ELEMENTOS DE MATEMÁTICA

Pedro G. Mattos

Elementos de Matemática

Pedro G. Mattos

Elementos de Matemática

Conjuntos, álgebra, geometria e dinâmica

Livros Livres

Pedro G. Mattos
pedrogmattos@ime.unicamp.br
Campinas, SP, Brasil

Arte da capa: *Scuola di Atene* (1511), Raffaello Sanzio.

Versão 0.20. Última revisão em 4 de novembro de 2023.

Licença Creative Commons (CC BY-SA 4.0).
Este livro pode ser compartilhado e redistribuído em qualquer suporte ou formato, adaptado, transformado e reeditado sob as seguintes condições.
Atribuição — Você deve dar o crédito apropriado, prover um link para a licença e indicar se mudanças foram feitas. Você deve fazê-lo em qualquer circunstância razoável, mas de maneira alguma que sugira ao licenciante a apoiar você ou o seu uso.
Compartilha Igual — Se você reorganizar, transformar ou criar a partir do material, tem de distribuir as suas contribuições sob a mesma licença que o original.
Sem restrições adicionais — Você não pode aplicar termos jurídicos ou medidas de caráter tecnológico que restrinjam legalmente outros de fazerem algo que a licença permita.
Para mais informações sobre a licença acesse:
https://creativecommons.org/licenses/by-sa/4.0/deed.pt_BR ou
https://wiki.creativecommons.org/wiki/Brazil.

Conteúdo

Figuras xxi

Tabelas xxii

1 Conjuntos 1

1 Os axiomas e as construções essenciais 2
1.1 Preliminares de lógica . . . 2
1.1.1 Lógica de ordem 0 . . . 3
1.1.2 Lógica de ordem 1 . . . 5
1.1.2.1 Abreviações . . . 7
1.1.2.2 Regras de inferência . . . 7
1.1.2.3 Igualdade e unicidade . . . 9
1.1.3 Conjunto e pertencimento . . . 10
1.1.3.1 Abreviações . . . 10
1.2 Axiomas do vazio e da extensão . . . 11
1.2.1 Vazio e igualdade . . . 11
1.2.2 Subconjuntos . . . 12
1.3 Axioma da especificação . . . 13
1.4 Axioma das partes . . . 14
1.5 Axioma do par . . . 14
1.6 Axioma da união . . . 15
1.6.1 União de um conjunto . . . 15
1.6.2 Interseção de um conjunto . . . 16
1.7 Axioma do infinito . . . 17
1.8 Axioma da escolha . . . 18
1.9 Axiomas da fundação e substituição . . . 20

2 Famílias e propriedades de conjuntos 22
2.1 Famílias e indexações . . . 22

2.2 Propriedades de união e interseção 24
2.3 Produto de conjuntos 26
2.4 Coproduto de conjuntos 27
2.4.1 Propriedades de produto e coproduto 28
2.5 Complementares e diferença simétrica 29
2.5.1 Propriedades 29

3 Relações **31**
3.1 Funções 31
3.1.1 Definição e propriedades básicas 32
3.1.2 Composição de funções 33
3.1.3 Função inversa, injetividade e sobrejetividade 34
3.1.4 Imagem inversa de função e propriedades 36
3.1.5 Propriedades de imagem e imagem inversa 37
3.1.6 Os funtores imagem e imagem inversa 38
3.2 Equivalências 39
3.2.1 Funções bem definidas 40
3.3 Ordens 41
3.3.1 Ordens parciais, estritas e totais 41
3.3.2 Conjuntos parcialmente ordenados 43
3.3.3 Funções monótonas 45
3.3.3.1 Multi-índices estritamente crescentes 46
3.3.4 Conjuntos totalmente ordenados e cadeias 46
3.3.5 Conjuntos bem ordenados 48
3.3.5.1 Números ordinais 49
3.3.6 Pré-ordens 49
3.3.7 Conjunto direcionado 50
3.3.8 Reticulados 50
3.3.9 Álgebras booleanas 51
3.3.9.1 Função indicadora 53
3.4 Coberturas e partições 54
3.4.1 Coberturas e subcoberturas 54
3.4.2 Refinamento de coberturas 54
3.4.3 Correfinamento de coberturas 55
3.4.4 Coberturas puxadas 57
3.4.5 Partições 58
3.4.6 Refinamento de partições 59
3.4.7 Correfinamento de partições 60

4 Cardinalidade de conjuntos **62**
4.1 Relações . . . 62
4.1.1 Igualdade de cardinais . . . 62
4.1.2 Ordenação de cardinais . . . 63
4.2 Operações . . . 64
4.2.1 Cardinalidade de soma (ou união disjunta) . . . 64

2 Álgebra **67**

5 Estruturas básicas **68**
5.1 Conjuntos numéricos . . . 68
5.1.1 Números naturais . . . 68
5.1.1.1 Adição . . . 69
5.1.1.2 Multiplicação . . . 71
5.1.1.3 Ordenação . . . 74
5.1.2 Números inteiros . . . 76
5.1.2.1 Adição e subtração . . . 77
5.1.2.2 Multiplicação . . . 79
5.1.2.3 Ordenação . . . 79
5.2 Grupoides . . . 80
5.2.1 Operação binária e grupoide . . . 80
5.2.2 Associatividade . . . 81
5.2.3 Comutatividade . . . 88
5.2.4 Subgrupoide . . . 91
5.2.5 Homomorfismo de grupoide . . . 92
5.3 Grupídeos . . . 93
5.3.1 Identidades e grupídeo . . . 93
5.3.2 Monoide (grupídeo associativo) . . . 95
5.3.3 Subgrupídeo . . . 95
5.3.4 Homomorfismos de grupídeo . . . 96
5.4 Grupíneo . . . 97
5.4.1 Inversos e grupíneo . . . 97

6 Grupos **102**
6.1 Conceitos básicos . . . 102
6.1.1 Grupo e subgrupo . . . 102
6.1.2 Coclasses e índice de subgrupo . . . 106
6.1.3 Subgrupo normal e grupo quociente . . . 108
6.1.4 Homomorfismo de grupo . . . 110
6.1.5 Núcleo, imagem e isomorfismo . . . 112

6.1.6 Teoremas de isomorfismo . . . 114
6.2 Construções categóricas . . . 116
6.2.1 Produto de grupos . . . 116
6.2.2 Grupo livre . . . 117
6.2.3 Coproduto de grupos . . . 119
6.3 Construções específicas . . . 121
6.3.1 Translações e conjugações . . . 121
6.3.2 Centro, automorfismos internos e externos . . . 122
6.3.3 Centralizador e normalizador . . . 122
6.3.4 Produto semidireto . . . 122
6.3.5 Grupo simples e subgrupo normal maximal . . . 124
6.3.6 Sequência subnormal . . . 125
6.3.7 Conjunto gerador . . . 125
6.4 Ação de grupos . . . 126
6.4.1 Grupo simétrico . . . 126
6.4.2 Ações . . . 128
6.4.3 Órbitas e estabilizadores . . . 130
6.4.4 Permutações finitas . . . 130
6.4.4.1 Ciclos e transposições . . . 132
6.5 Grupo linear geral . . . 132
6.6 Representação de grupos . . . 133
6.7 Sequências Exatas . . . 133

7 Anéis **135**
7.1 Conceitos básicos . . . 135
7.1.1 Anel e subanel . . . 135
7.1.2 Ideais e anéis quocientes . . . 138
7.1.3 Homomorfismo de anel . . . 142
7.1.4 Teoremas de isomorfismo . . . 148
7.1.5 Produto de anéis . . . 152
7.1.6 Domínios e corpos . . . 153
7.1.7 Ideais primos e ideais maximais . . . 154
7.1.8 Radicais . . . 156
7.1.9 Matrizes . . . 157
7.1.9.1 Soma de matrizes . . . 158
7.1.9.2 Produto de matrizes e produto por escalar . . . 159
7.1.9.3 Matrizes quadradas . . . 161
7.1.9.4 Traço . . . 162
7.2 Divisão em anéis . . . 163
7.2.1 Anel de frações . . . 163
7.2.2 Corpo de frações e números racionais $\mathbb{Q}$ . . . 169

7.2.2.1 Os números racionais $\mathbb{Q}$ . . . 170
7.2.2.2 Funções de arredondamento . . . 170
7.2.3 Números reais $\mathbb{R}$ . . . 172
7.2.3.1 Raiz de números reais . . . 175
7.2.4 Divisão e associação . . . 176
7.2.4.1 Divisão . . . 176
7.2.4.2 Relação de associação . . . 180
7.2.5 Domínios euclidianos . . . 181
7.2.6 Domínios principais . . . 183
7.2.7 Domínios de fatoração única . . . 183
7.2.7.1 Irredutíveis e primos . . . 183
7.2.7.2 Fatoração . . . 184
7.3 Anéis polinomiais . . . 190
7.3.1 Anel de polinômios . . . 190
7.3.1.1 Anel de polinômios multivariados . . . 194
7.3.2 Raízes de polinômios . . . 195
7.3.3 Anel de séries . . . 196
7.3.3.1 Topologia do anel de séries . . . 197
7.4 Corpos . . . 200
7.4.1 Corpo e subcorpo . . . 200
7.4.2 Extensões de corpos . . . 200
7.4.3 Extras . . . 203
7.5 Ação de anel . . . 205
7.5.1 Anel de endomorfismos de grupo . . . 205

8 Módulos **209**
8.1 Módulos e submódulos . . . 209
8.2 Homomorfismo de módulo . . . 210
8.2.1 Módulo de homomorfismos de módulo . . . 211

9 Espaços lineares **214**
9.1 Espaço e subespaço lineares . . . 214
9.2 Combinação linear de vetores . . . 220
9.3 Soma de subespaços vetoriais . . . 224
9.4 Bases de espaços vetoriais . . . 226
9.5 Funções lineares . . . 228
9.6 Produto e coproduto de espaços vetoriais . . . 231
9.6.1 Produto . . . 231
9.6.2 Coproduto (soma) . . . 233
9.7 Projeções lineares . . . 234
9.8 Funções multilineares . . . 235

9.8.1 Simetria, antissimetria e alternância 237
9.8.2 Espaços de funções multilineares 239
9.9 Formas multilineares 240
9.9.1 Produto tensorial de formas multilineares 240
9.9.2 Produto alternado de formas multilineares 241
9.9.3 Formas puxadas e determinante 243
9.9.4 Formas bilineares simétricas e formas quadráticas 244
9.9.4.1 2-Formas reais e os espaços $\mathbb{R}^{p,q}$ 246
9.10 Produto tensorial de espaços lineares 248
9.10.1 Tensores 251

10 Álgebras **252**
10.1 Álgebra e conceitos básicos 252
10.1.1 Ação adjunta 254
10.1.2 Colchete comutador 254
10.1.3 Derivação 255
10.2 Álgebra derivativa 257
10.3 Álgebra involutiva 259
10.4 Álgebras reais 262
10.4.1 Complexos $\mathbb{R}^2$ 263
10.4.1.1 Rotações em $\mathbb{R}^2$ 267
10.4.2 Quatérnios $\mathbb{R}^4$ 268
10.4.2.1 Multiplicação quaterniônica 268
10.4.2.2 Produtos interno, escalar e vetorial 278
10.4.2.3 Rotações em $\mathbb{R}^3$ 282
10.4.2.4 Rotações em $\mathbb{R}^4$ 285
10.4.3 Classificação das álgebras de divisão reais 286
10.4.4 Octônios $\mathbb{R}^8$ 289
10.4.4.1 Multiplicação octoniônica 289
10.4.5 2-Hiperbólicos $\mathbb{R}^{1,1}$ 292
10.4.6 4-Hiperbólicos $\mathbb{R}^{1,3}$ 294

11 Ordem em estruturas algébricas **296**
11.1 Corpos ordenados 296
11.1.1 Imersão dos inteiros e racionais 299
11.1.2 Cones positivos 300
11.1.3 Corpos ordenados completos 302
11.1.4 Corpos ordenados infinitesimais 303
11.1.5 Valor absoluto e distância 304
11.1.6 Conicidade 306
11.1.7 Convexidade 306

11.1.7.1 Simplexos 307

12 Espaços afins **310**
12.1 Espaço e subespaço afins 310
12.2 Transformação afim 312
12.3 Bases e coordenadas 314
12.3.1 Coordenadas baricêntricas 314

13 Álgebra universal **317**
13.1 Álgebra e estrutura algébrica 317
13.2 Morfismos 318
13.3 Subálgebra 318
13.4 Produto 318
13.5 Congruências e quociente 319
13.6 Núcleo e imagem 320

3 Geometria **322**

14 Topologia **323**
14.1 Espaços topológicos 323
14.1.1 Topologia, abertos e fechados 323
14.1.1.1 Interior e fecho 325
14.1.1.2 Fronteira 329
14.1.2 Topologias geradas, bases e sub-bases 329
14.1.2.1 Topologias finas e grossas e topologias geradas . . . 329
14.1.2.2 Bases e sub-bases 330
14.1.3 Funções contínuas 330
14.1.3.1 Suporte de funções vetoriais 331
14.1.4 Topologias induzidas 331
14.1.4.1 Topologias puxada e inicial 331
14.1.4.2 Topologias empurrada e final 333
14.1.4.3 Produto de espaços topológicos 334
14.1.4.4 Coproduto de espaços topológicos 335
14.1.4.5 Subespaços topológicos 336
14.1.4.6 Quociente de espaços topológicos 337
14.1.5 Topologia de ordem 339
14.1.5.1 Topologia de um corpo ordenado 339
14.2 Separação 341
14.2.1 Noções de separação de conjuntos 341
14.2.2 Espaços distinguíveis 342

14.2.3 Espaços acessíveis . . . 343
14.2.4 Espaços separados . . . 344
14.2.5 Espaços regulares . . . 344
14.2.6 Espaços completamente regulares . . . 345
14.2.7 Espaços normais . . . 345
14.3 Convergência . . . 347
14.3.1 Redes . . . 347
14.4 Conexidade . . . 348
14.4.1 Conexidades . . . 348
14.4.1.1 Componentes conexas . . . 351
14.4.1.2 Conexidade por caminhos . . . 352
14.5 Compacidade . . . 352
14.5.1 Compacidade . . . 352
14.5.1.1 Compacidade relativa . . . 354
14.5.2 Compacidade contável . . . 354
14.5.3 Compacidade regional e local . . . 354
14.5.4 Paracompacidade . . . 355
14.5.5 Compacidade sequencial . . . 355
14.6 Contabilidades . . . 355
14.6.1 Base de vizinhanças contável (1º contável) . . . 355
14.6.2 Base contável (2º contável) . . . 355
14.7 Espaços de funções . . . 355
14.7.1 Topologia pontual . . . 356
14.7.2 Topologia compacto-aberto . . . 356
14.7.3 Topologia compacto-cocompacto . . . 356
14.8 Topologia algébrica . . . 357
14.8.1 Homotopia . . . 357
14.8.2 Caminhos e laços . . . 358
14.8.3 Homotopia de caminhos . . . 359
14.8.4 Grupo fundamental . . . 363

15 Medida **364**
15.1 Espaço mensurável . . . 364
15.1.1 Sigma-álgebras e sub-sigma-álgebras . . . 364
15.1.2 Sigma-álgebras geradas . . . 365
15.1.3 Limites de conjuntos . . . 366
15.2 Funções mensuráveis . . . 367
15.2.1 Sigma-álgebras puxadas e empurradas . . . 367
15.3 Produto de espaços mensuráveis . . . 369
15.4 Espaços mensuráveis com estrutura adicional . . . 371
15.4.1 Espaços mensuráveis topológicos . . . 371

15.4.2 Funções mensuráveis com valores vetoriais 371
15.4.3 Funções mensuráveis com valores em espaços métricos 371
15.5 Medida e espaço de medida 373
15.5.1 Medidas 373
15.5.2 Medida exterior 376
15.6 Medida em espaços topológicos 379
15.6.1 Medidas regulares 379
15.6.2 Medidas localmente finitas 379
15.7 Espaço linear de medidas 380
15.8 Quase 381
15.8.1 Quase todo 381
15.8.2 Quase igualdade de conjuntos 382
15.8.3 Quase igualdade de funções 385
15.8.3.1 Abordagem alternativa 386
15.9 Integração 388
15.9.1 Integral de funções mensuráveis simples 388
15.9.2 Integral de funções mensuráveis positivas 393
15.9.3 Integral de funções mensuráveis 393
15.9.4 Teoremas de convergências 393
15.9.5 Mudança de variáveis na integração 393
15.9.6 Integral em espaços normados completos 394
15.9.7 Funções simples 395
15.9.8 Desintegração de medida 397

16 Espaços métricos **398**
16.1 Espaço métrico 398
16.1.1 Métricas 398
16.1.2 Diâmetro, bolas e limitação 402
16.2 Topologia dos espaços métricos 404
16.2.1 Interior e pontos interiores 404
16.2.2 Limites e convergência de sequências 405
16.2.3 Fecho e pontos aderentes 408
16.2.4 Conjuntos densos 409
16.2.5 Conjuntos compactos 410
16.2.6 Continuidade 410
16.2.7 Ponto limite e conjunto derivado 412
16.2.8 Distância e bolas de conjuntos e separação métrica 412
16.3 Estrutura uniforme 414
16.3.1 Sequências aproximantes 414
16.3.2 Continuidade uniforme 416
16.3.3 Espaços métricos completos 416

16.3.4 Limitação uniforme (ou total) 418
16.4 Funções métricas 420
16.4.1 Funções controladas 420
16.4.2 Homometrias e isometrias 423
16.4.3 Contrações 425
16.4.4 Semelhanças 427
16.4.5 Funções α-controladas 428
16.5 Medida em espaços métricos 429
16.5.1 Medidas exteriores métricas 429
16.5.2 Medidas por coberturas métricas 430
16.5.3 Dimensão métrica e fractais 434
16.5.4 Comprimento de caminhos 434

17 Grupos topológicos **438**
17.1 Grupo topológico 438
17.2 Topologia em grupos 442
17.3 Distância em grupos 443
17.4 Homomorfismos contínuos 444
17.5 Subgrupos topológicos 444
17.6 Conexidade em grupos topológicos 445
17.7 Grupo de homeomorfismos 445
17.8 Ação contínua 447
17.9 Medidas em grupos topológicos 447

18 Espaços lineares topológicos **449**
18.1 Anéis e corpos topológicos 449
18.2 Espaço linear topológico 449
18.3 Espaço de funções a valores vetoriais 449
18.4 Funções lineares contínuas 452
18.5 Espaço dual contínuo 452
18.6 Teoremas de representação 453
18.6.1 Representação linear de produto interno 453
18.6.2 Representação contínua de medida 454

19 Espaços normados **455**
19.1 Normas 455
19.1.1 Norma em corpos (valor absoluto) 455
19.1.2 Seminormas 458
19.1.3 Normas, espaços normados e métricas lineares 459
19.1.3.1 Funções limitadas e norma do supremo 462
19.1.4 Bolas e esferas unitárias e topologia 464

19.1.5 Equivalência de normas 465
19.2 Funções lineares contínuas 465
19.3 Espaços normados completos 468
19.3.1 Sequências absolutamente somáveis 470
19.3.2 Espaços normados de dimensão finita 471
19.3.3 Espaços de funções absolutamente somáveis 473
19.3.4 Espaços de funções absolutamente integráveis 475
19.3.4.1 Desigualdades das normas p 479
19.3.4.2 Os espaços de funções absolutamente integráveis são normados (e completos) 481
19.3.4.3 O caso anterior de espaços de funções absolutamente somáveis 484
19.3.5 Dualidade e mergulho de espaços absolutamente integráveis . 485
19.4 Isometrias lineares 486
19.4.1 Os grupos lineares geral e especial de transformações e de isometrias 486
19.4.2 Perturbação de uma transformação linear 487
19.5 Álgebras normadas 489
19.5.1 Função exponencial 490
19.5.1.1 Exponencial dos complexos 492
19.5.1.2 Exponencial dos quatérnios 493
19.6 Funções multilineares 494
19.6.1 Espaço de funções multilineares contínuas 496
19.7 Diferenciação em espaços normados 498
19.7.1 Diferenciabilidade 498
19.7.1.1 Diferencial e propriedades 498
19.7.1.2 Operações com funções diferenciáveis 501
19.7.1.3 Diferenciabilidade contínua 504
19.7.1.4 Diferenciabilidade forte 505
19.7.1.5 Diferenciabilidade da norma 507
19.7.2 Derivabilidade 507
19.7.2.1 Derivada de uma trajetória 507
19.7.2.2 Operaões entre trajetórias deriváveis 509
19.7.2.3 Diferenciabilidade e derivabilidade 511
19.7.3 Derivabilidade direcional 512
19.7.3.1 Fluxo constante 513
19.7.3.2 Derivada direcional 515
19.7.4 Extremos locais e pontos críticos 517
19.7.4.1 Campos escalares e pontos extremos locais 517
19.7.4.2 Pontos críticos 520

19.7.5 Valor médio 521
19.7.5.1 Valor médio para campos escalares 521
19.7.5.2 Majoração de trajetória 523
19.7.5.3 Desigualdade do valor médio e consequências . . . 525
19.7.5.4 Diferencial nula 527
19.7.6 Diferenciabilidade parcial 527
19.7.6.1 Contradomínio produto 528
19.7.6.2 Domínio produto 529
19.7.6.3 Diferenciabilidade das normas p 535
19.7.7 Multidiferenciabilidade 536
19.7.7.1 Multidiferencial e propriedades 536
19.7.7.2 Multidiferenciabilidade contínua 544
19.7.7.3 Multiderivabilidade 545
19.7.7.4 Multiderivabilidade contínua 546
19.7.8 Extremos locais e pontos críticos degenerados 547
19.7.8.1 Extremos locais e a segunda diferencial 547
19.7.8.2 Pontos críticos degenerados 548
19.7.9 Invertibilidade 549
19.7.9.1 Difeomorfismos 550
19.7.9.2 Teorema da Função Inversa 552
19.7.9.3 Teorema da Função Implícita 555
19.7.9.4 Teorema do Posto 558
19.7.9.5 Extremos locais com restrições 558
19.8 Integração em espaço normados 558
19.8.1 Integração de trajetórias 558
19.8.1.1 Integração em intervalos 558
19.8.1.2 Trajetórias primitivas 561
19.8.1.3 Integração orientada 562
19.8.2 Integração de campos vetoriais 564
19.8.2.1 Trajetórias integrais 564
19.8.2.2 Existência e unicidade de trajetórias integrais . . . 567
19.8.2.3 Fluxo de um campo vetorial 570
19.8.2.4 Propriedades das condições iniciais 574
19.8.2.5 Propriedades geométricas do fluxo 575
19.8.2.6 Continuidade e diferenciabilidade do fluxo 575
19.9 Campos tensoriais 575
19.9.1 Campos escalares 577
19.9.1.1 Álgebra dos campos escalares 577
19.9.2 Campos vetoriais 577
19.9.2.1 Álgebra dos campos vetoriais 577

19.9.2.2 Derivada ao longo de um campo vetorial 577
19.9.2.3 Colchete de campos vetoriais 579
19.9.3 Álgebra dos campos covetoriais alternados 583

20 Espaços lineares com produto interno 584
20.1 Produto interno 584
20.2 Norma induzida, ortogonalidade e ângulo 586
20.2.1 Norma 586
20.2.2 Perpendicularidade e paralelismo 593
20.2.3 Projeções paralela e perpendicular 594
20.2.3.1 Processo de ortogonalização 596
20.2.4 Projeções ortogonais 596
20.2.5 Ângulo 598
20.2.5.1 Definição de uma função ângulo 603
20.2.6 Espaço projetivo 624
20.2.6.1 Estrutura topológica 624
20.2.6.2 Estrutura métrica 625
20.2.7 Funções ortogonais e conformes 627
20.3 Espaço de funções quadrado somáveis 630

21 Variedades 631
21.1 Variedades topológicas e diferenciais 631
21.1.1 Cartas e atlas 631
21.1.2 Variedades e estrutura topológica e diferencial 634
21.1.3 Exemplos de variedades 636
21.1.3.1 Esfera 636
21.1.3.2 Espaço projetivo 638
21.1.3.3 Toro 639
21.1.3.4 Hiperboloide 640
21.1.4 Propriedades topológicas 640
21.2 Funções diferenciáveis 641
21.2.1 Funções diferenciáveis 641
21.2.2 Funções separadoras e partições da unidade 643
21.3 Espaço tangente 644
21.3.1 Espaço tangente em um ponto 644
21.3.2 Diferencial de uma função 647
21.3.3 Fibrado tangente 648
21.3.3.1 Referencial local coordenado 650
21.3.4 Álgebra dos campos escalares 651
21.3.5 Definições alternativas do espaço tangente 652
21.3.5.1 Derivações em pontos 652

