] * [\pi^1 \circ g] \times [\pi^2 \circ g] = \phi_*([f]) * \phi_*([g])
 \end{aligned}$$

Onde fizemos uso do lema B.3 na passagem para a terceira linha e a operação $*$ nas duas últimas linhas é aquela definida em $\pi_1(X, x_0) \times \pi_1(Y, y_0)$, isto é, o produto de caminhos de $X \times Y$ é o produto de suas componentes. Por fim, note que ϕ é injetiva e sobrejetiva, de modo que ϕ_* estabelece um isomorfismo entre grupos.

□