21.3.5.2 Diferencial de uma função diferenciável 653
21.3.5.3 Vetores tangentes e espaços tangentes ao $\mathbb{R}^d$ 655
21.3.5.4 Curvas equivelozes 657
21.4 Subvariedades 659
21.4.1 Imersão 659
21.4.2 Submersão 659
21.4.3 Mergulho 659
21.4.4 Subvariedades imersas e mergulhadas 659
21.5 Transversalidade 659
21.5.1 Conjuntos nulos 659
21.5.2 Valor regular 661
21.5.3 Ponto crítico 661
21.5.4 Transversalidade 661
21.6 Campos tensoriais 663
21.6.1 Campos tensoriais, vetoriais, e derivações 663
21.6.2 Derivações e colchete de campos vetoriais 665
21.6.3 Álgebra de campos tensoriais 667
21.6.4 Fluxo de campos vetoriais 667
21.6.5 Espaço cotangente 672
21.6.5.1 Referencial local coordenado 672
21.6.6 Formas diferenciáveis 672
21.6.6.1 Representação coordenada 673
21.6.6.2 Formas volume 674
21.6.6.3 Formas puxadas 674
21.6.7 Derivada exterior 675
21.6.7.1 Formas no espaço real 675
21.6.7.2 Formas em variedades diferenciais 678
21.6.8 Derivada por fluxo 679
21.6.8.1 Derivada de campos escalares 680
21.6.8.2 Definição algébrica 680
21.7 Orientação 681
21.7.1 Orientação de espaços lineares 681
21.7.1.1 Aritmética de orientações 682
21.7.2 Orientação de variedades 683
21.7.2.1 Atlas orientados 683
21.8 Folheações 684
21.8.1 Cartas e atlas folheados 684
21.8.2 Componentes tangencial e transversal da carta 686
21.8.3 Folheação 687
21.8.3.1 Definições alternativas 687

22 Fibrados **689**
22.1 Fibrados topológicos 689
22.2 Fibrados vetoriais 690
22.2.1 Funções de transição 692
22.2.2 Seções locais e globais 702
22.2.3 Grupo estrutural 704
22.2.4 Referenciais locais e globais 704
22.3 Fibrados principais 705
22.4 Conexões em fibrados vetoriais 707
22.4.1 Espaços verticais e horizontais 707
22.4.2 Conexão/derivada covariante 708

23 Grupos diferenciais **711**
23.1 Definições básicas 711
23.1.1 Álgebra infinitesimal de um grupo diferencial 714
23.1.1.1 Campos vetoriais invariantes 714
23.1.1.2 Álgebra infinitesimal 715

4 Dinâmica **717**

24 Sistemas dinâmicos **718**
24.1 Definições gerais 718
24.1.1 Sistemas dinâmicos discretos e contínuos 718
24.1.2 Órbitas 719
24.1.3 Conjuntos invariantes 721
24.1.3.1 Invariância positiva e negativa 721
24.1.3.2 Subsistemas 722
24.1.3.3 Núcleo de recorrência 723
24.1.4 Periodicidade 723
24.2 Conjugação de sistemas dinâmicos 724
24.3 Translação em grupos 727
24.3.1 Translação no $\mathbb{T}^1$ 727
24.3.2 Translação no $\mathbb{T}^d$ 728
24.4 Expansão no $\mathbb{T}^1$ 728
24.5 Deslocamentos simbólicos 728
24.5.1 Deslocamento unilateral 728
24.5.2 Deslocamento bilateral 730
24.5.3 Subdeslocamentos 731
24.5.4 Itinerários 733
24.5.5 Automorfismos hiperbólicos no $\mathbb{T}^2$ 733

25 Sistemas dinâmicos lineares **738**
25.1 Sistemas dinâmicos lineares hiperbólicos 738
25.1.1 Cones 739

26 Sistemas dinâmicos topológicos **741**
26.1 Conjugação topológica 741
26.2 Transitividade topológica e minimalidade 741
26.3 Pontos errantes, conjuntos limites e pontos recorrentes 743
26.3.1 Pontos errantes 743
26.3.2 Pontos limite 744
26.3.3 Pontos recorrentes 745
26.4 Conjuntos estável e instável 746
26.5 Exemplos 747
26.5.1 Topologia dos deslocamentos 747
26.5.2 Itinerários topológicos 748
26.5.3 Automorfismos no toro 749

27 Sistemas dinâmicos métricos **753**
27.1 Conjuntos estáveis e instáveis 753
27.2 Conjuntos estáveis e instáveis locais 754
27.3 Pseudo órbitas 755
27.4 Distância no espaços de símbolos 755

28 Sistemas dinâmicos de medida **756**
28.1 Definições gerais 756
28.1.1 Quase invariância 756
28.2 Conjugação que preserva medida 757
28.3 Recorrência 758
28.4 Teoremas ergódicos 760
28.4.1 Tempo médio de visita 760
28.4.2 Média temporal e média espacial 760
28.4.3 Teorema ergódico da média 761
28.4.4 Teorema ergódico pontual 761
28.5 Ergodicidade 762
28.6 Sistemas misturadores 765
28.7 Entropia de medida 771
28.8 Medida invariante para contrações e expansões 772
28.9 Exemplos 774
28.9.1 Medida dos deslocamentos 774
28.9.2 Itinerários de medida 775
28.9.3 Rotação no círculo 776

29 Sistemas dinâmicos diferenciais **778**
29.1 Estrutura hiperbólica 778
29.2 Axioma A 779
29.3 Anosov 779
29.4 Partições de Markov 780

Bibliografia **781**

Figuras

10.1 Grafo para memorizar multiplicação quaterniônica. 269
10.2 Grafo para memorizar multiplicação octoniônica. 290

11.1 Simplexos unitários reais de dimensões 0, 1, 2 e 3 308

14.1 Ilustração de uma homotopia de caminhos. 359

16.1 Bolas aberta e fechada de centro c e raio r, respectivamente. 403
16.2 Gráfico de $H^d(C)$ em função de d. 434

19.1 Gráfico da função $^*\colon [1,\infty] \to [1,\infty], p \mapsto \frac{p}{p-1}$. 479

20.1 Gráfico da função $\cos^{-1}\colon [-1,1] \to \left[0, \frac{\tau}{2}\right]$. 599
20.2 Representação de vetores e o ângulo entre eles. 599
20.3 Soma e diferença de vetores de mesma norma 615

21.1 Esfera 2-dimensional $\mathbb{S}^2$ 638
21.2 Toro 639

Tabelas

5.1 Tabela de pedra-papel-tesoura (r, p e t). 81
5.2 Sequências de associação 84
5.3 Sequências de associação 85

10.1 Tabela de multiplicação dos quatérnios. 269
10.2 Tabela de produto vetorial em $\vec{\mathbb{R}}^3$. 278
10.3 Tabela de multiplicação dos octônios. 290
10.4 Tabela de produto vetorial em $\vec{\mathbb{R}}^7$. 291

Parte 1

Conjuntos

Capítulo 1

Os axiomas e as construções essenciais

1.1 Preliminares de lógica

Alguns conceitos da lógica formal serão brevemente introduzidos nesta seção para que a abordagem nas próximas seções façam sentido. Comentaremos sobre lógicas de ordem 0 e de ordem 1, tradicionalmente chamadas de lógica proposicional e lógica de predicados.

As teorias da lógica formal costumam ter axiomas — sentenças assumidas válidas a partir das quais devem-se inferir todas as outras sentenças da teoria. Ao longo do livro, os axiomas, as definições e as proposições serão enunciados, não como sentenças simbólicas, mas como sentenças em português, como se faz tradicionalmente na matemática. Isso facilita o entendimento e a naturalidade das sentenças. No entanto, ao longo das seções que seguem essas sentenças serão enunciadas também formalmente, para estimular o pensamento e a manipulação formais.

Alguns símbolos lógicos frequentemente facilitam e deixam mais claros os enunciados de sentenças na matemática. O símbolo

$$\forall$$

será usados para substituir as expressões "para todo/a", "para todos/as", "para qualquer", "para quaisquer" e outras possíveis flexões gramaticais. O símbolo

$$\exists$$

será usados para substituir as expressões "para algum/ma", "para alguns/mas", "existe", "existem" e outras possíveis flexões gramaticais. Eles indicam que uma

propriedade vale para todo/algum conjunto. O símbolo $\exists!$ significa que existe e é único. Além desses, serão usados também os símbolos lógicos

$$\leftarrow \qquad \rightarrow \qquad \leftrightarrow$$

e as versões informais matemáticas

$$\Longrightarrow \qquad \Longleftarrow \qquad \Longleftrightarrow$$

para significar a implicação material em cada sentido e a equivalência lógica. Por fim, para os conectivos 'e' e 'ou' são usados os símbolos lógicos

$$\wedge \qquad \vee$$

e as versões informais matemáticas

$$\text{e} \qquad \text{ou} \ .$$

Esse conectivos indicam, informalmente, que sentenças são ambas verdadeiras, no caso de 'e', ou ao menos uma das duas é, no caso de 'ou'. Os parênteses, que são comumente usados na lógica formal, serão substituídos por espaços, de modo que não haja ambiguidade. Mais detalhes sobre lógica formal e o uso dos símbolos lógicos serão suprimidos. Para aprofundamento em lógica e sistemas dedutivos, um livro indicado é **liv:Tarski-IntroductionLogic**, de **liv:Tarski-IntroductionLogic**.

1.1.1 Lógica de ordem 0

Na lógica de ordem 0, existem contáveis *proposições simples* (ou *variáveis proposicionais*). Elas são representadas por símbolos especificados e no que segue serão representadas pela letra P ou por P seguida de uma quantidade finita de apóstrofes '′', como P' ou P''''.

Além disso, existem os *operadores lógicos*, que são os símbolos

$$\neg \qquad \wedge \qquad \vee \qquad \rightarrow \qquad \leftrightarrow$$

respectivamente chamados de *negação*, *conjunção*, *disjunção*, *implicação* e *equivalência*. Podem-se considerar também os símbolos

$$\top \qquad \bot$$

que são chamados de *verdade* e *falsidade*, mas não adotaremos eles neste livro. Os símbolos $\top$ e $\bot$ são operadores 0-ários, o símbolo $\neg$ é um operador 1-ário e os operadores $\wedge$, $\vee$, $\rightarrow$ e $\leftrightarrow$ são operadores 2-ários.

Os conectivos lógicos podem ser combinados com as proposições atômicas para formar outras proposições. Para isso, usam-se os símbolos

$$(\qquad)$$

que são os *parênteses inicial* e *final*, respectivamente. As regras de formação aceitas são que, se P e P' são proposições (atômicas ou não), então também são proposições

$$\neg P \qquad (P \wedge P') \qquad (P \vee P') \qquad (P \to P') \qquad (P \leftrightarrow P')$$

e os parênteses podem ser omitidos quando for claro o que se pretende dizer. Qualquer fórmula formada indutivamente por esses passos é uma proposição da linguagem.

Além disso, existem as regras de inferência ou dedução sintática, as regras que nos permitem, a partir de um conjunto de proposições Γ, deduzir outra proposição P. Denota-se isso por

$$\Gamma \vdash P$$

Adotaremos as seguintes regras. Para quaisquer proposições P, P' e P'', valem as seguintes deduções.

Introdução da Negação:

$$(P \to P'), (P \to \neg P') \vdash (\neg P)$$

Eliminação da Negação:

$$(\neg P) \vdash (P \to P')$$

Eliminação da Dupla-Negação:

$$(\neg\neg P) \vdash P$$

Introdução da Conjunção:

$$P, P' \vdash (P \wedge P')$$

Eliminações da Conjunção:

$$(P \wedge P') \vdash P$$

$$(P \wedge P') \vdash P'$$

Introduções da Disjunção:

$$P \vdash (P \vee P')$$

$$P' \vdash (P \vee P')$$

Eliminação da Disjunção:

$$(P \vee P'), (P \to P''), (P' \to P'') \vdash P''$$

Introdução da Equivalência:

$$(P \to P'), (P' \to P) \vdash (P \leftrightarrow P')$$

Eliminações da Equivalência:

$$(P \leftrightarrow P') \vdash (P \to P')$$

$$(P \leftrightarrow P') \vdash (P' \to P)$$

Introdução da Implicação:

$$(P \vdash P') \vdash (P \to P')$$

Eliminação da Implicação:

$$P, (P \to P') \vdash P'$$

A partir de um conjunto de proposições Γ, *deduzimos* uma proposição P com uma sequência finita de proposições obtidas de Γ usando as regras de inferência enunciadas.

1.1.2 Lógica de ordem 1

A lógica de ordem 1 é uma linguagem formal. O *alfabeto* é composto por

1. letras latinas (em diferentes fontes e tamanhos)

$$a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z$$

$$A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z$$

2. letras gregas

$$\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, \xi, o, \pi, \rho, \sigma, \tau, \upsilon, \phi, \chi, \psi, \omega$$

$$\Gamma, \Delta, \Theta, \Lambda, \Xi, \Pi, \Sigma, \Upsilon, \Phi, \Psi, \Omega$$

3. numerais arábicos

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9$$

4. símbolos lógicos (ou *operadores*)

$$\neg, \wedge, \vee, \to, \leftrightarrow, \forall, \exists$$

5. outros símbolos gráficos
$$(,),[,],\{,\},|,',,$$

A *expressões básicas* podem ser

1. as *variáveis*, (em uma quantidade contável), representadas aqui por letras dos alfabetos latino e grego, possivelmente indexadas por outras letras ou numerais, ou pelo apóstrofe ';
2. as *constantes (individuais)* (em uma quantidade contável), representadas aqui por letras dos alfabetos latino e grego, possivelmente indexadas por outras letras ou numerais, ou pelo apóstrofe ', ou representadas por símbolos específicos;
3. para cada número natural n, os *predicados* n-ários (em uma quantidade contável), representadas aqui por letras dos alfabetos latino e grego, possivelmente indexadas por outras letras ou numerais, ou pelo apóstrofe ', ou representadas por símbolos específicos. Os predicados 0-ários são as *proposições* (ou *letras sentenciais*), os predicados 1-ários são as *propriedades* e os predicados 2-ários são as *relações*.

Não restringiremos quais letras representam o que, isso sempre ficará claro pelo contexto. De fato, a quantidade de símbolos de uma linguagem é finita, mas expressões formadas por eles são contáveis, mas não especificaremos esses detalhes aqui.

As *expressões bem formadas* da linguagem são

1. Os *termos*, que são as variáveis ou constantes individuais;
2. As *fórmulas*, expressões definidas recursivamente por
   2.1. (Fórmulas simples ou atômicas) Para um predicado n-ário P e termos $t_1, \cdots, t_n$, $P(t_1, \cdots, t_n)$ é uma fórmula;
   2.2. (Fórmulas moleculares) Para fórmulas F e F', $\neg F$, $(F \wedge F')$, $(F \vee F')$, $(F \to F')$ e $(F \leftrightarrow F')$ são fórmulas. As fórmulas F e F' são *subfórmulas* dessas fórmulas;
   2.3. (Fórmulas gerais) Para fórmula F e variável v que ocorre em F, $(\forall v F)$ e $(\exists v F)$ são fórmulas. A fórmula F é uma *subfórmula* dessas fórmulas;
   2.4. Nenhuma outra expressão é uma fórmula.

Uma *variável livre* de uma fórmula é definida recursivamente como

1. Para uma fórmula simples F e v uma variável, v é livre em F se, e somente se, v ocorre em F;
2. Para uma fórmula F da forma $\neg F'$, $(F' \wedge F'')$, $(F' \vee F'')$, $(F' \to F'')$, $(F' \leftrightarrow F'')$, e v uma variável, v é livre em F se, e somente se, é livre em F' ou em F'';

3. Para uma fórmula composta F da forma $\forall v' F'$ ou $\exists v' F'$ e v uma variável, v é livre em F se, e somente se, v é livre em F' e v é diferente de v'. No caso em que v é livre em F' mas v é igual a v', a variável v está *ligada* ao quantificador $\forall$ ou $\exists$, respectivamente.

Uma *sentença* (ou *fórmula fechada*) é uma fórmula sem variáveis livres. Uma *fórmula aberta* é uma fórmula que não é uma sentença.

1.1.2.1 Abreviações

Por simplicidade, os parênteses de fórmulas moleculares e gerais podem ser omitidos sempre que possível. Fórmulas gerais da forma $\forall v F$ podem ser abreviadas por

$$\forall_v F$$

e fórmulas gerais da forma $\exists v F$ podem ser abreviadas por

$$\exists_v F.$$

Ainda, para fórmulas F e F' e variável v livre em alguma das fórmulas, proposições categóricas do tipo

$$\forall v (F \to F')$$

podem ser abreviadas por

$$\forall_{v|F} F'$$

e proposições categóricas do tipo

$$\exists v (F \wedge F')$$

podem ser abreviadas por

$$\exists_{v|F} F'$$

em que o símbolo | lê-se 'tal que'.

Esse é o modo em que pensamos nessas afirmações na linguagem natural e em matemática.

1.1.2.2 Regras de inferência

As seguintes regras de inferência são definidas para fórmulas F e F':

1. *Introdução da Negação*:

$$(F \to F'), (F \to \neg F') \vdash (\neg F)$$

2. *Eliminação da Negação*:
$$(\neg F) \vdash (F \to F')$$
3. *Eliminação da Dupla-Negação*:
$$(\neg\neg F) \vdash F$$
4. *Introdução da Conjunção*:
$$F, F' \vdash (F \wedge F')$$
5. *Eliminações da Conjunção*:
$$(F \wedge F') \vdash F$$
$$(F \wedge F') \vdash F'$$
6. *Introduções da Disjunção*:
$$F \vdash (F \vee F')$$
$$F' \vdash (F \vee F')$$
7. *Eliminação da Disjunção*:
$$(F \vee F'), (F \to F''), (F' \to F'') \vdash F''$$
8. *Introdução da Equivalência*:
$$(F \to F'), (F' \to F) \vdash (F \leftrightarrow F')$$
9. *Eliminações da Equivalência*:
$$(F \leftrightarrow F') \vdash (F \to F')$$
$$(F \leftrightarrow F') \vdash (F' \to F)$$
10. *Introdução da Implicação*:
$$(F \vdash F') \vdash (F \to F')$$
11. *Eliminação da Implicação*:
$$F, (F \to F') \vdash F'$$

Para uma fórmula F, uma variável v e um termo t tal que, ao substituirmos cada ocorrência livre de v em F por t, o termo t não é ligado a nenhum quantificador de F, denotamos a *substituição* de v por t na fórmula F por
$$F{\downarrow}^{v}_{t}\,,$$
comumente denotado $F[v \nearrow t]$.

1. *Introdução do Quantificador Universal*: Para fórmula F, variável v e termo t, t não ocorrente em premissa ou hipótese vigente,
$$F\downarrow_t^v \vdash \forall v F$$
2. *Eliminação do Quantificador Universal*: Para fórmula F, variável v e termo t,
$$\forall v F \vdash F\downarrow_t^v$$
3. *Introdução do Quantificador Particular*: Para fórmula F, variável v e termo t,
$$F\downarrow_t^v \vdash \exists v F$$
4. *Eliminação do Quantificador Particular*: Para fórmulas F e F', variável v e termo t, t não ocorrente em premissa ou hipótese vigente e t não ocorre em F',
$$(\exists v F), (F\downarrow_t^v \vdash F') \vdash F'$$

Os detalhes sobre essas regras de inferência não serão comentados, mas elas simplesmente formalizam as regras de inferência usuais na prática de matemática.

1.1.2.3 Igualdade e unicidade

O predicado 2-ário mais importante em linguagens de lógica de ordem 1 é a *igualdade* $=$. Os detalhes sobre a definição e uso da igualdade não serão especificados aqui, mas basicamente para termos t e t', $t = t'$ é uma fórmula, e sempre que vale $t = t'$, pode-se substituir todos os t por t' em uma fórmula que ainda se tem uma fórmula equivalente. O símbolo *desigualdade* é uma abreviação para uma fórmula de negação de uma igualdade. Para termos t e t', a expressão
$$t \neq t'$$
abrevia a fórmula
$$\neg(t = t').$$

Com o conceito de igualdade, podemos definir o que significa que uma *única* constante c satisfaz uma propriedade P. Para uma propriedade P e uma variável v, a expressão
$$\exists! v P v$$
significa
$$\exists v(Pv \wedge \forall v'(Pv' \to v' = v)).$$

Algumas definições equivalentes são
$$\exists v \forall v'(Pv' \leftrightarrow v' = v),$$

que é mais breve, mas na prática mais difícil de se usar em demonstrações, e

$$\exists v(Pv \wedge \neg\exists v'(Pv' \wedge v' \neq v)),$$

e

$$\exists v Pv \wedge \forall v \forall v'((Pv \wedge Pv') \rightarrow v = v').$$

1.1.3 Conjunto e pertencimento

A noção de um *conjunto* é uma noção primitiva na matemática. Intuitivamente, um conjunto é um objeto que tem *elementos*. Cada elemento tem para com o conjunto em que está a relação de *pertencimento*. Abstraindo mais essa noção, pensamos que todas as propriedades de um conjunto se resumem aos elementos que a ele pertencem, de modo que um conjunto é, de fato, seus elementos. A *Teoria de Conjuntos* é uma teoria da lógica formal que procura formalizar essas ideias e estudar suas consequências. Neste livro, o tratamento da teoria de conjuntos será um tratamento parcialmente formal, parcialmente informal, embora muita ênfase seja dada nos axiomas que constituem uma base para a teoria de conjuntos.

A lógica formal estuda sentenças formadas a partir de símbolos pré-determinados e fixos e as regras que dizem como essas sentenças se relacionam para formar novas sentenças. No tratamento formal da teoria de conjuntos, não há distinção entre conjunto e elemento. Ambos são somente denotados por letras de um alfabeto específico, e a relação de pertencimento é geralmente denotada pelo o símbolo $\in$. Se C e C' são conjuntos, a sentença "o conjunto C pertence ao conjunto C'" ou "o conjunto C é elemento do conjunto C'" é denotada por

$$C \in C'.$$

Para afirmar que um conjunto C não é elemento de um conjunto C', ou seja, negar $C \in C'$, o símbolo usado é $\notin$ e se denota $C \notin C'$.

Isso quer dizer que a teoria de conjuntos é uma teoria da lógica de ordem 1 com igualdade $=$ e um predicado 2-ário $\in$.

1.1.3.1 Abreviações

Comentamos anteriormente que, para fórmulas F e F' e variável v livre em alguma das fórmulas, proposições categóricas do tipo $\forall v(F \rightarrow F')$ podem ser abreviadas por $\forall_{v|F} F'$ e proposições categóricas do tipo $\exists v(F \wedge F')$ podem ser abreviadas por $\exists_{v|F} F'$. No caso em que F é da forma $v \in C$, as fórmulas podem ser ainda mais abreviadas respectivamente por

$$\forall_{v \in C} F'$$

e

$$\exists_{v \in C} F'.$$

1.2 Axiomas do vazio e da extensão

1.2.1 Vazio e igualdade

Os conceitos definidos nesta seção são *igualdade* e *contenção* de conjuntos. O primeiro axioma a ser considerado é o que define que existe um conjunto sem nenhum elemento, o *conjunto vazio*. Esse conjunto tem um papel semelhante ao número zero. Ele é, de certo modo, um "objeto neutro" na teoria de conjuntos. Ao decorrer do desenvolvimento da teoria, essa frase sem significado matemática de fato ganhará um significado intuitivo e, em vários casos, uma definição mais precisa.

:⊢ **Definição 1.1.** Um conjunto *vazio* é um conjunto que não possui elementos.

Axioma 0 (Vazio)**.** Algum conjunto é vazio.

$$\exists C \forall x (x \notin C)$$

O axioma não estabelece explicitamente a unicidade desse conjunto, mas ele é de fato único e é denotado $\emptyset$. Como o conjunto vazio não possui elementos, sempre que se conclui que existe um elemento em $\emptyset$, ou seja, que existe $x \in \emptyset$, chega-se em uma contradição e a conclusão é que o que se assumiu para chegar na contradição é falso. Essa é uma forma padrão de se demonstrarem diversas proposições na lógica e na matemática.

O segundo axioma considerado é um axioma baseado em uma dos primeiras propriedades de um conjunto quando pensado intuitivamente: a ideia de que, quando abstrai-se da realidade, um conjunto é totalmente definido pelos elementos que a ele pertencem. Esse axioma se chama axioma da extensão e é, de certa forma, a definição de *igualdade* entre conjuntos.

Axioma 1 (Extensão)**.** Sejam C e C' conjuntos. Os conjuntos C e C' são *iguais* se, e somente se, todo elemento de C pertence a C' e todo elemento de C' pertence a C. Denota-se $C = C'$.

$$\forall C \forall C' (\forall x (x \in C \leftrightarrow x \in C') \rightarrow C = C')$$

Caso contrário, denota-se $C \neq C'$.

$$\forall C \forall C' (C \neq C' \leftrightarrow \neg (C = C'))$$

A recíproca do axioma da extensão segue das propriedades da igualdade. O axioma da extensão nos permite provar a proposição afirmada anteriormente de que um único conjunto é vazio.

$\vdash$ **Proposição 1.1.** *Um único conjunto é vazio.*

$$\exists! C \forall x (x \notin C)$$

□ *Demonstração.* Sejam C e C' conjuntos vazios. Se $C \neq C'$, então (pelo axioma da extensão) para algum conjunto Y, $Y \in C$ e $Y \notin C'$, ou $Y \in C'$ e $Y \notin C$. O primeiro caso leva à contradição $Y \in C$ e o segundo caso leva à contradição $Y \in C'$, pois C e C' são vazios, não possuem elementos. Logo $C = C'$. ■

$:\vdash$ **Definição 1.2.** O único conjunto vazio é denotado $\emptyset$.

$$\forall C (C = \emptyset \leftrightarrow \forall x (x \notin C))$$

1.2.2 Subconjuntos

Quando se consideram conjuntos, é muito útil falar apenas de alguns de seus elementos, um conjunto desses elementos, possivelmente com alguma propriedade específica. Essa noção é a de um subconjunto, um conjunto cujos elementos pertencem todos a um outro conjunto considerado anteriormente. A definição de um subconjunto pode ser dada simplesmente a partir das noções primitivas já fornecidas, pois na ideia de subconjunto só são necessárias as noções de conjunto e pertencimento, além dos símbolos lógicos.

$:\vdash$ **Definição 1.3.** Seja C um conjunto. Um *subconjunto* (ou uma *parte*) de C é um conjunto C' tal que, para todo $x \in C'$, $x \in C$. Denota-se $C' \subseteq C$.

$$\forall C \forall C' (C' \subseteq C \leftrightarrow \forall x (x \in C' \rightarrow x \in C))$$

Caso contrário, denota-se $Y \nsubseteq X$.

$$\forall C \forall C' (C' \nsubseteq C \leftrightarrow \neg C' \subseteq C)$$

Um subconjunto *próprio* de C é um subconjunto $C' \subseteq C$ tal que $C' \neq C$. Denota-se $C' \subset C$.

$$\forall C \forall C' (C' \subset C \leftrightarrow C' \subseteq C \wedge C' \neq C)$$

Com essa definição, o axioma da extensão pode ser re-enunciado como

$$\forall C \forall C' (C = C' \leftrightarrow C \subseteq C' \wedge C' \subseteq C)$$

$\vdash$ **Proposição 1.2.** *Para todo conjunto C,*

1. *O conjunto vazio é subconjunto de todo conjunto.*

$$\forall C (\emptyset \subseteq C)$$

2. *O único subconjunto do conjunto vazio é ele mesmo.*

$$\forall C(C \subseteq \emptyset \to C = \emptyset)$$

□ *Demonstração.* 1. Suponha que $\emptyset$ não é subconjunto de C. Então existe $x \in \emptyset$ tal que $x \notin C$. Mas $x \in \emptyset$ é uma contradição, o que mostra que $\emptyset \subseteq C$.
2. Exercício.

■

1.3 Axioma da especificação

A noção intuitiva de subconjunto está diretamente relacionada à ideia de formar, a partir de um conjunto e uma propriedade, o subconjunto dos elementos que têm essa propriedade. A existência desse subconjunto é um axioma, chamado axioma da especificação porque a propriedade dada é um especificação dos elementos do conjunto original.

Axioma 2 (Especificação (Esquema))**.** Seja f uma fórmula com variáveis livres $x, C, v_0, \cdots, v_n$. Para todos $v_0, \cdots, v_n$ e todo conjunto C, algum conjunto S tem como seus únicos elementos os elementos de C tais que vale $f(x, C, v_0, \cdots, v_n)$.

$$\forall v_0, \cdots, v_n \forall C \exists S \forall x (x \in S \leftrightarrow (x \in C \wedge f(x, C, v_0, \cdots, v_n)))$$

⊢ **Proposição 1.3** (Unicidade da Especificação)**.** *Seja f uma fórmula com variáveis livres $x, C, v_0, \cdots, v_n$. Para todos $v_0, \cdots, v_n$ e todo conjunto C, um único conjunto S tem como seus únicos elementos os elementos de C tais que vale $f(x, C, v_0, \cdots, v_n)$.*

$$\forall v_0, \cdots, v_n \forall C \exists! S \forall x (x \in S \leftrightarrow (x \in C \wedge f(x, C, v_0, \cdots, v_n)))$$

□ *Demonstração.* Sejam f, $x, C, v_0, \cdots, v_n$ como no enunciado. Pelo axioma da especificação, algum conjunto S tem como seus únicos elementos os elementos de C tais que vale $f(x, C, v_0, \cdots, v_n)$. Sejam S, S' conjuntos com essas propriedades. Mostraremos que $S = S'$. Seja x um conjunto. Então $x \in S$ é equivalente a $x \in C$ e $f(x, C, v_0, \cdots, v_n)$, o que é equivalente $x \in S'$. Pelo axioma da extensão, $S = S'$. ■

:⊢ **Definição 1.4.** Seja f uma fórmula com variáveis livres $x, C, v_0, \cdots, v_n$. Para todos $v_0, \cdots, v_n$ e todo conjunto C, o único conjunto S que tem como seus únicos elementos os elementos de C tais que vale $f(x, C, v_0, \cdots, v_n)$ é denotado

$$\{x \in C \mid f(x, C, v_0, \cdots, v_n)\} := S.$$

⊢ **Proposição 1.4.** *Seja f uma fórmula com variáveis livres $x, C, v_0, \cdots, v_n$.*

$$\{x \in C \mid f(x, C, v_0, \cdots, v_n)\} \subseteq C.$$

□ *Demonstração.* Se $x \in S$, então $x \in C$ e $f(x, C, v_0, \cdots, v_n)$, logo $x \in C$, o que mostra que $S \subseteq C$. ■

1.4 Axioma das partes

O próximo axioma considerado é o que garante que os subconjuntos de um conjunto dado formam um conjunto.

Axioma 3 (Partes)**.** Para todo conjunto C, algum conjunto tem todos os subconjuntos de C como elementos.

$$\forall C \exists P \forall p (p \subseteq C \to p \in P)$$

O conjunto cujos elementos precisamente pelos subconjuntos de um conjunto X é chamado de conjunto das partes de X e denotado $\wp(X)$. O axioma das partes não estabelece explicitamente sua existência, mas podemos mostrar que ele existe usando o axioma da especificação, e mais precisamente mostrar que ele é único pela unicidade da especificação.

$\vdash$ **Proposição 1.5** (Conjuntos das Partes)**.** *Para todo conjunto C, um único conjunto tem exatamente os subconjuntos de C como elementos.*

$$\forall C \exists! P \forall p (p \subseteq C \leftrightarrow p \in P)$$

$\square$ *Demonstração.* Seja C um conjunto. Pelo axioma das partes, algum conjunto Q satisfaz que, para todo $p \subseteq C$, $p \in Q$. Consideremos a fórmula $p \subseteq C$. Pela unicidade da especificação, existe único P definido como

$$P := \{p \in Q \mid p \subseteq C\}. \qquad \blacksquare$$

O conjunto P não depende de Q pela unicidade da especificação.

$:\vdash$ **Definição 1.5.** Para todo conjunto C, o *conjunto das partes* de C é o único conjunto cujos elementos são os subconjuntos de C. Denota-se

$$\wp(C) = \{p \mid p \subseteq C\}.$$

$\vdash$ **Proposição 1.6.** *O conjunto das partes de um subconjunto de um conjunto é subconjunto do conjunto das partes do conjunto.*

$$\forall C \forall C' (C \subseteq C' \to \wp(C) \subseteq \wp(C'))$$

1.5 Axioma do par

O próximo axioma garante, a partir da existência de dois conjuntos, a existência de um novo conjunto cujos elementos são os dois conjuntos iniciais. Esse é o axioma do par. Embora a princípio sua necessidade não seja óbvia, esse axioma é importante — ao menos útil — para o desenvolvimento da teoria de conjuntos.

Axioma 4 (Par)**.** Para todos conjuntos C e C', algum conjunto P tem como elementos C e C'.

$$\forall C \forall C' \exists P (C \in P \wedge C' \in P)$$

Como o axioma do conjunto vazio, o axioma da especificação e o axioma das partes, esse conjunto P não é explicitamente único pelo axioma do par. Devemos derivar a unicidade de um conjunto que tem exatamente os dois conjuntos como elementos a partir da unicidade da especificação.

$\vdash$ **Proposição 1.7.** *Para todos conjuntos C e C', um único conjunto P tem como seus únicos elementos C e C'.*

$$\forall C \forall C' \exists! P \forall x (x \in P \leftrightarrow x = C \vee x = C')$$

$\square$ *Demonstração.* Pelo axioma do par, algum conjunto Q tem como elementos os conjunto C e C'. pela unicidade da especificação, definimos

$$P := \{x \in Q \mid x = C \vee x = C'\}. \qquad \blacksquare$$

$:\vdash$ **Definição 1.6.** Para todos conjuntos C e C', o *par* de C e C' é o único conjunto que tem como seus únicos elementos C e C'. Denota-se

$$\{C, C'\} = \{x \mid x = C \vee x = C'\}.$$

A partir do axioma do par pode-se formar o conjunto que tem como único elemento um conjunto X formando o par de X e X. Esse conjunto é o conjunto unitário com único elemento X.

$:\vdash$ **Definição 1.7.** Seja C um conjunto. O *conjunto unitário* de elemento C é o par de C e C. Denota-se

$$\{C\} := \{C, C\}.$$

1.6 Axioma da união

1.6.1 União de um conjunto

Nesta seção são apresentadas duas das construções mais importantes da teoria de conjuntos: a união e a interseção. A união de um conjunto de conjuntos denotado C é o conjunto cujos elementos pertencem a algum conjunto que pertence C. O axioma da união afirma que esse conjunto existe.

Axioma 5 (União)**.** Para todo conjunto C, algum conjunto tem como elementos os elementos que pertencem a algum elemento de C.

$$\forall C \exists U \forall x (\exists X (X \in C \wedge x \in X) \rightarrow x \in U)$$

Novamente, a unicidade de um conjunto que tem exatamente esses elementos segue do unicidade da especificação.

$\vdash$ **Proposição 1.8.** *Para todo conjunto C, um único conjunto tem como seus únicos elementos os elementos que pertencem a algum elemento de C.*

$$\forall C \exists! U \forall x (\exists X (X \in C \wedge x \in X) \leftrightarrow x \in U)$$

$\square$ *Demonstração.* Pelo axioma da união, existe V cujos elementos são os elementos que pertencem a algum elemento de C. Pela unicidade da especificação, definimos

$$U := \{x \in V \mid \exists X (X \in C \wedge x \in X)\}. \qquad \blacksquare$$

$:\vdash$ **Definição 1.8.** Para todo conjunto C, a *união* de C é o único conjunto que tem como seus únicos elementos os elementos que pertencem a algum elemento de C. Denota-se

$$\bigcup C = \{x \mid \exists X (X \in C \wedge x \in X)\}.$$

A união do par $\{C, C'\}$ e denotada $C \cup C'$.

$\vdash$ **Proposição 1.9.** *1. A união do conjunto vazio é o conjunto vazio.*

$$\bigcup \emptyset = \emptyset$$

2. A união de um subconjunto de um conjunto é subconjunto da união do conjunto.

$$\forall C \forall C' (C \subseteq C' \rightarrow \bigcup C \subseteq \bigcup C')$$

3. Todo elemento de um conjunto é subconjunto da união do conjunto.

$$\forall C \forall X (X \in C \rightarrow X \subseteq \bigcup C)$$

$\square$ *Demonstração.*

1. Suponha que $x \in \bigcup \emptyset$. Então $\exists X \in \emptyset$ tal que $x \in X$, o que é absurdo porque não pode existir $X \in \emptyset$.
2. Seja $x \in \bigcup C$. Então existe $X \in C$ tal que $x \in X$. Como $C \subseteq C'$, segue que $X \in C'$, portanto $x \in \bigcup C'$. $\blacksquare$

1.6.2 Interseção de um conjunto

A interseção de um conjunto não vazio de conjuntos denotado C é o conjunto cujos elementos pertencem a todos conjuntos que pertencem C. O conjunto interseção existe por consequência do axioma da especificação.

$\vdash$ **Proposição 1.10.** *Para todo conjunto não vazio C, um único conjunto tem como elementos exatamente os elementos que pertencem a todos elementos de C.*

$$\forall C \exists! I \forall x (\forall X (X \in C \rightarrow x \in X) \leftrightarrow x \in I)$$

□ *Demonstração.* Como C é não vazio, tome $X' \in C$. Pela unicidade da especificação, definimos

$$I := \{x \in X' \mid \forall X(X \in C \to x \in X)\}\,. \qquad \blacksquare$$

:⊢ **Definição 1.9.** Para todo conjunto não vazio C, a *interseção* de C é o único conjunto que tem como elementos exatamente os elementos que pertencem a todos elementos de C. Denota-se

$$\bigcap C = \{x \mid \forall X(X \in C \to x \in X)\}\,.$$

A interseção do par $\{C, C'\}$ e denotada $C \cap C'$.

⊢ **Proposição 1.11.** *A interseção de um conjunto não vazio é subconjunto da interseção de um subconjunto não vazio do conjunto.*

$$\forall C \forall C'(C \neq \emptyset \wedge C' \neq \emptyset \wedge C \subseteq C' \to \bigcap C' \subseteq \bigcap C)$$

□ *Demonstração.* Seja $x \in \bigcap C'$. Então, para todo $X \in C'$, $x \in X$. Como $C \subseteq C'$, então todo para $X \in C$, $X \in C'$, logo $x \in \bigcap C$. ■

⊢ **Proposição 1.12.** *A interseção de um conjunto não vazio é subconjunto de todo elemento do conjunto.*

$$\forall C \forall X(C \neq \emptyset \wedge X \in C \to \bigcap C \subseteq X)$$

1.7 Axioma do infinito

:⊢ **Definição 1.10.** Seja C um conjunto. O *sucessor* de C é o conjunto

$$\vdash\!\!\!\vdash(C) := C \cup \{C\}.$$

Um conjunto *indutivo* é um conjunto que contém $\emptyset$ e contém o sucessor de cada um de seus elementos.

Axioma 6. Algum conjunto é indutivo.

$$\exists I(\emptyset \in I \wedge \forall x \in I(\vdash\!\!\!\vdash(x) \in I))$$

⊢ **Proposição 1.13.** *Seja F um conjunto cujos elementos são conjuntos indutivos. $\bigcap F$ é indutivo.*

□ *Demonstração.* Como todo $C \in F$ é indutivo, para todo $C \in F$ vale $\emptyset \in C$, portanto $\emptyset \in \bigcap F$. Seja $x \in \bigcap F$. Então, para todo $C \in F$, $x \in C$ e, como C é indutivo, $\vdash\!\!\!\vdash(x) \in C$, logo $\vdash\!\!\!\vdash(x) \in \bigcap F$. Isso mostra que $\bigcap F$ é indutivo. ■

$\vdash$ **Proposição 1.14.** *Um único conjunto indutivo é subconjunto de todo conjunto indutivo.*

$\square$ *Demonstração.* Pelo axioma do infinito, algum conjunto I é indutivo. Consideremos o conjunto F dos subconjuntos indutivos de I:

$$F := \{C \in \wp(I) \mid \emptyset \in C \wedge \forall x \in C(\mathrm{H}(x) \in C)\}.$$

Definimos

$$\omega := \bigcap F.$$

O conjunto ω é indutivo pela proposição anterior. Mostremos que ω é subconjunto de todo conjunto indutivo. Seja J um conjunto indutivo. Consideremos o conjunto $\bigcap\{I, J\} = I \cap J$. Como I e J são indutivos, $I \cap J$ é indutivo pela proposição anterior; como $I \cap J \subseteq I$, $I \cap J \in F$, logo $\omega \subseteq I \cap J \subseteq J$. ■

$:\vdash$ **Definição 1.11.** O *conjunto dos números ordinais finitos* (ou *números naturais*) é o único conjunto indutivo que é subconjunto de todo conjunto indutivo. Denota-se esse conjunto por ω.

$\vdash$ **Proposição 1.15.**
1. *(Princípio da Indução Finita) O único subconjunto indutivo de ω é ω;*
2. *Todo elemento de ω é o conjunto vazio ou o sucessor de um outro elemento.*

$$\forall n(n \in \omega \to (n = \emptyset \vee \exists n'(n' \in \omega \wedge n = \mathrm{H}(n')))$$

$\square$ *Demonstração.* Segue direto da definição de ω, pois se $I \subseteq \omega$ é indutivo, então $\omega \subseteq I$, logo $I = \omega$. ■

1.8 Axioma da escolha

Para que o axioma da escolha seja compreensível, devem-se definir alguns conceitos antes. Essencialmente, o axioma da escolha é sobre produto de conjuntos e sobre funções. O nome escolha, de fato, vem de uma função, a função escolha. Para definir o conceito de função, é necessário primeiro definir o que é um par ordenado de elementos de dois conjuntos e o que é o conjunto de pares ordenados desses conjunto, que é chamado produto dos conjuntos. A partir desse produto de dois conjuntos, definem-se função e, a partir de função, define-se o produto de qualquer conjunto.

Pares ordenados, produto de par e função

:⊢ **Definição 1.12.** Sejam X e Y conjuntos. O *par ordenado* com *primeira coordenada* X e *segunda coordenada* Y é o conjunto

$$(X, Y) := \{\{X\}, \{X, Y\}\} \in \wp(\wp(X \cup Y)).$$

⊢ **Proposição 1.16.** *Sejam X, Y, Z e W conjuntos. Então*

$$(X, Y) = (Z, W) \iff X = Z \text{ e } Y = W.$$

:⊢ **Definição 1.13.** Sejam X e Y conjuntos. O *produto* de X por Y é o conjunto

$$X \times Y := \{(x, y) \in \wp(\wp(X \cup Y)) \mid x \in X \text{ e } y \in Y\}.$$

A existência desse conjunto depende da união de pares, do conjunto das partes e do axioma de especificação.

:⊢ **Definição 1.14.** Sejam X e Y conjuntos. Uma *função* de X para Y é um conjunto $f \subseteq X \times Y$ que satisfaz

$$\forall x \in X \; \exists! y \in Y \qquad (x, y) \in f.$$

Esse y é a *imagem* de x, denotada por $f(x)$. Denotam-se $f : X \to Y$ e $f(x) := y$. Para qualquer conjunto $K \subseteq X$, defini-se a *imagem* de K

$$f(K) = \{y \in Y \mid \exists k \in K \quad y = f(k)\},$$

que é subconjunto de Y. Diz-se que o conjunto $f(X)$ é a *imagem* de f.

⊢ **Proposição 1.17.** *Seja $f : A \to B$.*

1. $A = \emptyset \iff f = \emptyset$.
2. $B = \emptyset \implies A = \emptyset$.

□ *Demonstração.* 1. Suponhamos que $A = \emptyset$. Primeiro, notemos que $f = \emptyset$ é uma função de $\emptyset$ em B. Claramente, $f = \emptyset \subseteq \emptyset \times B$. Ainda, se f não fosse função de $\emptyset$ em B, existiria $a \in \emptyset$ tal que não existe único $b \in B$ satisfazendo $(a, b) \in f$. Mas existir $a \in \emptyset$ é um absurdo. Logo f é função. Por fim, se $g : \emptyset \times B$ é uma função, como $\emptyset \times B = \emptyset$, então $g \subseteq \emptyset \times B = \emptyset$, logo $g = \emptyset = f$.

Reciprocamente, suponhamos $f = \emptyset$. Se $A \neq \emptyset$, seja $a \in A$. Como f é função, existe $b \in B$ tal que $(a, b) \in f = \emptyset$, o que é absurdo. Portanto $A = \emptyset$.

2. Suponhamos que $A \neq \emptyset$. Então existe $a \in A$ e, como f é função, existe único $b \in \emptyset$ tal que $(a, b) \in f$. Mas $b \in \emptyset$ é absurdo, o que mostra que $A = \emptyset$.

■

O axioma da escolha e produto de conjuntos

:⊢ **Definição 1.15.** Seja C um conjunto. O *produto* de C é o conjunto

$$\prod C := \left\{ f : C \to \bigcup C \;\middle|\; \forall X \in C \quad f(X) \in X \right\}.$$

$$\prod C := \left\{ f \in \left(\bigcup C\right)^C \;\middle|\; \forall X \in C \quad f(X) \in X \right\}.$$

⊢ **Proposição 1.18.** *Seja C um conjunto. Então*

1. $C = \emptyset \implies \prod C = \{\emptyset\}$;
2. $\emptyset \in C \implies \prod C = \emptyset$.

□ *Demonstração.* 1. Como $C = \emptyset$, então $\bigcup \emptyset = \emptyset$. A função $\emptyset : \emptyset \to \emptyset$ é uma função em $\prod C$, pois satisfaz por vacuidade que $\forall X \in C \quad f(X) \in X$. Se não satisfizesse, existiria $X \in \emptyset$ tal que $f(X) \notin X$, o que é contradição. Isso mostra que $\emptyset \in \prod \emptyset$. Agora, seja $f \in \prod \emptyset$ função de $\emptyset$ em $\emptyset$. Como o domínio de f é $\emptyset$, segue que $f = \emptyset$.

2. Suponha que existe $f \in \prod C$. Então $f : C \to \bigcup C$ satisfaz que $\forall X \in C \quad f(X) \in X$. Como $\emptyset \in C$, existe $f(\emptyset) \in \bigcup C$ e, pela propriedade, $f(\emptyset) \in \emptyset$, contradição. Portanto $\prod C = \emptyset$.

■

Axioma 7 (Escolha)**.** Seja C um conjunto tal que $\emptyset \notin C$. Então $\prod C \neq \emptyset$.

1.9 Axiomas da fundação e substituição

Citamos aqui os dois últimos axiomas da teoria de conjuntos, mas não estudaremos eles aqui.

Axioma 8 (Fundação)**.** Todo conjunto não vazio tem algum elemento com o qual sua interseção é vazia.

$$\forall x (x \neq \emptyset \to \exists y (y \in x \wedge x \cap y = \emptyset))$$

Esse axioma é essencial na teoria de conjuntos, e também é conhecido como axioma da regularidade.

Axioma 9 (Substituição (Esquema))**.** Seja f uma fórmula com variáveis livres $x, C, v_0, \cdots, v_n$. Para todos $v_0, \cdots, v_n$ e todo conjunto C, algum conjunto S tem como seus únicos elementos os elementos de C tais que vale $f(x, C, v_0, \cdots, v_n)$.

$$\begin{aligned} \forall v_0, \cdots, v_n \forall C \Big(& (\forall x (x \in C \to \exists! y f(x, y, C, v_0, \cdots, v_n))) \\ & \to \exists B \forall y (y \in B \leftrightarrow \exists x \in C f(x, y, C, v_0, \cdots, v_n)) \Big) \end{aligned}$$

Os axiomas da especificação e do par são consequência do axioma da substituição.

Propriedades gerais

Contenção

⊢ **Proposição 1.19.** *Sejam X, Y e Z conjuntos. Então*

1. $X \subseteq X$;
2. $X \subseteq Y$ *e* $Y \subseteq X \iff X = Y$;
3. $X \subseteq Y$ *e* $Y \subseteq Z \implies X \subseteq Z$.

□ *Demonstração.* 1. Se $X = \emptyset$, então $\emptyset \subseteq X = \emptyset$. Logo $X \subseteq X$. Caso contrário, seja $x \in X$. Então $x \in X$. Logo $X \subseteq X$.

2. $X \subseteq Y$ e $Y \subseteq X$ se, e somente se, $\forall x \in X\ x \in Y$ e $\forall y \in Y\ y \in X$, o que é equivalente a $X = Y$ pelo axioma da extensão.
3. Se $X = \emptyset$, então $X \subseteq Z$. Caso contrário, seja $x \in X$. Então, como $X \subseteq Y$, $x \in Y$ e, como $Y \subseteq Z$, $x \in Z$. Logo $X \subseteq Z$.

■

União e interseção

⊢ **Proposição 1.20.** *Sejam X, Y e Z conjuntos. Então*

1. $X \cup \emptyset = X$;
2. $X \cup Y = Y \cup X$;
3. $(X \cup Y) \cup Z = X \cup (Y \cup Z)$;
4. $X \cup X = X$;
5. $X \subseteq Y \iff X \cup Y = Y$.

⊢ **Proposição 1.21.** *Sejam X, Y e Z conjuntos. Então*

1. $X \cap \emptyset = \emptyset$;
2. $X \cap Y = Y \cap X$;
3. $(X \cap Y) \cap Z = X \cap (Y \cap Z)$;
4. $X \cap X = X$;
5. $X \subseteq Y \iff X \cap Y = X$.

⊢ **Proposição 1.22.** *Sejam X, Y e Z conjuntos. Então*

1. $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$;
2. $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$.

Capítulo 2

Famílias e propriedades de conjuntos

2.1 Famílias e indexações

Os axiomas já foram todos enunciados no capítulo anterior e as bases da teoria de conjuntos clássica está construída. Sendo assim, é necessário mudar a linguagem com que muitas das operações sobre conjuntos são tratadas, entre elas a união, a interseção e o produto. O conceito de uma família será definido nesta seção. Embora inicialmente a notação do capítulo anterior seja mais simples, eventualmente a notação de famílias com índices será necessária, por dois motivos principais. O primeiro é que esse conceito facilitará muito os enunciados de várias propriedades e teoremas na matemática eventualmente. O segundo é que tradicionalmente os matemáticos usam famílias e índices para denotar uniões, interseções, produtos e muitas outras noções. A ideia básica de uma família é a seguinte. Quando se define a união de um conjunto C na teoria de conjuntos, a ideia intuitiva por trás da definição é que se estão unindo os conjuntos que são elementos de C. A união de um par $\{X, Y\}$ é denotada $X \cup Y$ não por acaso, essa notação indica um conjunto que está sendo formado com os elementos de X e Y, não com os elementos dos elementos de C, como no caso de $\bigcup C$.

Generalizando essa ideia, a união de três conjuntos X_1, X_2, X_3 pode ser denotada $X_1 \cup X_2 \cup X_3$ e o mesmo pode ser feito para qualquer quantidade finita de conjuntos. Mas para fazer o mesmo para uma quantidade qualquer de conjuntos, não é possível escrever esses conjuntos numa lista. Por isso surgiu a ideia de *indexar* os conjuntos de C que se pretende unir, usando a notação $(X_i)_{i \in I}$, sendo que cada X_i é um elemento de C e i seu índice. Em seguida, indica-se na parte inferior do símbolo de

união que os conjuntos indexados estão sendo unidos, de modo que $\bigcup C$ é denotado

$$\bigcup_{i \in C} X_i.$$

Essa notação tem a vantagem de estar mais próxima da intuição e também permite trabalhar com duplas uniões mais facilmente. As mesmas ideias são aplicadas para interseções e produtos. No entanto, ainda resta um problema, o problema principal. Tendo já especificada qual é a notação que pretende-se aplicar, ainda falta definir o que é uma família somente a partir dos conceitos da teoria de conjuntos. Essa definição vem a seguir.

:⊢ **Definição 2.1.** Sejam C e I conjuntos não vazios. Uma *família* de elementos de C indexados por I é uma função $F : I \to C$. O conjunto I é o *conjunto de índices* da família. Denota-se isso por $(F_i)_{i \in I}$ e a imagem de $i \in I$ por F é denotada F_i e chamada de *i-ésimo membro* da família. Uma *sequência* é uma família em que $I = \mathbb{N}$, e uma *sequência finita* é uma família em que $I \in \mathbb{N}$ (ou seja, $I = \{0, \ldots, n-1\}$ para algum $n \in \mathbb{N}$, e nesse caso diz-se n-sequência).

Vale notar que uma família é vazia se, e somente se, $I = \emptyset$. Uma família é uma função e, portanto, quando se afirma que uma família, afirma-se que uma função é vazia, ou que é a função vazia. Mas isso ocorre se, e somente se, seu domínio, no caso o conjunto de índices, é vazio.

:⊢ **Definição 2.2.** Seja X um conjunto não vazio. Uma *indexação* de X é uma família bijetiva $(x_i)_{i \in I}$ de elementos de X. Nesse caso, X é um conjunto indexado por I e denota-se $X = \{x_i\}_{i \in I}$.

A noção de uma família é, de fato, mais motivada por notação do que por um conceito teórico, já que uma família é simplesmente uma função sem nenhuma restrição, e a única diferença entre uma família é uma função é o contexto. Uma pergunta relevante, ainda, é se todo conjunto pode ser indexado por meio de uma família. Essa pergunta tem uma resposta óbvia e uma não óbvia, e ambas afirmam que sim. A resposta óbvia é que, para se indexar um conjunto C basta considerar a função $F : C \to C$ definida para todo $X \in C$ por $F(X) = X$. Desse modo, essa é uma indexação do conjunto X. Mas essa resposta não satisfaz a tradição de indexar um quantidade finita de conjuntos $\{X, Y\}$ com números naturais. A resposta menos óbvia é que todo conjunto pode ser bem ordenado e, dessa forma, existe uma função de um número ordinal para o conjunto, logo uma indexação desse conjunto por um número ordinal. Os números naturais são os números ordinais finitos, o que significa que essa resposta menos óbvia condiz com a indexação que se faz usualmente de uma quantidade finita de conjuntos. Esse tópicos, no entanto, não serão abordados nesse capítulo.

2.2 Propriedades de união e interseção

A partir da definição de família, pode-se definir a união e a interseção de uma família de conjuntos a partir da imagem do conjunto de índices I pela função C, o conjunto $C(I) = \{C_i \mid i \in I\}$. No entanto, um problema teórico se manifesta para se definir uma família de conjuntos. Se uma família é uma função de um conjunto de índices em um conjunto de elementos, para se definir uma família de conjuntos deveria existir um conjunto de todos conjuntos para fazer o papel de contradomínio de uma família. Esse conjunto, no entanto, não existe na teoria de conjuntos abordada neste livro, o que sugere que a definição de uma família de conjuntos depende, de fato, de um conjunto cujos elementos são os conjuntos da família de conjuntos. A existência desse conjunto de conjuntos é suposta, mas ele não é o conjunto de todos os conjuntos. Sendo assim, sempre que se enunciar uma família de conjuntos, essas ressalvas serão assumidas.

:⊢ **Definição 2.3.** A *união* de uma família $(C_i)_{i\in I}$ de conjuntos é o conjunto

$$\bigcup_{i\in I} C_i := \bigcup C(I).$$

A *interseção* de uma família não vazia $(C_i)_{i\in I}$ de conjuntos é o conjunto

$$\bigcap_{i\in I} C_i := \bigcap C(I).$$

Quando I for finito, pode-se denotar

$$C_1 \cup \cdots \cup C_n := \bigcup_{i\in I} C_i \qquad \text{e} \qquad C_1 \cap \cdots \cap C_n := \bigcap_{i\in I} C_i.$$

⊢ **Proposição 2.1.** *Seja $(C_i)_{i\in I}$ uma família de conjuntos. Então*

1. $\forall i \in I \quad C_i = \emptyset \qquad \Leftrightarrow \qquad \bigcup_{i\in I} C_i = \emptyset.$
2. $\exists i \in I \quad C_i = \emptyset \qquad \Rightarrow \qquad \bigcap_{i\in I} C_i = \emptyset.$

⊢ **Proposição 2.2.** *Sejam X um conjunto e $(C_i)_{i\in I}$ uma família não vazia de subconjuntos de X. Então*

1. $\left(\bigcap_{i\in I} C_i\right)^{\complement} = \bigcup_{i\in I} (C_i)^{\complement}$
2. $\left(\bigcup_{i\in I} C_i\right)^{\complement} = \bigcap_{i\in I} (C_i)^{\complement}$

□ *Demonstração.* 1. Para isso, basta notar que $c \in \left(\bigcap_{i\in I} C_i\right)^{\complement}$ se, e somente se, $c \notin \bigcap_{i\in I} C_i$. Mas isso ocorre se, e somente se, existe $i \in I$ tal que $c \notin C_i$. Essa afirmação é equivalente a $c \in (C_i)^{\complement}$ que, por sua vez, é equivalente a $c \in \bigcup_{i\in I}(C_i)^{\complement}$.

2. Como, para todo conjunto C, $(C^{\complement})^{\complement} = C$, segue do item anterior que

$$\left(\bigcup_{i\in I} C_i\right)^{\complement} = \left(\bigcup_{i\in I}((C_i)^{\complement})^{\complement}\right)^{\complement} = \left(\left(\bigcap_{i\in I}(C_i)^{\complement}\right)^{\complement}\right)^{\complement} = \bigcap_{i\in I}(C_i)^{\complement}.$$

■

⊢ **Proposição 2.3.** *Seja (C_{ij}) uma família não vazia de conjuntos. Então*

$$\bigcup_{i\in I}\left(\bigcap_{j\in J} C_{ij}\right) \subseteq \bigcap_{j\in J}\left(\bigcup_{i\in I} C_{ij}\right)$$

⊢ **Proposição 2.4.** *Seja (C_{ij}) uma família não vazia de conjuntos. Então*

1. $\displaystyle\bigcap_{i\in I} \wp\left(C_i\right) = \wp\left(\bigcap_{i\in I} C_i\right)$;
2. $\displaystyle\bigcup_{i\in I} \wp\left(C_i\right) \subseteq \wp\left(\bigcup_{i\in I} C_i\right)$.

2.3 Produto de conjuntos

:⊢ **Definição 2.4.** Seja $(C_i)_{i\in I}$ uma família de conjuntos. O *produto* de $(C_i)_{i\in I}$ é o conjunto

$$\prod_{i\in I} C_i := \{(c_i)_{i\in I} \mid \forall i \in I \quad c_i \in C_i\}.$$

As famílias $(c_i)_{i\in I}$ são de elementos em $\bigcup_{i\in I} C_i$.

:⊢ **Definição 2.5.** Seja $(C_i)_{i\in I}$ uma família de conjuntos e $i \in I$. A *projeção canônica* de $\prod_{i\in I} C_i$ em C_i é a função

$$\begin{aligned} \pi_i \colon \prod_{i\in I} C_i &\longrightarrow C_i \\ (c_i)_{i\in I} &\longmapsto c_i. \end{aligned}$$

⊢ **Proposição 2.5** (Propriedade Universal)**.** *Sejam $(C_i)_{i\in I}$ uma família de conjuntos, X um conjunto e, para todo $i \in I$, $f_i : X \to C_i$ uma função. Então existe uma única função $f : X \to \prod_{i\in I} C_i$ tal que, para todo $i \in I$, $\pi_i \circ f = f_i$ (o diagrama comuta).*

$$\begin{tikzcd} & \prod_{i\in I} C_i \arrow[d, "\pi_i"] \\ X \arrow[ur, dashed, "f"] \arrow[r, "f_i"'] & C_i \end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} f \colon X &\longrightarrow \prod_{i\in I} C_i \\ x &\longmapsto (f_i(x))_{i\in I}. \end{aligned}$$

Para todo $x \in X$ e para todo $i \in I$,

$$\pi_i \circ f(x) = \pi_i(f(x)) = \pi_i((f_i(x))_{i\in I}) = f_i(x).$$

Portanto $\pi_i \circ f = f_i$. Isso mostra a existência da f. Para a unicidade, seja $\overline{f} : X \to \prod_{i\in I} C_i$ função tal que, para todo $i \in I$, $\pi_i \circ \overline{f} = f_i$. Seja $x \in X$. Como $\overline{f}(x) \in \prod_{i\in I} C_i$, $\overline{f}(x) = (x_i)_{i\in I}$. Da propriedade comutativa de $\overline{f}$, segue que, para todo $i \in I$,

$$x_i = \pi_i \circ \overline{f}(x) = f_i(x).$$

Como $f(x) = (f_i(x))_{i\in I}$, isso mostra que $\overline{f}(x) = f(x)$. Portanto $\overline{f} = f$. ■

2.4 Coproduto de conjuntos

:⊢ **Definição 2.6.** Seja $(C_i)_{i\in I}$ uma família não vazia de conjuntos. O *coproduto* de $(C_i)_{i\in I}$ é o conjunto

$$\bigsqcup_{i\in I} C_i := \{(i,c) \mid i \in I \ \text{ e } \ c \in C_i\}.$$

:⊢ **Definição 2.7.** Seja $(C_i)_{i\in I}$ uma família de conjuntos e $i \in I$. A *inclusão canônica* de C_i em $\bigsqcup_{i\in I} C_i$ é a função

$$\begin{aligned} \iota_i \colon C_i &\longrightarrow \bigsqcup_{i\in I} C_i \\ c &\longmapsto (i,c). \end{aligned}$$

⊢ **Proposição 2.6** (Propriedade Universal)**.** *Sejam $(C_i)_{i\in I}$ uma família de conjuntos, X um conjunto e, para todo $i \in I$, $f_i : C_i \to X$ uma função. Então existe uma única função $f : \bigsqcup_{i\in I} C_i \to X$ tal que, para todo $i \in I$, $f \circ \iota_i = f_i$ (o diagrama comuta).*

$$\begin{tikzcd} \bigsqcup_{i\in I} C_i \arrow[dr, dashed, "f"] & \\ C_i \arrow[u, "\iota_i"] \arrow[r, "f_i"'] & X \end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} f \colon \bigsqcup_{i\in I} C_i &\longrightarrow X \\ (i,c) &\longmapsto f_i(c). \end{aligned}$$

Seja $i \in I$ e $c \in C_i$. Então

$$f \circ \iota_i(c) = f(\iota_i(c)) = f(i,c) = f_i(c).$$

Portanto $f \circ \iota_i = f_i$. Isso mostra a existência da f. Para a unicidade, seja $\overline{f} : \bigsqcup_{i\in I} C_i \to X$ função tal que, para todo $i \in I$, $\overline{f} \circ \iota_i = f_i$. Seja $x \in \bigsqcup_{i\in I} C_i$. Existem $i \in I$ e $c \in C_i$ tais que $x = (i,c)$. Da propriedade comutativa de $\overline{f}$, segue que

$$\overline{f}(x) = \overline{f}(i,c) = \overline{f}(\iota_i(x)) = \overline{f} \circ \iota_i(c) = f_i(c) = f(i,c) = f(x).$$

Isso mostra que $\overline{f} = f$. ■

2.4.1 Propriedades de produto e coproduto

⊢ **Proposição 2.7.** *Seja* $(C_{ij})_{(i,j)\in I\times J}$ *uma família de conjuntos. Então*

1. $\bigcup_{j\in J}\left(\prod_{i\in I} C_{ij}\right) \subseteq \prod_{i\in I}\left(\bigcup_{j\in J} C_{ij}\right)$;
2. $\bigcap_{j\in J}\left(\prod_{i\in I} C_{ij}\right) = \prod_{i\in I}\left(\bigcap_{j\in J} C_{ij}\right)$.

□ *Demonstração.* 1.

$$\begin{aligned}
c \in \bigcup_{j\in J}\left(\prod_{i\in I} C_{ij}\right) &\implies \exists j \in J \left(c \in \prod_{i\in I} C_{ij}\right) \\
&\implies \exists j \in J\ \forall i \in I\,(c_i \in C_{ij}) \\
&\implies \forall i \in I \left(c \in \bigcup_{j\in J} C_{ij}\right) \\
&\implies c \in \prod_{i\in I}\left(\bigcup_{j\in J} C_{ij}\right).
\end{aligned}$$

2.

$$\begin{aligned}
c \in \bigcap_{j\in J}\left(\prod_{i\in I} C_{ij}\right) &\iff \forall j \in J \left(c \in \prod_{i\in I} C_{ij}\right) \\
&\iff \forall j \in J\ \forall i \in I\,(c_i \in C_{ij}) \\
&\iff \forall i \in I\ \forall j \in J\,(c_i \in C_{ij}) \\
&\iff \forall i \in I \left(c_i \in \bigcup_{j\in J} C_{ij}\right) \\
&\iff c \in \prod_{i\in I}\left(\bigcap_{j\in J} C_{ij}\right).
\end{aligned}$$

■

Notemos que a inclusão contrária no primeiro item nao vale. Suponhamos que para um $j_0 \in J$, todos os C_{ij_0} são vazios, mas para todos outros $j \in J$, os C_{ij} não são vazios. Então o produto desses C_{ij} será sempre vazio, pois sempre tem um dos elementos do produto vazio, e então a união desses produtos será vazia; no entanto, a união desses C_{ij} não será nenhuma vazia e, então, o produto não será vazio (pelo axioma da escolha).

$\vdash$ **Proposição 2.8.** *Sejam X um conjunto, $(Y_i)_{i\in I}$ uma família de conjuntos, $(S_i)_{i\in I}$ uma família de subconjuntos de $(Y_i)_{i\in I}$, $f : X \to \prod_{i\in I} Y_i$ uma função e, para todo $i \in I$, $f_i := \pi_i \circ f$. Então*

$$f^{-1}\left(\prod_{i\in I} S_i\right) = \bigcap_{i\in I} f_i^{-1}(S_i).$$

$\square$ *Demonstração.* Note que $x \in f^{-1}(\prod_{i\in I} S_i)$ é equivalente a $f(x) \in \prod_{i\in I} S_i$, que por sua vez ocorre se, e somente se, para todo $i \in I$, $\pi_i(f(x)) \in S_i$. Como $f_i(x) = \pi_i(f(x)) \in S_i$, isso é equivalente a, para todo $i \in I$, $x \in f_i^{-1}(S_i)$. ■

Notação alternativa

$$\prod_{i\in I} C_i = \{\lceil c_i \rceil_{i\in I} \mid \forall i \in I\ c_i \in C_i\}$$

$\lceil c_i \rceil_{i\in I} = c : I \to \bigcup_{i\in I} C_i$

$$\bigsqcup_{i\in I} C_i = \{\lfloor c \rfloor_i \mid i \in I, c \in C_i\}$$

$\lfloor c \rfloor_i = (i, c)$

2.5 Complementares e diferença simétrica

:$\vdash$ **Definição 2.8.** Sejam X e Y conjuntos. O *complementar relativo* de Y em X é o conjunto

$$X \setminus Y := \{x \in X \mid x \notin Y\}.$$

:$\vdash$ **Definição 2.9.** Sejam X um conjunto e S um subconjunto de X. O *complementar* de S em X é o conjunto

$$S^{\complement} := X \setminus S.$$

:$\vdash$ **Definição 2.10.** Sejam X e Y conjuntos. A *diferença simétrica* de X e Y é o conjunto

$$X \triangle Y := (X \setminus Y) \cup (Y \setminus X).$$

2.5.1 Propriedades

$\vdash$ **Proposição 2.9.** *Sejam X, Y subconjuntos de U. Então*

1. $(X^{\complement})^{\complement} = X$;
2. $\emptyset^{\complement} = U$ *e* $U^{\complement} = \emptyset$;

3. $X \cap X^{\complement} = \emptyset$ *e* $X \cup X^{\complement} = U$;
4. $X \subseteq Y \iff Y^{\complement} \subseteq X^{\complement}$.
5. $(X \cup Y)^{\complement} = X^{\complement} \cap Y^{\complement}$ *e* $(X \cap Y)^{\complement} = X^{\complement} \cup Y^{\complement}$.

Capítulo 3

Relações

Definimos brevemente o que são relações e, em especial, relações binárias, antes de abordar os casos específicos de funções, equivalências e ordens.

:⊢ **Definição 3.1.** Sejam X e Y conjuntos. Uma *relação* R de X para Y é um subconjunto de $X \times Y$. Os conjuntos X e Y são, respectivamente, o *domínio* e o *contradomínio* de R. Denota-se $x\ R\ y$ para $(x, y) \in R$.

:⊢ **Definição 3.2.** Seja A um conjunto não vazio. Uma *relação binária* R em A é uma relação R de A para A.

:⊢ **Definição 3.3.** Seja A um conjunto não vazio e R uma relação binária em A. Definem-se as seguintes propriedades de R:

1. (Reflexividade) $\forall a \in A \qquad aRa$;
2. (Irreflexividade) $\nexists a \in A \qquad aRa$;
3. (Simetria) $\forall a_1, a_2 \in A \qquad a_1Ra_2 \Leftrightarrow a_2Ra_1$;
4. (Antissimetria) $\forall a_1, a_2 \in A \qquad a_1Ra_2 \text{ e } a_2Ra_1 \Rightarrow a_1 = a_2$;
5. (Transitividade) $\forall a_1, a_2, a_3 \in A \qquad a_1Ra_2 \text{ e } a_2Ra_3 \Rightarrow a_1Ra_3$;
6. (Totalidade) $\forall a_1, a_2 \in A \qquad a_1Ra_2 \text{ ou } a_2Ra_1$.

Uma relação que satisfaz as propriedades acima é, respectivamente, reflexiva, simétrica, antissimétrica, transitiva e total.

3.1 Funções

:⊢ **Definição 3.4.** Seja R uma relação de X em Y. A *relação inversa* de R é a relação R^{-1} de Y em X definida por

$$\forall x \in X\ \forall y \in Y \qquad x\ R\ y \Leftrightarrow y\ R^{-1}\ x.$$

3.1.1 Definição e propriedades básicas

:⊢ **Definição 3.5.** Sejam A e B conjuntos. Uma *função* de A para B é uma relação f de A para B tal que

1. (Funtorialidade) Para todo $a \in A$, existe único $b \in B$ tal que $(a, b) \in f$.

Denota-se $f : A \to B$. O conjunto das funções de A para B é denotado $\mathcal{F}(A, B)$ ou B^A. O conjunto das funções de A para A é denotado $\mathcal{F}(A)$.

A *imagem* de $a \in A$ é o único $b \in B$ que satisfaz $(a, b) \in f$. Denota-se $b = f(a)$. Ambas informações podem ser denotadas por

$$\begin{aligned} f\colon A &\longrightarrow B \\ a &\longmapsto b. \end{aligned}$$

⊢ **Proposição 3.1.** *Seja $f : A \to B$ uma função. Então*

1. $A = \emptyset \iff f = \emptyset$.
2. $B = \emptyset \implies A = \emptyset$.

□ *Demonstração.* 1. Suponhamos que $A = \emptyset$. Primeiro, notemos que $f = \emptyset$ é uma função de $\emptyset$ em B. Claramente, $f = \emptyset \subseteq \emptyset \times B$. Ainda, se f não fosse função de $\emptyset$ em B, existiria $a \in \emptyset$ tal que não existe único $b \in B$ satisfazendo $(a, b) \in f$. Mas existir $a \in \emptyset$ é uma contradição. Logo f é função. Por fim, se $g : \emptyset \times B$ é uma função, como $\emptyset \times B = \emptyset$, então $g \subseteq \emptyset \times B = \emptyset$, logo $g = \emptyset = f$.

Reciprocamente, suponhamos $f = \emptyset$. Se $A \neq \emptyset$, seja $a \in A$. Como f é função, existe $b \in B$ tal que $(a, b) \in f = \emptyset$, o que é contradição. Portanto $A = \emptyset$.

2. Suponhamos que $A \neq \emptyset$. Então existe $a \in A$ e, como f é função, existe único $b \in \emptyset$ tal que $(a, b) \in f$. Mas $b \in \emptyset$ é absurdo, o que mostra que $A = \emptyset$.

■

⊢ **Proposição 3.2.** *Sejam $f : A \to B$ e $g : A' \to B'$. Então*

$$f = g \iff A = A' \ \text{ e } \ \forall a \in A \quad f(a) = g(a).$$

□ *Demonstração.* Suponhamos que $f = g$. Se $A = \emptyset$, então $f = \emptyset$ e $g = f = \emptyset$, o que implica $A' = \emptyset$. Ainda, para todo $a \in A$, $f(a) = g(a)$ pois, se isso fosse falso, existiria $a \in \emptyset$ tal que $f(a) \neq g(a)$, mas existir $a \in \emptyset$ é absurdo. Se $A \neq \emptyset$, seja $a \in A$. Então existe $b \in B$ tal que $(a, b) \in f$ e, como $f = g$, $(a, b) \in g$. Isso implica $a \in A'$ e concluímos que $A \subseteq A'$. Por outro lado, seja $a \in A'$. Então existe $b \in B'$ tal que $(a, b) \in g$ e, como $f = g$, $(a, b) \in f$. Isso implica $a \in A$ e concluímos que $A' \subseteq A$. Portanto $A = A'$. Agora, seja $a \in A$. Então existem $f(a) \in B$ e $g(a) \in B'$.

Como $(a, f(a)) \in f$ e $f = g$, então $(a, f(a)) \in g$. Como f é função, existe único $b \in B$ tal que $(a, b) \in f$, o que implica $f(a) = g(a)$.

Reciprocamente, suponhamos que $A = A'$ e que, para todo $a \in A$, $f(a) = g(a)$. Se $A = \emptyset$, então $f = \emptyset$ e $g = \emptyset$, logo $f = g$. Se $A \neq \emptyset$, então seja $p \in f$. Existe $a \in A$ tal que $p = (a, f(a))$. Como $f(a) = g(a)$, então $p = (a, g(a))$; mas $(a, g(a)) \in g$, o que implica $p \in g$ e, portanto, $f \subseteq g$. Agora, seja $p \in g$. Existe $a \in A'$ tal que $p = (a, g(a))$. Como $f(a) = g(a)$, então $p = (a, f(a))$; mas $(a, f(a)) \in f$, o que implica $p \in f$ e, portanto, $f \subseteq g$. Assim, concluímos que $f = g$. ■

:⊢ **Definição 3.6.** Sejam $f : A \to B$ uma função e $C \subseteq A$ um conjunto. O *conjunto imagem* de C sob f é

$$f(C) = \{y \in Y \mid \exists c \in C \quad y = f(c)\}.$$

O conjunto $f(A)$ é o *imagem* de f.

⊢ **Proposição 3.3.** *Seja $f : A \to B$. Então $f : A \to f(A)$.*

:⊢ **Definição 3.7.** Sejam $f : A \to B$ uma função e $A' \subseteq A$ um conjunto. A *restrição* de f a A' é a função

$$\begin{aligned} f|_{A'} \colon A' &\longrightarrow B \\ a &\longmapsto f(a). \end{aligned}$$

⊢ **Proposição 3.4.** *Sejam $f : A \to B$, $A' \subseteq A$ e $B' \subseteq B$. Então a restrição $f|_{A'}$ é uma função de A' em B' se, e somente se, $f(A') \subseteq B'$.*

□ *Demonstração.* Se que $f|_{A'}$ é uma função de A' em B', então o contradomínio de $f_{A'}$ é B', o que significa que, para todo $a \in A'$, $f(a) = f|_{A'}(a) \in B'$, logo $f(A') \subseteq B$. Reciprocamente, se, para todo $a \in A'$, $f(a) \in B'$, então $f|_{A'}$ é uma função de A' em B'. ■

3.1.2 Composição de funções

:⊢ **Definição 3.8.** Sejam $f : A \to B'$ e $g : B \to C$ funções tais que $B' \subseteq B$. A *função composta* de g com f é a função

$$\begin{aligned} g \circ f \colon A &\longrightarrow C \\ a &\longmapsto g(f(a)). \end{aligned}$$

⊢ **Proposição 3.5.** *Sejam $f : A \to B'$, $g : B \to C'$ e $h : C \to D$ funções tais que $B' \subseteq B$ e $C' \subseteq C$. Então*

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

□ *Demonstração.* Primeiro, notemos que $g \circ f$ é uma função de A em C', o que implica que $h \circ (g \circ f)$ é uma função de A em D. Anda, notemos que $h \circ g$ é uma função de B em D, o que implica que $(h \circ g) \circ f$ é uma função de A em D. Logo os domínios de $h \circ (g \circ f)$ e $(h \circ g) \circ f$ são iguais. Se $A = \emptyset$, então $h \circ (g \circ f) = (h \circ g) \circ f = \emptyset$. Suponhamos, então, que $A \neq \emptyset$ e seja $a \in A$. Então

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

o que mostra que $h \circ (g \circ f) = (h \circ g) \circ f$. ■

⊢ **Proposição 3.6.** *Seja $f\colon A \to B$. Então*

1. *$f \circ \emptyset = \emptyset$;*
2. *$\emptyset \circ f = \emptyset$.*

□ *Demonstração.* Para a primeira igualdade, notemos que $f \circ \emptyset$ é uma função de $\emptyset$ em B e, portanto, $f \circ \emptyset = \emptyset$. Para a segunda igualdade, notemos que $\emptyset \circ f$ é uma função de A em $\emptyset$ e, portanto, $A = \emptyset$, o que é equivalente a $\emptyset \circ f = \emptyset$. ■

:⊢ **Definição 3.9.** Seja A um conjunto não vazio. A *função identidade* em A é a função

$$\begin{aligned} \mathrm{I}_A\colon A &\longrightarrow A \\ a &\longmapsto a. \end{aligned}$$

⊢ **Proposição 3.7.** *Seja $f\colon A \to B$ uma função. Então*

$$f \circ \mathrm{I}_A = f \quad e \quad \mathrm{I}_B \circ f = f.$$

□ *Demonstração.* Primeiro, notemos que $f \circ \mathrm{I}_A$ e $\mathrm{I}_B \circ f$ são funções de A em B e, portanto, têm o mesmo domínio de f. Se $A = \emptyset$, então $f : \emptyset \to B$ e, portanto, $f = \emptyset$. Notemos que $\mathrm{I}_\emptyset = \emptyset$. De fato, $\emptyset$ é função e, se não fosse identidade de $\emptyset$ em $\emptyset$, existiria $a \in \emptyset$ tal que $f(a) \neq a$; mas $a \in \emptyset$ é absurdo. Assim, $f \circ \mathrm{I}_A$ é uma função de $\emptyset$ em B e, portanto, $f \circ \mathrm{I}_A = \emptyset = f$. Ainda, $\mathrm{I}_B \circ f$ é uma função de $\emptyset$ em B e, portanto, $\mathrm{I}_B \circ f = \emptyset = f$. Se $A \neq \emptyset$, seja $a \in A$. Então $(f \circ \mathrm{I}_A)(a) = f(\mathrm{I}_A(a)) = f(a) = \mathrm{I}_B(f(a)) = (\mathrm{I}_B \circ f)(a)$. ■

3.1.3 Função inversa, injetividade e sobrejetividade

:⊢ **Definição 3.10.** Seja $f\colon A \to B$ uma função. Uma *função inversa* de f é uma função $g\colon B \to A$ tal que

$$g \circ f = \mathrm{I}_A \quad \text{e} \quad f \circ g = \mathrm{I}_B.$$

:⊢ **Definição 3.11.** Uma *função injetiva* (ou *injeção*) é uma função $f\colon A \to B$ que satisfaz

$$\forall a_1, a_2 \in A \qquad f(a_1) = f(a_2) \Rightarrow a_1 = a_2.$$

O conjunto das funções injetivas de A para B é denotado $\overset{\hookrightarrow}{\mathcal{F}}(A, B)$. O conjunto das funções injetivas de A para A é denotado $\overset{\hookrightarrow}{\mathcal{F}}(A)$.

:⊢ **Definição 3.12.** Uma *função sobrejetiva* sobre um conjunto B é uma função $f\colon A \to B$ que satisfaz $f(A) = B$. O conjunto das funções sobrejetivas de A para B é denotado $\overset{\twoheadrightarrow}{\mathcal{F}}(A, B)$. O conjunto das funções sobrejetivas de A para A é denotado $\overset{\twoheadrightarrow}{\mathcal{F}}(A)$.

:⊢ **Definição 3.13.** Sejam A e B conjunto. Uma *bijeção* entre A e B é uma função injetiva $f\colon A \to B$ que é sobrejetiva sobre B. O conjunto das funções bijetivas de A para B é denotado $\overset{\leftrightarrow}{\mathcal{F}}(A, B)$.

⊢ **Proposição 3.8.** *Seja $f\colon A \to B$. Então f é injetiva se, e somente se, existe $g\colon B \to A$ tal que $g \circ f = \mathrm{I}_A$.*

□ *Demonstração.* Suponhamos que f é injetiva. Se $A = \emptyset$. Então $f = \emptyset$ e, portanto, tomando $g = \mathrm{I}_B$, temos que $g \circ f = \mathrm{I}_B \circ \emptyset = \mathrm{I}_\emptyset = \emptyset$. Se $A \neq \emptyset$, seja $a \in A$. ■

⊢ **Proposição 3.9.** *Seja $f\colon A \to B$. Então f é sobrejetiva sobre B se, e somente se, existe $g\colon B \to A$ tal que $f \circ g = \mathrm{I}_B$.*

□ *Demonstração.* Suponhamos que f é sobrejetiva sobre B. Então $B = f(A)$; ou seja, para todo $b \in B$, existe $a \in A$ tal que $f(a) = b$ e, portanto, definimos a função $g\colon B \to A$ para cada elemento de B como $g(b) := a$. Assim, segue que $g \circ f = \mathrm{I}_B$. ■

⊢ **Proposição 3.10.** *Seja $f\colon A \to B$. Se $g\colon B \to A$ e $g'\colon B \to A$ são funções inversas de f, então $g = g'$.*

⊢ **Proposição 3.11.** *Sejam $f\colon A \to B'$ e $g\colon B \to C$ funções tais que $B' \subseteq B$. Se f e g são funções injetivas, então $g \circ f$ é uma função injetiva.*

□ *Demonstração.* Sejam $a_1, a_2 \in A$ tais que $g \circ f(a_1) = g \circ f(a_2)$. Então $g(f(a_1)) = g(f(a_2))$. Como g é injetiva, então $f(a_1) = f(a_2)$ e, como f é injetiva, então $a_1 = a_2$. Portanto $g \circ f$ é injetiva. ■

⊢ **Proposição 3.12.** *Sejam $f\colon A \to B$ e $g\colon B \to C$ funções. Se f e g são funções sobrejetivas, então $g \circ f$ é uma função sobrejetiva.*

□ *Demonstração.* Como f é sobrejetiva, então $f(A) = B$. Ainda, como g é sobrejetiva, então $g(B) = C$. Então $g \circ f(A) = g(f(A)) = g(B) = C$. Portanto $g \circ f$ é sobrejetiva. ■

3.1.4 Imagem inversa de função e propriedades

:⊢ **Definição 3.14.** Seja $f : A \to B$ uma função e $B' \subseteq B$. A *imagem inversa* de B sob f é o conjunto

$$f^{-1}(B') := \{a \in A \mid f(a) \in B'\}.$$

⊢ **Proposição 3.13.** *Seja $f : A \to B$ uma função, $B' \subseteq B$ e $(B_i)_{i\in I} \subseteq \wp(B)$ uma família de subconjuntos de B. Então*

1. $f^{-1}(\emptyset) = \emptyset$;
2. $f^{-1}(B) = A$;
3. $f^{-1}\left((B')^{\complement}\right) = (f^{-1}(B'))^{\complement}$;
4. $f^{-1}\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^{-1}(B_i)$;
5. $f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i)$.

□ *Demonstração.*

1. Suponha, por absurdo, que existe $a \in f^{-1}(\emptyset)$. Então $f(a) \in \emptyset$, o que é absurdo, e conclui-se $f^{-1}(\emptyset) = \emptyset$.
2. Seja $a \in A$. Como f é função de A em B, então existe $b \in B$ tal que $f(a) = b$, o que implica $a \in f^{-1}(B)$ e, então, $a \subseteq A$. Como a inclusão contrária vale por definição, então$f^{-1}(B) = A$.
3. Seja $a \in f^{-1}((B')^{\complement})$. Então $f(a) \in (B')^{\complement}$. Mas isso implica $a \notin f^{-1}(B')$, pois, caso contrário, seguiria que $f(a) \in B'$, o que contradiz a hipótese. Portanto $a \in (f^{-1}(B'))^{\complement}$; ou seja, $f^{-1}((B')^{\complement}) \subseteq (f^{-1}(B'))^{\complement}$. Reciprocamente, seja $a \in (f^{-1}(B'))^{\complement}$. Se, por absurdo, $f(a) \in B'$, então $a \notin f^{-1}(B')$, o que contradiz a hipótese. Portanto $f(a) \in (B')^{\complement}$, o que implica $a \in f^{-1}((B')^{\complement})$. Assim conclui-se que $(f^{-1}(B'))^{\complement} \subseteq f^{-1}((B')^{\complement})$ e, portanto, $f^{-1}((B')^{\complement}) = (f^{-1}(B'))^{\complement}$.
4. Seja $a \in f^{-1}(\bigcup_{i\in I} B_i)$. Então $f(a) \in \bigcup_{i\in I} B_i$. Isso significa que exite $i \in I$ tal que $f(a) \in B_i$. Portanto $a \in f^{-1}(B_i)$, e segue que $a \in \bigcup_{i\in I} f^{-1}(B_i)$; ou seja, $f^{-1}(\bigcup_{i\in I} B_i) \subseteq \bigcup_{i\in I} f^{-1}(B_i)$. Reciprocamente, seja $a \in \bigcup_{i\in I} f^{-1}(B_i)$. Então existe $i \in I$ tal que $a \in f^{-1}(B_i)$. Então $f(a) \in B_i$. Mas isso implica que $f(a) \in \bigcup_{i\in I} B_i$. Portanto $a \in f^{-1}(\bigcup_{i\in I} B_i)$; ou seja, $\bigcup_{i\in I} f^{-1}(B_i) \subseteq f^{-1}(\bigcup_{i\in I} B_i)$. Assim, conclui-se que $f^{-1}(\bigcup_{i\in I} B_i) = \bigcup_{i\in I} f^{-1}(B_i)$.
5. Seja $a \in f^{-1}(\bigcap_{i\in I} B_i)$. Então $f(a) \in \bigcap_{i\in I} B_i$. Isso significa que, para todo $i \in I$, $f(a) \in B_i$. Portanto, para todo $i \in I$, $a \in f^{-1}(B_i)$, e segue que $a \in \bigcap_{i\in I} f^{-1}(B_i)$; ou seja, $f^{-1}(\bigcap_{i\in I} B_i) \subseteq \bigcap_{i\in I} f^{-1}(B_i)$. Reciprocamente, seja $a \in \bigcap_{i\in I} f^{-1}(B_i)$. Então, para todo $i \in I$, $a \in f^{-1}(B_i)$. Então, para todo $i \in I$, $f(a) \in B_i$, o que implica que $f(a) \in \bigcap_{i\in I} B_i$. Portanto $a \in f^{-1}(\bigcap_{i\in I} B_i)$; ou seja, $\bigcap_{i\in I} f^{-1}(B_i) \subseteq f^{-1}(\bigcap_{i\in I} B_i)$. Assim, conclui-se que $f^{-1}(\bigcap_{i\in I} B_i) = \bigcap_{i\in I} f^{-1}(B_i)$. ■

3.1.5 Propriedades de imagem e imagem inversa

⊢ **Proposição 3.14.** *Sejam $f : D \to C$ uma função e $(C_i)_{i\in I}$ uma família de subconjuntos de C. Então*

1. $f(\emptyset) = \emptyset$;
2. $f(D) \subseteq C$;
3. $f\left(\bigcup_{i\in I} C_i\right) = \bigcup_{i\in I} f(C_i)$;

□ *Demonstração.*

1. Suponha, por absurdo, que existe $c \in f(\emptyset)$. Nesse caso, existe $d \in \emptyset$ tal que $f(d) = c$, o que é absurdo. Logo $f(\emptyset) = \emptyset$.
2. Se$f(D) = \emptyset$, então vale a proposição. Caso contrário, seja $c \in f(D)$. Então existe $d \in D$ tal que $f(d) = c \in C$.
3. Se $f\left(\bigcup_{i\in I} C_i\right) = \emptyset$, então $\bigcup_{i\in I} C_i = \emptyset$. Assim, segue que, para todo $i \in I$, $C_i = \emptyset$ e temos que $f(C_i) = \emptyset$. Portanto $\bigcup_{i\in I} f(C_i) = \emptyset$. Caso contrário, seja $d \in f\left(\bigcup_{i\in I} C_i\right)$. Então existe $c \in \bigcup_{i\in I} C_i$ tal que $f(c) = d$ e, consequentemente, existe $i \in I$ tal que $c \in C_i$. Assim, segue que $d = f(c) \in f(C_i) \subseteq \bigcup_{i\in I} f(C_i)$. Reciprocamente, se $\bigcup_{i\in I} f(C_i) = \emptyset$, então, para todo $i \in I$, $f(C_i) = \emptyset$, o que implica $C_i = \emptyset$. Assim, segue que $\bigcup_{i\in I} C_i = \emptyset$ e, portanto, $f\left(\bigcup_{i\in I} C_i\right) = \emptyset$. Caso contrário, seja $d \in \bigcup_{i\in I} f(C_i)$. Então existe $i \in I$ tal que $d \in f(C_i)$ e, consequentemente, existe $c \in C_i$ tal que $f(c) = d$. Assim, segue que $c \in \bigcup_{i\in I} C_i$ e, portanto, que $d \in f\left(\bigcup_{i\in I} C_i\right)$.

■

⊢ **Proposição 3.15.** *Sejam $f : D \to C$ uma função, $X \subseteq D$ e $Y \subseteq C$. Então*

1. $X \subseteq f^{-1}(f(X))$.
2. $X = f^{-1}(f(X))$ *se f é injetiva.*
3. $f(f^{-1}(Y)) \subseteq Y$.
4. $f(f^{-1}(Y)) = Y$ *se f é sobrejetiva.*

□ *Demonstração.*

1. Seja $x \in X$. Então $f(x) \in f(X)$, o que implica que $x \in f^{-1}(f(X))$.
2. Seja $x \in f^{-1}(f(X))$. Então $f(x) \in f(X)$. Portanto existe $x' \in X$ tal que $f(x) = f(x')$. Da injetividade, segue que $x = x' \in X$.
3. Seja $y \in f(f^{-1}(Y))$. Então existe $x \in f^{-1}(Y)$ tal que $f(x) = y$. Mas então $f(x) \in Y$, portanto $y \in Y$.
4. Seja $y \in Y$. Da sobrejetividade, existe $x \in X$ tal que $f(x) = y \in Y$. Isso implica que $x \in f^{-1}(Y)$ e, portanto, $y = f(x) = f(f^{-1}(Y))$.

■

3.1.6 Os funtores imagem e imagem inversa

Podemos entender a imagem e a imagem inversa de funções como funções definidas no conjunto das partes do domínio de contradomínio da função, como a seguir.

Seja $f\colon C \to C'$ uma função. A função imagem de f é a função

$$\begin{aligned} \wp_\star(f)\colon \wp(C) &\longrightarrow \wp(C') \\ A &\longmapsto f(A) = \{f(a) \mid a \in A\} \end{aligned}$$

e a função imagem inversa de f é a função

$$\begin{aligned} \wp^\star(f)\colon \wp(C') &\longrightarrow \wp(C) \\ B &\longmapsto f^{-1}(B) = \{c \in C \mid f(c) \in B\}\,. \end{aligned}$$

Sendo assim, elas satisfazem as seguintes propriedades funtoriais.

⊢ **Proposição 3.16** (Imagem é funtor covariante)**.** *1. Para todo conjunto C,*

$$\wp_\star(\mathrm{I}_C) = \mathrm{I}_{\wp(C)};$$

2. Para todos conjuntos C, C', C'' e todas funções $f\colon C \to C'$ e $f'\colon C' \to C''$,

$$\wp_\star(f' \circ f) = \wp_\star(f') \circ \wp_\star(f).$$

□ *Demonstração.* 1. Para todo $A \in \wp(C)$,

$$\wp_\star(\mathrm{I}_c)(A) = \{\mathrm{I}_C(a) \mid a \in A\} = \{a \mid a \in A\} = A,$$

portanto $\wp_\star(\mathrm{I}_C) = \mathrm{I}_{\wp(C)}$.

2. Para todo $A \in \wp(C)$,

$$\begin{aligned} \wp_\star(f' \circ f)(A) &= \{f' \circ f(a) \mid a \in A\} \\ &= \{f'(f(a)) \mid a \in A\} \\ &= \wp_\star(f')(\{f(a) \mid a \in A\}) \\ &= \wp_\star(f')(\wp_\star(f)(A)) \\ &= (\wp_\star(f') \circ \wp_\star(f))(A), \end{aligned}$$

portanto $\wp_\star(f' \circ f) = \wp_\star(f') \circ \wp_\star(f)$.

■

⊢ **Proposição 3.17** (Imagem inversa é funtor contravariante)**.** *Valem as seguintes propriedades:*

1. *Para todo conjunto C,*
$$\wp_\star(\mathrm{I}_C) = \mathrm{I}_{\wp(C)};$$
2. *Para todos conjuntos C, C', C'' e todas funções $f\colon C \to C'$ e $f'\colon C' \to C''$,*
$$\wp_\star(f' \circ f) = \wp_\star(f) \circ \wp_\star(f').$$

□ *Demonstração.* 1. Para todo $A \in \wp(C)$,
$$\wp^\star(\mathrm{I}_c)(A) = \{c \in C \mid \mathrm{I}_C(c) \in A\} = \{c \in C \mid c \in A\} = A,$$
portanto $\wp^\star(\mathrm{I}_C) = \mathrm{I}_{\wp(C)}$.

2. Para todo $A \in \wp(C'')$,
$$\begin{aligned}\wp^\star(f' \circ f)(A) &= \{c \in C \mid f' \circ f(c) \in A\}\\ &= \{c \in C \mid f'(f(c)) \in A\}\\ &= \{c \in C \mid f(c) \in \wp^\star(f')(A)\}\\ &= \{c \in C \mid c \in \wp^\star(f)(\wp^\star(f')(A))\}\\ &= \{c \in C \mid c \in \wp^\star(f) \circ \wp^\star(f')(A)\}\\ &= \wp^\star(f) \circ \wp^\star(f')(A),\end{aligned}$$
portanto $\wp_\star(f' \circ f) = \wp_\star(f') \circ \wp_\star(f)$.

■

⊢ **Proposição 3.18.** *Seja $f\colon C \to C'$ uma função.*

1. $\mathrm{I}_{\wp(C)} \subseteq \wp^\star \circ \wp_\star(f)$;
2. $\mathrm{I}_{\wp(C)} = \wp^\star \circ \wp_\star(f)$ *se f é injetiva;*
3. $\wp_\star \circ \wp^\star(f) \subseteq \mathrm{I}_{\wp(C')}$;
4. $\wp_\star \circ \wp^\star(f) = \mathrm{I}_{\wp(C')}$ *se f é sobrejetiva.*

3.2 Equivalências

:⊢ **Definição 3.15.** Seja A um conjunto. Uma *equivalência* em A é uma relação binária $\sim$ em A que é reflexiva, simétrica e transitiva.

Costumamos denotar uma relação de equivalência com símbolos $\sim, \simeq, \approx, \equiv$ ou outros símbolos semelhantes.

:⊢ **Definição 3.16.** Seja A um conjunto e $\sim$ uma relação de equivalência em A. A *classe de equivalência* de $a \in A$ é o conjunto
$$[\![a]\!] := \{b \in A \mid b \sim a\}.$$
O *conjunto quociente* de A por $\sim$ é o conjunto
$$A/\!\sim\; := \{[\![a]\!] \mid a \in A\}.$$

$\vdash$ **Teorema 3.19** (Teorema Fundamental das Relações de Equivalência)**.** *Seja A um conjunto. Se $\sim$ é uma relação de equivalência em A, então $A/\!\sim$ é uma partição de A. Reciprocamente, se $\mathcal{P}$ é uma partição de A, então existe uma relação de equivalência $\sim$ em A tal que $\mathcal{P} = A/\!\sim$.*

$\square$ *Demonstração.* Seja $\sim$ uma relação de equivalência em A e $\mathcal{P} := A/\!\sim$. Claramente, $\emptyset \notin \mathcal{P}$. Ainda, para todo $a \in A$, como $a \sim a$, então $a \in [\![a]\!]$. Logo

$$\bigcup_{[\![a]\!] \in \mathcal{P}} [\![a]\!] = A.$$

Por fim, sejam $[\![a_1]\!], [\![a_2]\!] \in \mathcal{P}$ tais que $[\![a_1]\!] \neq [\![a_2]\!]$. Se existir $a \in [\![a_1]\!] \cap [\![a_2]\!]$, então, para todo $b \in [\![a_1]\!]$, $b \sim a_1$ e $a_1 \sim a$, o que implica $b \sim a$. Ainda, $a \sim a_2$. Então $b \in [\![a_2]\!]$; ou seja, $[\![a_1]\!] \subseteq [\![a_2]\!]$. Por outro lado, $b \sim a_2 \sim a \sim a_1$, o que implica $[\![a_2]\!] \subseteq [\![a_1]\!]$. Isso implica $[\![a_1]\!] = [\![a_2]\!]$, contradição. Logo $[\![a_1]\!] \cap [\![a_2]\!] = \emptyset$. Assim, concluímos que $\mathcal{P}$ é uma partição de A.

Seja $\mathcal{P}$ uma partição de A. A relação binária $\sim$ em A, definida por

$$\forall a_1, a_2 \in A \qquad a_1 \sim a_2 \Leftrightarrow \exists Q \in \mathcal{P} \quad a_1, a_2 \in Q,$$

é uma relação de equivalência. Claramente, para todo $a \in A$, existe $Q \in \mathcal{P}$ tal que $a \in Q$, pois $\bigcup_{R \in \mathcal{P}} R = A$. Então $a \sim a$, o que mostra a reflexividade. Ainda, a simetria é trivial pela definição da relação $\sim$. Por fim, para $a_1, a_2, a_3 \in A$, se $a_1 \sim a_2$ e $a_2 \sim a_3$, existem conjuntos $Q, R \in \mathcal{P}$ tais que $a_1, a_2 \in Q$ e $a_2, a_3 \in R$. Como $a_2 \in Q \cap R$, pela definição de partição $Q = R$. Então $a_1 \sim a_3$, o que mostra a transitividade. Logo $\sim$ é uma relação de equivalência em A. ■

3.2.1 Funções bem definidas

$:\!\vdash$ **Definição 3.17.** Sejam $(E, \sim)$ e $(E', \sim')$ conjuntos com equivalência. Uma *função bem definida* (ou *função que preserva equivalência*) de $(E, \sim)$ para $(E', \sim')$ é uma função $f\colon E \longrightarrow E'$ tal que, para todos $x_0, x_1 \in E$ tais que $x_0 \sim x_1$,

$$f(x_0) \sim' f(x_1).$$

$\vdash$ **Proposição 3.20.** *Sejam $(E, \sim)$ e $(E', \sim')$ conjuntos com equivalência e $f\colon E \longrightarrow E'$ uma função. A função f é bem definida se, e somente se, a relação*

$$[\![f]\!] := \left\{ ([\![x]\!], [\![x']\!]') \in {}^{E}\!/_{\sim} \times {}^{E'}\!/_{\sim'} \;\middle|\; \exists_{x_0 \in [\![x]\!]} f(x_0) \in [\![x']\!]' \right\}$$

é uma função.

□ *Demonstração.* ($\Rightarrow$) Suponhamos que f é bem definida e seja $[\![x]\!] \in E/\sim$. Para todo $x_0 \in [\![x]\!]$, $x_0 \sim x$, portanto $f(x_0) \sim' f(x)$, o que mostra que $[\![f(x_0)]\!]' = [\![f(x)]\!]'$, portanto existe única classe $[\![f(x)]\!]' \in E'/\sim'$ tal que $([\![x]\!], [\![f(x)']\!]) \in [\![f]\!]$.

($\Leftarrow$) Suponhamos que $[\![f]\!]$ é função e sejam $x_0, x_1 \in E$ tais que $x_0 \sim x_1$, temos que $[\![x_0]\!] = [\![x_1]\!]$, portanto

$$[\![f(x_0)]\!]' = [\![f]\!]([\![x_0]\!]) = [\![f]\!]([\![x_1]\!]) = [\![f(x_1)]\!]',$$

o que mostra que $f(x_0) \sim' f(x_1)$.

■

Nesse caso, temos a função

$$\begin{aligned} [\![f]\!] \colon E\big/\!\sim &\longrightarrow E'\big/\!\sim' \\ [\![x]\!] &\longmapsto [\![f(x)]\!]'. \end{aligned}$$

Em geral, a função $[\![f]\!]$ também é denotada f por simplicidade. Comumente, considera-se somente um espaço com equivalência, e a equivalência no segundo conjunto é tomada como a igualdade.

3.3 Ordens

3.3.1 Ordens parciais, estritas e totais

:⊢ **Definição 3.18.** Seja X um conjunto. Uma *ordem parcial* em X é uma relação binária $\leq$ em X que é reflexiva, antissimétrica e transitiva. Uma *ordem total* é uma ordem parcial que é total.

Costumamos denotar uma relação de ordem com símbolos $\leq, \subseteq, \trianglelefteq$ ou outros símbolos semelhantes.

► **Exemplo 3.1.** *Seja A um conjunto. Então a relação $\subseteq$ entre elementos de $\wp(A)$ é uma relação de ordem parcial em $\wp(A)$.*

► **Exemplo 3.2.** *Seja $\mathbb{N}$ o conjunto dos naturais. Então a relação divide $|$, definida por*

$$a|b \Leftrightarrow \exists n \in \mathbb{N} \qquad an = b$$

é uma relação de ordem parcial nos naturais.

⊢ **Proposição 3.21.** *Seja X um conjunto e $\leq$ uma ordem parcial em X. Então a relação binária $\geq$ em X, definida para todos $x_1, x_2 \in X$ por*

$$x_1 \geq x_2 \Leftrightarrow x_2 \leq x_1,$$

é uma ordem parcial em X.

□ *Demonstração.* Vamos mostrar que valem as três propriedades de ordem parcial. Sejam $x_1, x_2, x_2 \in X$. Como $x_1 \leq x_1$, então $x_1 \geq x_1$. Agora suponha que $x_1 \geq x_2$. Por definição, temos $x_2 \leq x_1$, o que implica $x_1 \leq x_2$, que por sua vez implica $x_2 \geq x_1$. Por fim, suponha $x_1 \geq x_2$ e $x_2 \geq x_3$. Então $x_2 \leq x_1$ e $x_3 \leq x_2$, o que implica $x_3 \leq x_1$ e, portanto, $x_1 \geq x_3$. ■

:⊢ **Definição 3.19.** Seja X um conjunto e $\leq$ uma ordem parcial em X. A *ordem dual de* $\leq$ é a ordem parcial $\geq$ em X, definida para todos $x_1, x_2 \in X$ por

$$x_1 \geq x_2 \Leftrightarrow x_2 \leq x_1.$$

O conceito de dualidade é um conceito importante na teoria de ordem. De fato, toda definição ou teorema tem uma definição ou teorema dual, que consiste em trocar a ordem parcial $\leq$ por sua ordem dual $\geq$.

:⊢ **Definição 3.20.** Seja X um conjunto. Uma *ordem parcial estrita* em X é uma relação binária $<$ em X que é irreflexiva e transitiva. Uma *ordem total estrita* é uma ordem estrita que é total.

Costumamos denotar uma relação de ordem parcial estrita com símbolos $<, \prec, \subset, \lhd$ ou outros símbolos semelhantes.

► **Exemplo 3.3.** *Seja A um conjunto. Então a relação $\subset$ entre elementos de $\wp(A)$ é uma relação de ordem estrita em $\wp(A)$.*

⊢ **Proposição 3.22.** *Seja X um conjunto não vazio e $\leq$ uma ordem parcial em X. Então a relação binária $<$ em X, definida para todos $x_1, x_2 \in X$ por*

$$x_1 < x_2 \Leftrightarrow x_1 \leq x_2 \quad e \quad x_1 \neq x_2,$$

é uma ordem estrita em X.

□ *Demonstração.* Sejam $x_1, x_2, x_3 \in X$. Claramente, $<$ é irreflexiva por definição pois, se $x_1 < x_2$, então $x_1 \neq x_2$. Consideremos agora a transitividade de $<$. Se $x_1 < x_2$ e $x_2 < x_3$, então $x_1 \leq x_2$ e $x_2 \leq x_3$, e também $x_1 \neq x_2$ e $x_2 \neq x_3$. Pela transitividade de $\leq$, temos $x_1 \leq x_3$. Ainda, $x_1 = x_3$ implica $x_1 \leq x_2$ e $x_2 \leq x_1$ e, da antissimetria de $\leq$, temos $x_1 = x_2$, absurdo. Concluímos que $x_1 \neq x_3$ e, portanto, $x_1 < x_3$. ■

:⊢ **Definição 3.21.** Seja X um conjunto não vazio e $\leq$ uma ordem parcial em X. A *ordem estrita associada a* $\leq$ é a ordem estrita $<$ em X, definida para todos $x_1, x_2 \in X$ por

$$x_1 < x_2 \Leftrightarrow x_1 \leq x_2 \ \text{ e } \ x_1 \neq x_2.$$

$\vdash$ **Proposição 3.23.** *Seja X um conjunto não vazio e $<$ uma ordem estrita em X. Então a relação binária $\leq$ em X, definida para todos $x_1, x_2 \in X$ por*

$$x_1 \leq x_2 \Leftrightarrow x_1 < x_2 \quad ou \quad x_1 = x_2,$$

é uma ordem parcial em X.

$\square$ *Demonstração.* A demonstração é análoga à demonstração da proposição anterior. ■

$:\vdash$ **Definição 3.22.** Seja X um conjunto não vazio e $<$ uma ordem estrita em X. A *ordem parcial associada a* $<$ é a ordem parcial $<$ em X, definida para todos $x_1, x_2 \in X$ por

$$x_1 \leq x_2 \Leftrightarrow x_1 < x_2 \quad \text{ou} \quad x_1 = x_2.$$

3.3.2 Conjuntos parcialmente ordenados

$:\vdash$ **Definição 3.23.** Um *conjunto parcialmente ordenado* é um par $(X, \leq)$ em que X é um conjunto e $\leq$ é uma ordem parcial em X. Um *conjunto parcialmente ordenado estrito* é um par $(X, <)$ em que X é um conjunto e $<$ é uma ordem parcial estrita em X.

$:\vdash$ **Definição 3.24** (Maior e menor elementos)**.** Sejam $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$. Um *maior elemento* de Y é um elemento $m \in Y$ que satisfaz

$$\forall y \in Y \qquad y \leq m.$$

Dualmente, um *menor elemento* de Y é um elemento $m \in Y$ que satisfaz

$$\forall y \in Y \qquad m \leq y.$$

$\vdash$ **Proposição 3.24.** *Seja $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$. Se existe maior elemento de Y, ele é único. Dualmente, se existe menor elemento de Y, ele é único.*

$\square$ *Demonstração.* Seja m um maior elemento de Y. Então, se $n \in Y$ é um maior elemento de Y, então $m \leq n$. Mas, como m é um maior elemento de Y, então $n \leq m$ e, como $\leq$ é antissimétrica, $m = n$. A mesma demonstração vale para um menor elemento de Y, considerando a ordem parcial $\geq$, dual de $\leq$. ■

Notação. Sejam $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$. Se existirem, o maior e menor elementos de Y são denotados $\max Y$ e $\min Y$, respectivamente.

$\vdash$ **Proposição 3.25.** *Seja $(X, \leq)$ um conjunto parcialmente ordenado. Então*

1. *$\emptyset$ não tem maior nem menor elemento.*
2. $\forall x \in X \qquad \min\{x\} = \max\{x\} = x.$

$\vdash$ **Proposição 3.26.** *Sejam $(X, \leq)$ um conjunto parcialmente ordenado e Y e Z conjuntos tais que $Z \subseteq Y \subseteq X$. Então, se Y e Z têm maior elemento,*

$$\max Y = \max(\{\max Z\} \cup (Y \smallsetminus Z)).$$

Dualmente, se Y e Z têm menor elemento,

$$\min Y = \min(\{\min Z\} \cup (Y \smallsetminus Z)).$$

$\square$ *Demonstração.* Vamos mostrar que

$$\max Y \in \{\max Z\} \cup (Y \smallsetminus Z).$$

Como $\max Y \in Y$, $\max Y \notin (Y \smallsetminus Z)$ implica que $\max Y \in Z$. Portanto $\max Y \leq \max Z$; por outro lado, como $Z \subseteq Y$, então $\max Z \leq \max Y$, o que implica $\max Y = \max Z$ e, assim, concluímos que $\max Y \in \{\max Z\} \cup (Y \smallsetminus Z)$. Agora vamos mostrar que $\{\max Z\} \cup (Y \smallsetminus Z)$ tem maior elemento $\max Y$. Seja $y \in \{\max Z\} \cup (Y \smallsetminus Z)$. Se $y = \max Z$, como $Z \subseteq Y$, então $y \leq \max Y$. Se $y \in (Y \smallsetminus Z)$, como $(Y \smallsetminus Z) \in Y$, então $y \leq \max Y$. Portanto $\max Y = \max(\{\max Z\} \cup (Y \smallsetminus Z))$. ■

$:\vdash$ **Definição 3.25** (Elementos maximal e minimal)**.** Seja $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$ um conjunto não vazio. Um *elemento maximal* de Y é um elemento $m \in Y$ que satisfaz

$$\nexists y \in Y \qquad m < y.$$

Dualmente, um *elemento minimal* de Y é um elemento $m \in Y$ que satisfaz

$$\nexists y \in Y \qquad y < m.$$

$\vdash$ **Proposição 3.27.** *Seja $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$ um conjunto não vazio. Se Y tem maior elemento, então ele é o único elemento maximal de Y. Dualmente, se Y tem menor elemento, então ele é o único elemento minimal de Y.*

$\square$ *Demonstração.* Se Y tem maior elemento, então, para todo $y \in Y$, vale $y \leq \max Y$. Como $\max Y$ é único, não existe elemento $y \in Y$ tal que $y \neq \max$ e $\max Y \leq y$. Portanto $\max Y$ é um elemento maximal de Y. Agora, se existisse outro elemento maximal m de Y, teríamos $m \leq \max Y$, pois $\max Y$ é o maior elemento de Y, o que contradiz a maximalidade de m. Logo $\max Y$ é o único elemento maximal de Y. ■

:⊢ **Definição 3.26** (Limitantes superior e inferior)**.** Seja $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$. Um *limitante superior* de Y é um elemento $l \in X$ que satisfaz

$$\forall y \in Y \qquad y \leq l.$$

Dualmente, um *limitante inferior* de Y é um elemento $l \in X$ que satisfaz

$$\forall y \in Y \qquad l \leq y.$$

Um conjunto *limitado por cima* é um conjunto que possui limitante superior. Um conjunto *limitado por baixo* é um conjunto que possui limitante inferior. Um conjunto *limitado* é um conjunto limitado por cima e por baixo.

⊢ **Proposição 3.28.** *Sejam $(X, \leq)$ um conjunto parcialmente ordenado e Y e Z conjuntos tais que $Z \subseteq Y \subseteq X$. Então, se L_Z é o conjunto dos limitantes superiores de Z*

...

:⊢ **Definição 3.27** (Supremo e ínfimo)**.** Seja $(X, \leq)$ um conjunto parcialmente ordenado e $Y \subseteq X$. O *supremo* de Y, denotado $\sup Y$, é o menor elemento do conjunto de limitantes superiores de Y. Dualmente, o *ínfimo* de Y, denotado $\inf Y$, é o maior elemento do conjunto de limitantes inferiores de Y.

⊢ **Proposição 3.29.** *Sejam $(X, \leq)$ um conjunto parcialmente ordenado e Y e Z conjuntos tais que $Z \subseteq Y \subseteq X$. Então, se Y e Z têm supremo,*

$$\sup Y = \sup(\{\sup Z\} \cup (Y \smallsetminus Z)).$$

Dualmente, se Y e Z têm ínfimo,

$$\inf Y = \inf(\{\inf Z\} \cup (Y \smallsetminus Z)).$$

□ *Demonstração.* Seja $y \in \{\sup Z\} \cup (Y \smallsetminus Z)$. Se $y = \sup Z$, como $Z \subseteq Y$, então $\sup Z \leq \sup Y$; ■

3.3.3 Funções monótonas

:⊢ **Definição 3.28.** Sejam $\boldsymbol{X} = (X, \leq)$ e $\boldsymbol{X}' = (X', \leq')$ conjuntos parcialmente ordenados. Uma *função monótona* de $\boldsymbol{X}$ para $\boldsymbol{X}'$ é uma função $f\colon X \longrightarrow X'$ tal que, para todos $x_0, x_1 \in X$ tais que $x_0 \leq x_1$,

$$f(x_0) \leq' f(x_1).$$

Denota-se $f\colon \boldsymbol{X} \longrightarrow \boldsymbol{X}'$. O conjunto dessas funções é denotado $\mathcal{F}_{\leq}(\boldsymbol{X}, \boldsymbol{X}')$.

Uma *função monótona estrita* de $\boldsymbol{X}$ para $\boldsymbol{X}'$ é uma função $f\colon X \longrightarrow X'$ tal que, para todos $x_0, x_1 \in X$ tais que $x_0 < x_1$,

$$f(x_0) <' f(x_1).$$

▷ **Exercício 3.1.** *Sejam $\boldsymbol{X} = (X, \leq)$ e $\boldsymbol{X}' = (X', \leq')$ conjuntos parcialmente ordenados e $f\colon X \longrightarrow X'$ uma função. A função f é monótona e injetiva se, e somente se, é monótona estrita.*

□ *Demonstração.* ($\Rightarrow$) Suponhamos que f é monótona e injetiva e sejam $x_0, x_1 \in X$ tais que $x_0 < x_1$. Como $x_0 \leq x_1$, segue da monotonicidade que $f(x_0) \leq f(x_1)$. Como $x_0 \neq x_1$, segue da injetividade que $f(x_0) \neq f(x_1)$, portanto $f(x_0) < f(x_1)$.

($\Leftarrow$) Suponhamos que f é monótona estrita e sejam $x_0, x_1 \in X$ tais que $x_0 \leq x_1$. Se $x_0 = x_1$, então $f(x_0) = f(x_1)$, logo $f(x_0) \leq f(x_1)$. Caso contrário, $x_0 < x_1$, e segue da monotonicidade estrita que $f(x_0) < f(x_1)$. logo $f(x_0) \leq f(x_1)$, o que mostra que f é monótona.

■

:⊢ **Definição 3.29.** Sejam $\boldsymbol{X} = (X, \leq)$ e $\boldsymbol{X}' = (X', \leq')$ conjuntos parcialmente ordenados. Um *isomorfismo de ordem* (ou *isomorfismo monótono*) de $\boldsymbol{X}$ para $\boldsymbol{X}'$ é uma função monótona $f\colon \boldsymbol{X} \longrightarrow \boldsymbol{X}'$ invertível. O conjunto dessas funções é denotado $\overleftrightarrow{\mathcal{F}}_{\leq}(\boldsymbol{X}, \boldsymbol{X}')$. Nesse caso, os conjuntos $\boldsymbol{X}$ e $\boldsymbol{X}'$ são *isomorfos* e denota-se $\boldsymbol{X} \simeq \boldsymbol{X}'$.

⊢ **Proposição 3.30.** *Sejam $\boldsymbol{X} = (X, \leq)$ e $\boldsymbol{X}' = (X', \leq')$ conjuntos parcialmente ordenados e $f\colon \boldsymbol{X} \longrightarrow \boldsymbol{X}'$ um isomorfismo de ordem de $\boldsymbol{X}$ para $\boldsymbol{X}'$. A função inversa $f^{-1}\colon X' \longrightarrow X$ é uma função monótona de $\boldsymbol{X}'$ para $\boldsymbol{X}$.*

□ *Demonstração.* Sejam $x'_0, x'_1 \in X$ tais que $x'_0 \leq x'_1$. Suponhamos, por absurdo, que $f^{-1}(x'_0) > f^{-1}(x'_1)$. Da monotonicidade de f seguiria que $x'_0 = f(f^{-1}(x'_0)) > f(f^{-1}(x'_1)) = x'_1$, contradição. ■

3.3.3.1 Multi-índices estritamente crescentes

Definimos

$$[d]^{\uparrow k} := \left\{ (i_0, \ldots, i_{k-1}) \in [d]^k \;\middle|\; i_0 < \cdots < i_{k-1} \right\}.$$

Note que

$$\#\left([d]^{\uparrow k}\right) = \#\left(\binom{[d]}{k}\right) = \binom{d}{k},$$

em que $\binom{[d]}{k} = \{I \subseteq [d] \mid \#(I) = k\}$.

3.3.4 Conjuntos totalmente ordenados e cadeias

:⊢ **Definição 3.30.** Um *conjunto totalmente ordenado* é um conjunto parcialmente ordenado $(X, \leq)$ tal que $\leq$ é uma ordem total: para todos $x, x' \in X$,

$$x \leq x' \quad \text{ou} \quad x' \leq x.$$

Um *conjunto totalmente ordenado estrito* é um conjunto parcialmente ordenado estrito $(X, <)$ tal que $<$ é uma ordem total estrita.

:⊢ **Definição 3.31.** Seja $(X, \leq)$ um conjunto parcialmente ordenado. Uma *cadeia* de X é um conjunto $Y \subseteq X$ que satisfaz

$$\forall y_1, y_2 \in Y \qquad y_1 \leq y_2 \ \text{ ou } \ y_2 \leq y_1.$$

▷ **Exercício 3.2.** *Seja $(X, \leq)$ um conjunto totalmente ordenado e $Y \subseteq X$ um conjunto não vazio. Então Y é uma cadeia de X.*

:⊢ **Definição 3.32.** Seja $(X, \leq)$ um conjunto totalmente ordenado. Um *intervalo* de X é um conjunto $I \subseteq X$ tal que, para todos $i, i' \in I$ e $x \in X$ tal que $i \leq x \leq i'$, vale $x \in I$. Para $e, e' \in X$, definimos:

O *intervalo aberto* de extremos e e e' é o conjunto

$$]e, e'[:= \{x \in X \mid e < x < e'\}.$$

O *intervalo semi-aberto inferiormente* de extremos e e e' é o conjunto

$$]e, e'] := \{x \in X \mid e < x \leq e'\}.$$

O *intervalo semi-aberto superiormente* de extremos e e e' é o conjunto

$$[e, e'[:= \{x \in X \mid e \leq x < e'\}.$$

O *intervalo fechado* de extremos e e e' é o conjunto

$$[e, e'] := \{x \in X \mid e \leq x \leq e'\}.$$

A *semirreta fechada superior* com extremo e é o conjunto

$$[e, \infty[:= \{e \in X \mid e \leq x\}.$$

A *semirreta aberta superior* com extremo e é o conjunto

$$]e, \infty[:= \{e \in X \mid e < x\}.$$

A *semirreta fechada inferior* com extremo e é o conjunto

$$]-\infty, e] := \{e \in X \mid x \leq e\}.$$

A *semirreta aberta inferior* com extremo e é o conjunto

$$]-\infty, e[:= \{e \in X \mid x < e\}.$$

Claramente, supomos que ∞ e $-\infty$ não são símbolos usados para representar um elemento de X, de modo a não gerar confusão.

Aqui citamos o resultado conhecido como lema de Zorn, mas não apresentamos uma demonstração.

⊢ **Lema 3.31** (Lema de Zorn)**.** *Seja $(X, \leq)$ um conjunto parcialmente ordenado. Se toda cadeia de X possui limitante superior, então X tem elemento maximal.*

3.3.5 Conjuntos bem ordenados

:⊢ **Definição 3.33.** Um *conjunto bem ordenado* é um conjunto totalmente ordenado $(X, \leq)$ tal que todo conjunto não vazio $C \subseteq X$ tem elemento mínimo $\min C$. O *zero* de um conjunto bem ordenado não vazio é o elemento $0_X := \min X$. Um *conjunto bem ordenado estrito* é um conjunto totalmente ordenado estrito $(X, <)$ tal que $(X, \leq)$ é um conjunto bem ordenado.

⊢ **Proposição 3.32.** *Sejam $(X, \leq)$ um conjunto bem ordenado e $f\colon \boldsymbol{X} \longrightarrow \boldsymbol{X}$ uma função monótona.*

1. *Se f é injetiva, para todo $x \in X$ vale $x \leq f(x)$;*
2. *Se f é isomorfismo, $f = \mathrm{I}$.*

□ *Demonstração.*

1. Consideremos o conjunto
$$C := \{x \in X \mid f(x) < x\}.$$
Suponhamos que C é não vazio e seja $m := \min C$. Porque $m \in C$, vale $f(m) < m$, o que implica $f(m) \notin C$; da monotonicidade de f segue que $f(f(m)) < f(m)$, o que mostra que $f(m) \in C$, contradição.
2. Seja $x \in X$. Como f é isomorfismo, f e f^{-1} são injetivas, logo $x \leq f(x)$ e $f^{-1}(x) \leq x$; como f é monótona, segue que $x = f(f^{-1}(x)) \leq f(x)$, portanto $f(x) = \mathrm{I}$.

■

▷ **Exercício 3.3.** *Sejam $(X, \leq)$ e $(X', \leq')$ conjuntos bem ordenados isomorfos. Existe único isomorfismo de ordem $f\colon \boldsymbol{X} \longrightarrow \boldsymbol{X}$.*

:⊢ **Definição 3.34.** Sejam $(X, \leq)$ um conjunto bem ordenado e $e \in X$. O *segmento inicial* dado por e é o conjunto
$$[e] := [0_X, e[= \{x \in X \mid x < e\}.$$

⊢ **Proposição 3.33.** *Sejam $(X, \leq)$ um conjunto bem ordenado e $e \in X$. Não existe isomorfismo de ordem $f\colon X \longrightarrow [e]$.*

□ *Demonstração.* Se existisse isomorfismo $f\colon X \longrightarrow [e]$, teríamos $f(X) = [e] = \{x \in X \mid x < e\}$, o que implicaria $f(e) < e$; mas como f seria injetiva, teríamos $e \leq f(e)$, contradição. ■

▷ **Exercício 3.4** (Tricotomia)**.** *Sejam $(X, \leq)$ e $(X', \leq')$ conjuntos bem ordenados. Exatamente um dos três vale:*

1. *$(X, \leq)$ é isomorfo a $(X', \leq')$;*
2. *$(X, \leq)$ é isomorfo a um segmento inicial de $(X', \leq')$;*
3. *$(X', \leq')$ é isomorfo a um segmento inicial de $(X, \leq)$.*

3.3.5.1 Números ordinais

:⊢ **Definição 3.35.** Um conjunto *transitivo* é um conjunto T tal que, para todo $t \in T$, $t \subseteq T$.

:⊢ **Definição 3.36.** Um *número ordinal* é um conjunto transitivo O tal que $(O, \in)$ é um conjunto bem ordenado.

3.3.6 Pré-ordens

:⊢ **Definição 3.37.** Seja X um conjunto não vazio. Uma *pré-ordem* (ou *precedência*) em X é uma relação binária em X que é reflexiva e transitiva. O par $(X, \preceq)$ é um *conjunto pré-ordenado.*

:⊢ **Definição 3.38.** Seja $(X, \preceq)$ um conjunto pré-ordenado. A *equivalência induzida* por $\preceq$ é a relação binária $\sim$ definida por: para todos $x, x' \in X$,

$$x \sim x' \quad \Longleftrightarrow \quad x \preceq x' \text{ e } x' \preceq x.$$

A *ordenação induzida* por $\leq$ é a relação binária em $X/\sim$ definida por: para todos $x, x' \in X$,

$$[x] \leq [x'] \quad \Longleftrightarrow \quad x \preceq x'.$$

⊢ **Proposição 3.34.** *Seja $(X, \preceq)$ um conjunto pré-ordenado. A relação $\sim$ em A é uma equivalência em X e a relação $\leq$ em $X/\sim$ é uma ordem em $X/\sim$.*

□ *Demonstração.* 1. (Equivalência $\sim$)

1.1. (Reflexividade) Para todo $x \in X$, vale que $x \preceq x$, portanto $x \sim x$.

1.2. (Simetria) Para todos $x, x' \in X$, se $x \preceq x'$ e $x' \preceq x$, então $x \sim x'$ por definição.

1.3. (Transitividade) Para todos $x, x', x'' \in X$, se $x \sim x'$ e $x' \sim x''$, então se $x \preceq x'$, $x' \preceq x$, $x' \preceq x''$ e $x'' \preceq x'$, o que implica pela transitividade de $\preceq$ que $x \preceq x''$ e $x'' \preceq x$, portanto $x \sim x''$.

2. (Ordem $\leq$) Primeiro devemos mostrar que a relação está bem definida. Sejam $[x], [x'] \in X$. Tomemos $x, y \in [x]$ e $x', y' \in [x']$; queremos mostrar que se $x \preceq x'$, então $y \preceq y'$. Como $x \sim y$, então $y \preceq x$, e como $x' \sim y'$, então $x' \preceq y'$; assim, da transitividade de $\preceq$ segue que

$$y \preceq x \preceq x' \preceq y'.$$

Isso mostra que $\leq$ está bem definida. Agora, mostremos que $\leq$ é ordem. (Reflexividade) Para todo $x \in X$, vale que $[x] \leq [x]$, pois $x \preceq x$. (Antissimetria) Para todos $x, x' \in X$, se $[x] \leq [x']$ e $[x]' \leq [x]$, então $x \preceq x'$ e $x' \preceq x$, o que

implica $x \sim x'$, portanto $[x] = [x']$. (Transitividade) Para todos $x, x', x'' \in X$, se $[x] \leq [x']$ e $[x'] \leq [x'']$, então $x \preceq x'$ e $x' \preceq x''$, o que implica que $x \preceq x''$, portanto $[x] \leq [x'']$.

■

3.3.7 Conjunto direcionado

:⊢ **Definição 3.39.** Um *conjunto direcionado* (*superiormente*) é um par $(X, \preceq)$ em que X é um conjunto não vazio e $\preceq$ é uma pré-ordem em X que satisfaz: para todos $x, x' \in X$, existe $s \in X$ tal que $x \leq s$ e $x' \leq s$.

▷ **Exercício 3.5.** *Sejam $(X, \preceq)$ um conjunto direcionado e $x_0, \ldots, x_{n-1} \in X$. Existe $s \in X$ tal que, para todo $i \in [n]$, $x_i \leq s$.*

3.3.8 Reticulados

:⊢ **Definição 3.40.** Um *reticulado* é um conjunto parcialmente ordenado $(X, \leq)$ em que, para todos $x_1, x_1 \in X$, o conjunto $\{x_1, x_2\}$ tem supremo e ínfimo, denotados, respectivamente, $x_1 \vee x_2$ e $x_1 \wedge x_2$.

⊢ **Proposição 3.35.** *Seja $(X, \leq)$ um reticulado e $Y \subseteq X$ um conjunto finito. Então Y tem supremo e ínfimo.*

Um reticulado também pode ser entendido como uma estrutura algébrica. As definições a seguir usam definições da parte de Álgebra do livro, e devem ser conferidas nessa parte.

:⊢ **Definição 3.41.** Um *reticulado* é uma tripla $(R, \vee, \wedge)$ em que

1. $(R, \vee)$ e $(R, \wedge)$ são semigrupos comutativos;
2. Valem as propriedades de *absorção*: para todos $a, b \in R$,
   2.1. $a \vee (a \wedge b) = a$;
   2.2. $a \wedge (a \vee b) = a$.

⊢ **Proposição 3.36.** *Seja $(R, \vee, \wedge)$ um reticulado. Valem as propriedades de* idempotência*: para todo $a \in R$,*

1. *$a \vee a = a$;*
2. *$a \wedge a = a$.*

:⊢ **Definição 3.42.** Um *reticulado limitado* é uma 5-sequência $(R, \vee, \wedge, 0, 1)$ em que $(R, \vee, \wedge)$ é um reticulado, 0 é elemento neutro de $(R, \vee)$ e 1 é elemento neutro de $(R, \wedge)$.

$\vee$ e $\wedge$ estão definidos para todo subconjunto não-vazio finito por indução, já que são operações associativas.

⊢ **Proposição 3.37.** *Todo reticulado finito é limitado.*

3.3.9 Álgebras booleanas

:⊢ **Definição 3.43.** Uma *álgebra booleana* é uma 5-sequência $(A, \vee, \wedge, 0, 1)$, em que A é um conjunto não vazio, que satisfaz

1. $(A, \vee, 0)$ e $(A, \wedge, 1)$ são magmas comutativos com elemento neutro;
2. As operações $\vee$ e $\wedge$ são distributivas uma sobre a outra;
3. Para todo $a \in A$ existe um elemento *complementar* $a' \in A$, que satisfaz $a \vee a' = 1$ e $a \wedge a' = 0$.

⊢ **Proposição 3.38.** *Seja A um conjunto e $\mathcal{A} \subseteq \wp(A)$ um conjunto de partes de A que satisfaz*

1. $\emptyset \in \mathcal{A}$;
2. $X \in \mathcal{A} \Rightarrow X^{\complement} \in \mathcal{A}$.

Então $(\mathcal{A}, \cup, \cap)$ é uma álgebra booleana.

□ *Demonstração.* Primeiramente, é necessário notar, embora os símbolos $\cup$ e $\cap$ não sejam funções propriamente ditas, ao fixarmos um conjunto A, podemos definir $\cup$ e $\cap$ como operações binárias em $\wp(A)$, dadas por $(X, Y) \mapsto X \cup Y$ e $(X, Y) \mapsto X \cap Y$, respectivamente. Para $X, Y \in \mathcal{A}$, temos que $X \cup Y, X \cap Y \in \mathcal{A}$, o que mostra que as operações estão bem definidas.

Sendo assim, podemos prosseguir com a demonstração. Se $\mathcal{A}$ satisfaz as propriedades do enunciado, então $A = \emptyset^{\complement} \in \mathcal{A}$. O par $(\mathcal{A}, \cup)$ é um magma comutativo com elemento neutro $\emptyset$, pois a união de dois conjuntos é comutativa por definição e a união de um conjunto qualquer com o conjunto vazio dá o próprio conjunto. Da mesma forma, o par $(\mathcal{A}, \wedge)$ é um magma comutativo com elemento neutro A, pois a interseção de dois conjuntos é comutativa por definição e a interseção de qualquer conjunto com o conjunto A é o próprio conjunto. Ainda, vale que, para todo $X, Y, Z \in \mathcal{A}$, $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$ e $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$; ou seja, as operações binárias $\cup$ e $\cap$ são distributivas uma sobre a outra. Por fim, nota-se que, dado $X \in \mathcal{A}$, $X^{\complement} \in \mathcal{A}$ e vale $X \cup X^{\complement} = A$ e $X \cap X^{\complement} = \emptyset$. Logo $(\mathcal{A}, \cup, \cap)$ é uma álgebra booleana. ■

⊢ **Proposição 3.39** (Princípio da Dualidade)**.** *Toda afirmação dedutível somente a partir da definição de álgebra booleana continua válida se são trocados entre si os símbolos $\vee$ e $\wedge$ e os símbolos 0 e 1 que aparecem na expressão.*

□ *Demonstração.* Todas as propriedades de uma álgebra booleana são definidas simetricamente e continuam iguais se trocamos entre si os símbolos $\vee$ e $\wedge$ e os símbolos 0 e 1. Logo isso também vale para qualquer afirmação dedutível dessas propriedades. ■

Como consequência do princípio da dualidade, qualquer afirmação dedutível das propriedades de álgebra booleana tem uma afirmação associada a ela ao trocarmos entre si os símbolos $\vee$ e $\wedge$ e os símbolos 0 e 1, que chamaremos que sua afirmação *dual*. Claramente, a afirmação dual da dual é a própria afirmação. Portanto só será necessário demonstrar a afirmação para demonstrar sua afirmação dual. Toda proposição, lema e teorema dessa seção exibirá sua proposição, lema e teorema dual, mas a afirmação dual não será demonstrada.

$\vdash$ **Teorema 3.40** (Identidades)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$\forall a \in A \qquad a \vee 1 = 1$$

$$\forall a \in A \qquad a \wedge 0 = 0$$

$\vdash$ **Teorema 3.41** (Absorção)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$\forall a, b \in A \qquad a \vee (a \wedge b) = a$$

$$\forall a, b \in A \qquad a \wedge (a \vee b) = a$$

$\vdash$ **Corolário 3.42** (Idempotência)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$\forall a \in A \qquad a \vee a = a$$

$$\forall a \in A \qquad a \wedge a = a$$

$\square$ *Demonstração.* Basta tomar $b = 1$ e $b = 0$ nas proposições anteriores. $\blacksquare$

$\vdash$ **Teorema 3.43** (Associatividade)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$(A, \vee) \text{ é associativo.}$$

$$(A, \wedge) \text{ é associativo.}$$

$\vdash$ **Teorema 3.44** (Unicidade do Complementar)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana e* $a \in A$*. Então o complementar de* a *é único.*

Note que esse teorema é seu próprio dual.

$\vdash$ **Teorema 3.45** (Dupla Complementação)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana e* $a \in A$*. Então o complementar de* a' *é* a*.*

$\vdash$ **Teorema 3.46** (Identidades Complementares)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$0' = 1$$

$$1' = 0$$

$\vdash$ **Teorema 3.47** (Leis de De Morgan)**.** *Seja* $(A, \vee, \wedge)$ *uma álgebra booleana. Então*

$$\forall a, b \in A \qquad (a \wedge b)' = a' \vee b'$$

$$\forall a, b \in A \qquad (a \vee b)' = a' \wedge b'$$

3.3.9.1 Função indicadora

:⊢ **Definição 3.44.** Sejam X um conjunto. A *função indicadora* em X é a função

$$\begin{aligned} \mathbf{1}\colon \wp(X) &\longrightarrow 2^X \\ C &\longmapsto \mathbf{1}_C\colon X \longrightarrow \{0,1\} \\ & \qquad\qquad x \longmapsto \begin{cases} 1, & x \in C \\ 0, & x \notin C. \end{cases} \end{aligned}$$

A função indicadora de um conjunto $C \subseteq X$ é a função $\mathbf{1}_C\colon X \longrightarrow \{0,1\}$.

A função indicadora é uma bijeção e mostra que os conjuntos $\wp(X)$ e 2^X têm a mesma cardinalidade. De fato, sabemos que $(\wp(X), \cap, \cup, \emptyset, X)$ é uma álgebra de conjuntos. Podemos também, usando a estrutura de álgebra em $\{0,1\}$, dada pelas operações mínimo e máximo $\wedge, \vee\colon \{0,1\} \times \{0,1\} \to \{0,1\}$ e pelos os elementos 0 e 1, induzir uma álgebra em 2^X com as operações definidas pontualmente e as funções constantes $0, 1 \in 2^X$. Assim, podemos mostrar que a bijeção $\mathbf{1}\colon \wp(X) \to 2^X$ é um isomorfismo de álgebras.

⊢ **Proposição 3.48.** *Seja X um conjunto. A função indicadora*

$$\mathbf{1}\colon \wp(X) \to 2^X$$

de X é um isomorfismo entre as álgebras $(\wp(X), \cap, \cup, \emptyset, X)$ e $(2^X, \wedge, \vee, 0, 1)$.

□ *Demonstração.* Para isso, devemos mostrar que $\mathbf{1}$ preserva as operações binárias e constantes das álgebras. É imediato verificar que, para todos $C, C' \in \wp(X)$, $\mathbf{1}_{C\cap C'} = \mathbf{1}_C \wedge \mathbf{1}_{C'}$, $\mathbf{1}_{C\cup C'} = \mathbf{1}_C \vee \mathbf{1}_{C'}$, e que $\mathbf{1}_\emptyset = 0$ e $\mathbf{1}_X = 1$. ■

Vale notar, também, que em $\{0,1\}$ vale que, para todos $n, n' \in \{0,1\}$, $n \wedge n' = nn'$ e $n \vee n' = n + n' - nn'$. Algumas outras relações da função indicadora estão expostas na proposição seguinte. Todas elas seguem diretamente do fato de $\mathbf{1}$ ser isomorfismo de álgebras. As demonstrações ficam como exercício.

▷ **Exercício 3.6.** *Sejam X um conjunto, $A, B \subseteq X$, $n \in \mathbb{N}$ $(A_i)_{i\in[n]} \subseteq X$. Então*

1. $\mathbf{1}_{A^\complement} = 1 - \mathbf{1}_A$;
2. $\mathbf{1}_{A\setminus B} = \mathbf{1}_A - \mathbf{1}_A\mathbf{1}_B$;
3. $\mathbf{1}_{A\triangle B} = \mathbf{1}_A + \mathbf{1}_B - 2\mathbf{1}_A\mathbf{1}_B$;
4. $\mathbf{1}_{\bigcap_{i\in[n]} A_i} = \bigtimes\limits_{i\in[n]} \mathbf{1}_{A_i}$;
5. $\mathbf{1}_{\bigcup_{i\in[n]} A_i} = \bigplus\limits_{\substack{S\subseteq[n] \\ S\neq\emptyset}} \left((-1)^{\#(S)-1} \bigtimes\limits_{i\in S} \mathbf{1}_{A_i} \right)$;
6. $\mathbf{1}_{\triangle_{i\in[n]} A_i} = \bigplus\limits_{\substack{S\subseteq[n] \\ S\neq\emptyset}} \left((-2)^{\#(S)-1} \bigtimes\limits_{i\in S} \mathbf{1}_{A_i} \right)$.

3.4 Coberturas e partições

3.4.1 Coberturas e subcoberturas

:⊢ **Definição 3.45** (Cobertura e subcobertura)**.** Seja X um conjunto. Uma *cobertura* de X é um conjunto $\mathcal{C} \subseteq \wp(X)$ tal que $X = \bigcup \mathcal{C}$. O conjunto das coberturas de X é denotado $\mathsf{C}(X)$. Uma *subcobertura* de uma cobertura $\mathcal{C}$ é uma cobertura $\mathcal{S}$ de X tal que $\mathcal{S} \subseteq \mathcal{C}$.

► **Exemplo 3.4.** *Seja X um conjunto.*

1. *O conjunto $\{X\}$ é a* cobertura trivial *de X;*
2. *O conjunto $\wp(X)$ é a* cobertura total *de X;*
3. *O conjunto*

$$\wp_1(X) = \{C \in \wp(X) \mid \#(C) = 1\} = \{\{x\} \mid x \in X\}$$

é a cobertura atômica *de X.*

Às vezes é conveniente considerar uma cobertura de um subconjunto $S \subseteq X$. Nesse caso, uma cobertura de S pode ser considerada como um conjunto $\mathcal{C} \subseteq \wp(X)$ tal que $S \subseteq \bigcup \mathcal{C}$. Não é preciso restringir a cobertura a S, isto é, considerar somente subconjuntos de S, pois para todo $C \in \mathcal{C}$, sendo $C \subseteq X$, basta considerar $C \cap S$ e têm-se $C \cap S \subseteq S$ e $S = \bigcup_{C \in \mathcal{C}} C \cap S$.

3.4.2 Refinamento de coberturas

Definiremos uma relação de pré-ordem entre coberturas que difere da relação de contenção usual entre conjuntos.

:⊢ **Definição 3.46** (Refinamento)**.** Sejam X um conjunto e $\mathcal{C}$ uma cobertura de X. Um *refinamento* de $\mathcal{C}$ é uma cobertura $\mathcal{C}'$ de X tal que, para todo $C' \in \mathcal{C}'$, algum $C \in \mathcal{C}$ satisfaz $C' \subseteq C$. A cobertura $\mathcal{C}'$ é dita *mais fina* que $\mathcal{C}$ e a cobertura $\mathcal{C}$ é dita *mais grossa* que $\mathcal{C}'$ e um *engrossamento* de $\mathcal{C}'$. Denota-se $\mathcal{C} \preceq \mathcal{C}'$ e a relação binária $\preceq$ em $\mathsf{C}(X)$ é a relação de *refinamento*.

Aqui o importante é notar que ser mais fino significa ter conjuntos menores, por isso as ordens de $\subseteq$ e $\preceq$ são contrárias.

⊢ **Proposição 3.49** (Refinamento é pré-ordem)**.** *Seja X um conjunto. A relação de refinamento $\preceq$ é uma relação de pré-ordem no conjunto de coberturas $\mathsf{C}(X)$.*

□ *Demonstração.* Devemos mostrar que a relação é reflexiva e transitiva.

- (Reflexividade) Seja $\mathcal{C} \in \mathsf{C}(X)$ uma cobertura. Para todo $C \in \mathcal{C}$, vale $C \subseteq C$, portanto $\mathcal{C} \preceq \mathcal{C}$.
- (Transitividade) Sejam $\mathcal{C}, \mathcal{C}', \mathcal{C}'' \in \mathsf{C}(X)$. Para todo $C'' \in \mathcal{C}''$, algum $C' \in \mathcal{C}'$ satisfaz $C'' \subseteq C'$, pois $\mathcal{C}' \preceq \mathcal{C}''$; além disso, algum $C \in \mathcal{C}$ satisfaz $C' \subseteq C$, pois $\mathcal{C} \preceq \mathcal{C}'$, o que implica $C'' \subseteq C' \subseteq C$, e isso mostra que $\mathcal{C} \preceq \mathcal{C}''$. ■

No entanto, não é possível mostrar que a relação de refinamento é antissimétrica. O exemplo a seguir deixa claro um dos problemas que pode ocorrer.

► **Exemplo 3.5.** *Seja $X := \{0, 1\}$, $\mathcal{C} := \{\{0\}, X\}$ e $\mathcal{C}' := \{\{1\}, X\}$. Vale $\mathcal{C} \preceq \mathcal{C}'$, pois $\{1\} \subseteq X$ e $X \subseteq X$, e vale $\mathcal{C}' \preceq \mathcal{C}$, pois $\{0\} \subseteq X$ e $X \subseteq X$. Como $\mathcal{C} \neq \mathcal{C}'$, isso mostra que a relação de refinamento não pode ser antissimétrica.*

Nas seções seguintes, estudaremos um tipo de cobertura chamado partição. Quando restrita a essas coberturas, a relação será antissimétrica, portanto uma ordem.

▷ **Exercício 3.7.** *Seja X um conjunto.*

1. *Para toda cobertura $\mathcal{C}$ de X e toda subcobertura $\mathcal{S} \subseteq \mathcal{C}$, vale $\mathcal{C} \preceq \mathcal{S}$;*
2. *Para toda cobertura $\mathcal{C}$ de X, $\{X\} \preceq \mathcal{C} \preceq \wp_1(X)$.*

3.4.3 Correfinamento de coberturas

:⊢ **Definição 3.47** (Correfinamento de coberturas)**.** Sejam X um conjunto e $\mathcal{C}$ e $\mathcal{C}'$ coberturas de X. O *correfinamento* de $\mathcal{C}$ e $\mathcal{C}'$ é o conjunto

$$\mathcal{C} \curlyvee \mathcal{C}' := \{C \cap C' \mid C \in \mathcal{C}, C' \in \mathcal{C}'\}.$$

Seja $(\mathcal{C}_i)_{i \in I}$ uma família de coberturas de X. O *correfinamento* de $(\mathcal{C}_i)_{i \in I}$ é o conjunto

$$\bigcurlyvee_{i \in I} \mathcal{C}_i := \left\{ \bigcap_{i \in I} C_i \;\middle|\; \forall_{i \in I} C_i \in \mathcal{C}_i \right\}.$$

As definições coincidem se consideramos uma família $(\mathcal{C}_0, \mathcal{C}_1)$ de duas coberturas. Nesse caso,

$$\bigcurlyvee_{i \in [2]} \mathcal{C}_i = \left\{ \bigcap_{i \in [2]} C_i \;\middle|\; \forall_{i \in [2]} C_i \in \mathcal{C}_i \right\} = \{C_0 \cap C_1 \mid C_0 \in \mathcal{C}_0, C_1 \in \mathcal{C}_1\} = \mathcal{C}_0 \curlyvee \mathcal{C}_1.$$

Algumas nomenclaturas utilizadas na literatura para essa operação entre coberturas são *soma* e *junção*, mas adotamos aqui *correfinamento* para ressaltar que esse conjunto é a menor cobertura que é mais fina que ambas as coberturas, portanto o refinamento comum delas. A proposição seguinte formaliza esse fato.

⊢ **Proposição 3.50.** *Sejam X um conjunto e $\mathcal{C}$ e $\mathcal{C}'$ coberturas de X. O correfinamento $\mathcal{C} \curlyvee \mathcal{C}'$ é a menor cobertura de X (com respeito à pré-ordem $\preceq$) que é mais fina que $\mathcal{C}$ e $\mathcal{C}'$.*

□ *Demonstração.* O conjunto $\mathcal{C} \curlyvee \mathcal{C}'$ é uma cobertura pois

$$\begin{aligned} X &= \bigcup_{C \in \mathcal{C}} C \\ &= \bigcup_{C \in \mathcal{C}} C \cap X \\ &= \bigcup_{C \in \mathcal{C}} C \cap \left(\bigcup_{C' \in \mathcal{C}'} C' \right) \\ &= \bigcup_{C \in \mathcal{C}} \bigcup_{C' \in \mathcal{C}'} C \cap C' \\ &= \bigcup_{C \cap C' \in \mathcal{C} \curlyvee \mathcal{C}'} C \cap C'. \end{aligned}$$

Mostremos agora que $\mathcal{C} \curlyvee \mathcal{C}'$ refina ambas coberturas. Seja $C \cap C' \in \mathcal{C} \curlyvee \mathcal{C}'$. Como $C \cap C' \subseteq C$ e $C \cap C' \subseteq C'$, segue que $\mathcal{C} \preceq \mathcal{C} \curlyvee \mathcal{C}'$ e $\mathcal{C}' \preceq \mathcal{C} \curlyvee \mathcal{C}'$.

Por fim, para mostrar que ele é o menor refinamento com essa propriedade, seja $\mathcal{C}'' \in \mathsf{C}(X)$ tal que $\mathcal{C} \preceq \mathcal{C}''$ e $\mathcal{C}' \preceq \mathcal{C}''$. Seja $C'' \in \mathcal{C}''$. Por refinar $\mathcal{C}$ e $\mathcal{C}'$, segue que para algum $C \in \mathcal{C}$ e algum $C' \in \mathcal{C}'$, vale $C'' \subseteq C$ e $C'' \subseteq C'$. Isso implica que $C'' \subseteq C \cap C'$, portanto $\mathcal{C} \curlyvee \mathcal{C}' \preceq \mathcal{C}''$. ■

▷ **Exercício 3.8.** *Sejam X um conjunto e $(\mathcal{C}_i)_{i \in I}$ uma família de coberturas de X. O correfinamento $\bigcurlyvee_{i \in I} \mathcal{C}_i$ é a menor cobertura de X (com respeito à pré-ordem $\preceq$) que, para todo $i \in I$, é mais fina que $\mathcal{C}_i$.*

⊢ **Proposição 3.51.** *Seja X um conjunto.*

1. *Para todas coberturas $\mathcal{C}$, $\mathcal{C}'$, $\mathcal{R}$ e $\mathcal{R}'$ de X tais que $\mathcal{C} \preceq \mathcal{R}$ e $\mathcal{C}' \preceq \mathcal{R}'$, vale*
$$\mathcal{C} \curlyvee \mathcal{C}' \preceq \mathcal{R} \curlyvee \mathcal{R}';$$
2. *Para todas coberturas $\mathcal{C}$ e $\mathcal{C}'$ de X e todas subcoberturas $\mathcal{S} \subseteq \mathcal{C}$ e $\mathcal{S}' \subseteq \mathcal{C}'$, vale*
$$\mathcal{S} \curlyvee \mathcal{S}' \subseteq \mathcal{C} \curlyvee \mathcal{C}';$$

□ *Demonstração.*
1. Seja $R \cap R' \in \mathcal{R} \curlyvee \mathcal{R}'$, com $R \in \mathcal{R}$ e $R' \in \mathcal{R}'$. Como $\mathcal{R}$ é refinamento de $\mathcal{C}$ e $\mathcal{R}'$ é refinamento de $\mathcal{C}'$, segue que, para algum $C \in \mathcal{C}$ e algum $C' \in \mathcal{C}'$, vale que $R \subseteq C$ e $R' \subseteq C'$, o que implica que $R \cap R' \subseteq C \cap C'$. Como $C \cap C' \in \mathcal{C} \curlyvee \mathcal{C}'$, segue que $\mathcal{C} \curlyvee \mathcal{C} \preceq \mathcal{R} \curlyvee \mathcal{R}'$.
2. Exercício.

■

▷ **Exercício 3.9** (Propriedades de reticulado)**.** *Seja X um conjunto.*

1. *(Associatividade) Para todas coberturas $\mathcal{C}, \mathcal{C}', \mathcal{C}'' \in \mathsf{C}(X)$,*

$$(\mathcal{C} \curlyvee \mathcal{C}') \curlyvee \mathcal{C}'' = \mathcal{C} \curlyvee (\mathcal{C}' \curlyvee \mathcal{C}'');$$

2. *(Comutatividade) Para todas coberturas $\mathcal{C}, \mathcal{C}' \in \mathsf{C}(X)$,*

$$\mathcal{C} \curlyvee \mathcal{C}' = \mathcal{C}' \curlyvee \mathcal{C};$$

3. *(Idempotência) Para toda cobertura $\mathcal{C} \in \mathsf{C}(X)$,*

$$\mathcal{C} \curlyvee \mathcal{C} = \mathcal{C};$$

4. *(Limitação inferior) Para toda cobertura $\mathcal{C} \in \mathsf{C}(X)$,*

$$\mathcal{C} \curlyvee \{X\} = \mathcal{C};$$

O correfinamento pode ser calculado repetidas vezes entre várias coberturas e a ordem em que ele é calculado não importa por causa da associatividade; por isso, podemos usar, para coberturas $\mathcal{C}_0, \ldots, \mathcal{C}_{n-1} \in \mathsf{C}(X)$, a notação

$$\mathcal{C}_0 \curlyvee \cdots \curlyvee \mathcal{C}_{n-1} := \bigcurlyvee_{i \in [n]} \mathcal{C}_i.$$

3.4.4 Coberturas puxadas

:⊢ **Definição 3.48.** Sejam X e X' conjuntos, $f \colon X \longrightarrow X'$ uma função e $\mathcal{C}'$ uma cobertura de X'. A *cobertura puxada de $\mathcal{C}'$ por f* é o conjunto

$$f^{-1}(\mathcal{C}') := \left\{ f^{-1}(C') \;\middle|\; C' \in \mathcal{C}' \right\}.$$

⊢ **Proposição 3.52.** *Sejam X e X' conjuntos, $f \colon X \longrightarrow X'$ uma função e $\mathcal{C}'$ uma cobertura de X'. A cobertura puxada $f^{-1}(\mathcal{C}')$ é uma cobertura de X.*

□ *Demonstração.* Segue direto de

$$X = f^{-1}(X') = f^{-1}\left(\bigcup_{C' \in \mathcal{C}'} C' \right) = \bigcup_{C' \in \mathcal{C}'} f^{-1}(C'). \qquad \blacksquare$$

⊢ **Proposição 3.53.** *Sejam X, X' e X'' conjuntos, $f \colon X \longrightarrow X'$ e $f' \colon X' \longrightarrow X''$ funções e $\mathcal{C}''$ uma cobertura de X''. Então*

$$(f' \circ f)^{-1}(\mathcal{C}'') = f^{-1}((f')^{-1}(\mathcal{C}'')).$$

□ *Demonstração.* Basta notarmos que, para todo $C'' \in \mathcal{C}''$, vale

$$(f' \circ f)^{-1}(C'') = f^{-1}((f')^{-1}(C'')). \qquad \blacksquare$$

Isso nos permite usar sem ambiguidade, para $f\colon X \longrightarrow X$ e uma cobertura $\mathcal{C}$ de X, a notação

$$f^{-n}(\mathcal{C}) := \left\{ f^{-n}(C) \;\middle|\; C \in \mathcal{C} \right\}.$$

▷ **Exercício 3.10.** *Sejam X e X' conjuntos, $f\colon X \longrightarrow X'$ uma função e $\mathcal{C}'_0$ e $\mathcal{C}'_1$ coberturas de $\boldsymbol{X}'$.*

1. *Se $\mathcal{C}'_0 \preceq \mathcal{C}'_1$, então $f^{-1}(\mathcal{C}'_0) \preceq f^{-1}(\mathcal{C}'_1)$;*
2. *$f^{-1}(\mathcal{C}'_0 \curlyvee \mathcal{C}'_1) = f^{-1}(\mathcal{C}'_0) \curlyvee f^{-1}(\mathcal{C}'_1)$.*

3.4.5 Partições

:⊢ **Definição 3.49** (Partição)**.** Seja X um conjunto. Uma *partição* de X é um conjunto $\mathcal{P} \subseteq \wp(X)$ que satisfaz

1. (Cobertura) $X = \bigcup \mathcal{P}$;
2. (Não vacuidade) $\emptyset \notin \mathcal{P}$;
3. (Não coincidência) Para todos $P, P' \in \mathcal{P}$, se $P \neq P'$ então $P \cap P = \emptyset$.

Os conjuntos $P \in \mathcal{P}$ são as *partes* de $\mathcal{P}$. O conjunto das partições de X é denotado $\mathsf{P}(X)$.

Uma partição é uma cobertura de X por conjuntos disjuntos (logo distintos) que não contém o conjunto vazio.

► **Exemplo 3.6.** *Seja X um conjunto.*

1. *O conjunto $\{X\}$ é a* partição trivial *de X;*
2. *O conjunto*

$$\wp_1(X) = \{C \in \wp(X) \mid \#(C) = 1\} = \{\{x\} \mid x \in X\}$$

é a partição atômica *de X.*

Alguns tipos especiais de partições são úteis na teoria de integração de Riemann. Em $\mathbb{R}^1$, essas partições são chamadas de partições de intervalo, e são representadas como um número finito de pontos em um intervalo. Quando generaliza-se para dimensões maiores, usam-se d-retângulos, que são conjuntos em $\mathbb{R}^d$ dados pelo produto de d intervalos limitados. Podemos fixar um critério a mais, o de que um d-retângulo é produto de intervalos fechados em baixo e abertos em cima. Nesse

caso, podemos definir que uma partição cujos elementos são d-retângulos é uma *malha*.

Alternativamente, quando temos uma medida, que é o caso de $\mathbb{R}^d$, podemos enfraquecer a restrição de que as partes de uma partição são iguais ou disjuntas para a de que são iguais ou *quase disjuntas* (a interseção tem medida zero) e a restrição de que cobrem para a restrição de que a união da partição é quase total (seu complementar tem medida nula) e por fim, de que nenhum elemento da partição é quase vazio (tem medida nula) e definir que isso é uma μ-partição ou *quase partição com respeito a* μ. Como um exemplo, poderíamos considerar retângulos em $\mathbb{R}^d$ formados como produto de intervalos fechados, de modo que eles se intersecionem somente nas fronteiras, que tem medida nula; alternativamente, poderíamos considerar retângulo formados como produtos de intervalos abertos, de modo que as fronteiras não seriam parte da partição, mas a união ainda teria medida total pois as fronteiras têm medida nula.

3.4.6 Refinamento de partições

Como partições são tipos específicos de coberturas, podemos considerar a pré-ordem de refinamento entre partições. Devido à estrutura mais específica das partições, essa pré-ordem, quando restrita ao conjunto de partições, é uma ordem.

$\vdash$ **Proposição 3.54** (Refinamento é ordem)**.** *Seja X um conjunto. A relação de refinamento $\preceq$ é uma relação de ordem no conjunto de partições $\mathsf{P}(X)$.*

$\square$ *Demonstração.* A proposição 3.49 mostra que $\preceq$ é reflexiva e transitiva. Mostremos que quando restrita ao conjunto $\mathsf{P}(X)$, vale a antissimetria. Sejam $\mathcal{P}, \mathcal{P}' \in \mathsf{P}(X)$ tais que $\mathcal{P} \preceq \mathcal{P}'$ e $\mathcal{P}' \preceq \mathcal{P}$. Seja $P \in \mathcal{P}$. Como $\mathcal{P} \preceq \mathcal{P}'$, algum $P' \in \mathcal{P}'$ satisfaz $P' \subseteq P$. Mas, como $\mathcal{P}' \preceq \mathcal{P}$, algum $Q \in \mathcal{P}$ satisfaz $Q \subseteq P'$. Isso implica que $Q \subseteq P$, logo que $Q \cap P = Q$; como $\mathcal{P}$ é uma partição, segue que $Q \cap P = Q \neq \emptyset$, portanto $Q = P$. Assim, $P \subseteq P' \subseteq P$, e concluímos que $P' = P$, logo $P \in \mathcal{P}'$, resultando em $\mathcal{P} \subseteq \mathcal{P}'$. A contenção contrária $\mathcal{P}' \subseteq \mathcal{P}$ segue por simetria, invertendo os papéis de $\mathcal{P}$ e $\mathcal{P}'$, e assim concluímos que $\mathcal{P} = \mathcal{P}'$. ■

$:\vdash$ **Definição 3.50** (Subpartição e superpartição)**.** Sejam X um conjunto e $\mathcal{P}$ uma partição de X. Uma *subpartição* de $\mathcal{P}$ é uma partição mais grossa que $\mathcal{P}$. Uma *superpartição* de $\mathcal{P}$ é uma partição mais fina que $\mathcal{P}$.

Notemos que o conceito de subpartição não coincide com o de subcobertura. Uma subcobertura é menor que uma cobertura referência com relação à ordem de contenção de conjuntos $\subseteq$, enquanto que uma subpartição é menor que uma partição referência com relação à ordem de refinamento de partições $\preceq$. Não falamos em subcoberturas com respeito ao refinamento $\preceq$ pois essa relação não é uma ordem no conjunto de coberturas, somente no conjunto de partições.

⊢ **Proposição 3.55.** *Sejam X um conjunto e $\mathcal{P}, \mathcal{P}'$ partições de X tais que $\mathcal{P} \preceq \mathcal{P}'$. Então*

1. $\#(\mathcal{P}) \leq \#(\mathcal{P}')$;
2. *Para todo $P \in \mathcal{P}$, o conjunto*

$$\mathcal{P}'|_P := \{P' \in \mathcal{P}' \mid P' \subseteq P\}$$

é uma partição de P.

□ *Demonstração.* 1. Por definição de refinamento, para todo $P' \in \mathcal{P}'$ existe $P \in \mathcal{P}$ tal que $P' \subseteq P$. Notemos que P é única pois, se existir $Q \in \mathcal{P}$ tal que $P' \subseteq Q$, então $P' \subseteq P \cap Q$ e, como $P' \neq \emptyset$, segue que $P = Q$. Assim, consideramos a função que mapeia, para cada $P' \in \mathcal{P}'$ a parte $P_{P'} \in \mathcal{P}$ tal que $P' \subseteq P_{P'}$:

$$\begin{aligned} f\colon \mathcal{P}' &\longrightarrow \mathcal{P} \\ P' &\longmapsto P_{P'}. \end{aligned}$$

Mostremos que essa função é sobrejetiva. Para isso, seja $P \in \mathcal{P}$. Como $\bigcup \mathcal{P}' = X$, para todo $x \in P \subseteq X$ existe $P' \in \mathcal{P}'$ tal que $x \in P'$. Como $P \neq \emptyset$, existe $x \in P$, logo existe $P' \in \mathcal{P}'$ tal que $x \in P'$. Por definição de refinamento, existe $Q \in \mathcal{P}$ tal que $P' \subseteq Q$, o que implica $x \in Q$. Como $x \in Q \cap P$, segue que $P = Q$, e concluímos que $P' \subseteq P$. Isso mostra que $f(P') = P$, logo que f é sobrejetiva. Concluímos, então, que $\#(\mathcal{P}) \leq \#(\mathcal{P}')$.

2. As propriedades de não coincidência e não vacuidade são evidentes porque $\mathcal{P}'$ é partição. Para a propriedade de cobertura, seja $P \in \mathcal{P}$ e $U := \bigcup \{P' \in \mathcal{P}' \mid P' \subseteq P\}$. Notemos que $P = U$. Para mostrar isso, seja $x \in P$. Então existe $P' \in \mathcal{P}'$ tal que $x \in P'$, pois $\bigcup \mathcal{P} = X$. Por definição de refinamento, existe $Q \in \mathcal{P}$ tal que $P' \subseteq Q$, portanto $x \in Q$. Como $x \in Q \cap P$, segue que $P = Q$. Concluímos que $P' \subseteq P$, portanto que $x \in U$, o que mostra $P \subseteq U$. Reciprocamente, para todo $P' \in U$, $P' \subseteq P$, portanto $U \subseteq P$, e concluímos que $P = U$. ■

3.4.7 Correfinamento de partições

A operação de correfinamento de coberturas também pode ser calculada em partições, pois elas são tipos específicos de cobertura. Porém, o que se deve mostrar é que o correfinamento de partições ainda é uma partição, para que assim a operação possa de fato ser restrita ao conjunto de partições. De fato isso não é verdade, mas somente pelo motivo técnico de que quando se calcula o correfinamento, o conjunto vazio pode passar a ser um elemento da cobertura resultante, e não permitimos conjuntos vazios em partições. Assim, adaptamos a definição de correfinamento de coberturas para excluir esse conjunto.

:⊢ **Definição 3.51** (Correfinamento de partições)**.** Sejam X um conjunto e $\mathcal{P}$ e $\mathcal{P}'$ partições de X. O *correfinamento* de $\mathcal{P}$ e $\mathcal{P}'$ é o conjunto

$$\mathcal{P} \curlyvee \mathcal{P}' := \{P \cap P' \mid P \in \mathcal{P}, P' \in \mathcal{P}'\} \smallsetminus \{\emptyset\}.$$

Seja $(\mathcal{P}_i)_{i\in I}$ uma família de partições de X. O *correfinamento* de $(\mathcal{P}_i)_{i\in I}$ é o conjunto

$$\bigvee_{i\in I} \mathcal{P}_i := \left\{ \bigcap_{i\in I} P_i \;\middle|\; \forall_{i\in I} P_i \in \mathcal{P}_i \right\} \smallsetminus \{\emptyset\}.$$

⊢ **Proposição 3.56.** *Sejam X um conjunto e $(\mathcal{P}_i)_{i\in I}$ uma família de partições de X. O correfinamento $\bigvee_{i\in I} \mathcal{P}_i$ de $(\mathcal{P}_i)_{i\in I}$ é a menor partição de que X (com respeito à ordem $\preceq$) que, para todo $i \in I$, é mais fina que $\mathcal{P}_i$.*

□ *Demonstração.* Denotemos $\mathcal{P} := \bigvee_{i\in I} \mathcal{P}_i$ por simplicidade. Primeiro, mostremos que $\mathcal{P}$ é uma partição.

- (Cobertura) Seja $x \in X$. Para todo $i \in I$, como $X = \bigcup \mathcal{P}_i$, segue que algum $P_i \in \mathcal{P}_i$ satisfaz $x \in P_i$. Sendo assim, $x \in \bigcap_{i\in I} P_i$, portanto $X \subseteq \bigcup \mathcal{P}$, e segue que $X = \bigcup \mathcal{P}$.
- (Não vacuidade) Por definição, $\emptyset \notin \mathcal{P}$.
- (Não coincidência) Sejam $P = \bigcap_{i\in I} P_i, P' = \bigcap_{i\in I} P'_i \in \mathcal{P}$. Se $P \neq P'$, então existe $x \in P \smallsetminus P'$ ou existe $x \in P' \smallsetminus P$. Sem perda de generalidade, suponhamos o primeiro. Então, existe $i \in I$ tal que $x \notin P'_i$. Como $x \in P$, então $x \in P_i$, portanto $P_i \neq P'_i$. Mas então, como $\mathcal{P}_i$ é partição, $P_i \cap P'_i = \emptyset$. Por fim, como $P \subseteq P_i$ e $P' \subseteq P'_i$, segue que $P \cap P' = \emptyset$.

Agora mostraremos que, para todo $i \in I$, $\mathcal{P}_i \preceq \mathcal{P}$. Sejam $i \in I$ e $P \in \mathcal{P}$. Então $P = \bigcap_{i\in I} P_i$, com $P_i \in \mathcal{P}_i$. Isso significa que, para todo $i \in I$, $P \subseteq P_i$, portanto $\mathcal{P}_i \preceq \mathcal{P}$.

Por fim, $\mathcal{P}$ é a menor partição com essa propriedade. Sejam $\mathcal{P}'$ partição de X que refina $\mathcal{P}_i$ para todo $i \in I$ e $P' \in \mathcal{P}'$. Então, para todo $i \in I$, algum $P_i \in \mathcal{P}_i$ satisfaz $P' \subseteq P_i$. Portanto $P' \subseteq \bigcap_{i\in I} P_i$, e como $\bigcap_{i\in I} P_i \in \mathcal{P}$, segue que $\mathcal{P} \preceq \mathcal{P}'$. ■

Capítulo 4

Cardinalidade de conjuntos

4.1 Relações

4.1.1 Igualdade de cardinais

:⊢ **Definição 4.1.** Sejam X e Y conjuntos. Diz-se que $\#(X) = \#(Y)$ (a *cardinalidade de X é igual à cardinalidade de Y*) se, e somente se, existe uma bijeção C entre X e Y. Caso contrário, diz-se que $\#(X) \neq \#(Y)$ (a *cardinalidade de X é diferente da cardinalidade de Y*).

As cardinalidades dos números naturais e dos números reais são denotadas, respectivamente

$$\aleph_0 := \#(\mathbb{N}) \quad \text{e} \quad \mathfrak{c} := \#(\mathbb{R}).$$

⊢ **Proposição 4.1.** *Sejam X, Y e Z conjuntos não vazios. Então*

1. $\#(X) = \#(X)$;
2. $\#(X) = \#(Y) \Rightarrow \#(Y) = \#(X)$;
3. $\#(X) = \#(Y)$ *e* $\#(Y) = \#(Z) \Rightarrow \#(X) = \#(Z)$.

□ *Demonstração.*

1. Claramente, a função identidade em X é uma bijeção entre X e X e, portanto, $\#(X) = \#(X)$.
2. Se $\#(X) = \#(Y)$, então existe bijeção $C : X \to Y$. Mas então $C^{-1} : Y \to X$ é uma bijeção de Y em X e, portanto, $\#(Y) = \#(X)$.
3. Se $\#(X) = \#(Y)$ e $\#(Y)$ e $\#(Y)$, então existem bijeções $C_1 : X \to Y$ e $C_2 : Y \to Z$. Mas então $C_2 \circ C_1 : X \to Z$ é uma bijeção de X em Y e, portanto, $\#(X) = \#(Z)$.

■

De certa forma, essa proposição mostra que a noção de cardinalidades iguais se comporta como uma relação de equivalência. Não podemos dizer que $=$ é, de fato,

uma relação de equivalência porque nao existe um conjunto de todos os conjuntos no qual defini-la. Todas proposições sobre cardinalidades são, na verdade, proposições sobre funções entre conjuntos e convém saber que as propriedades acima valem.

4.1.2 Ordenação de cardinais

:⊢ **Definição 4.2.** Sejam X e Y conjuntos não vazios.

1. Diz-se que $\#(X) \leq \#(Y)$ (a *cardinalidade de X é menor ou igual à cardinalidade de Y*) se, e somente se, existe função injetiva $C : X \to Y$.
   Diz-se que $\#(X) \geq \#(Y)$ (a *cardinalidade de X é maior ou igual à cardinalidade de Y*) se, e somente se, existe função sobrejetiva $C : X \to Y$.
2. Diz-se que $\#(X) < \#(Y)$ (a *cardinalidade de X é menor que a cardinalidade de Y*) se, e somente se, $\#(X) \leq \#(Y)$ e $\#(X) \neq \#(Y)$.
   Diz-se que $\#(X) > \#(Y)$ (a *cardinalidade de X é maior que a cardinalidade de Y*) se, e somente se, $\#(X) \geq \#(Y)$ e $\#(X) \neq \#(Y)$.

:⊢ **Definição 4.3.** Um conjunto *enumerável* (ou *contável*) é um conjunto X tal que $\#X \leq \aleph_0$. Uma função injetiva $E : X \to \mathbb{N}$ é uma *enumeração* de X.

:⊢ **Definição 4.4.** Um conjunto *finito* é um conjunto X tal que $\#X < \aleph_0$. Um conjunto *infinito* é um conjunto que não é finito.

A seguir, demonstraremos algumas proposições para mostrar que o símbolo $\leq$ se comporta como uma relação de ordem total. Novamente, não podemos dizer formalmente que $\leq$ é uma relação, pois não existe o conjunto de todos os conjuntos no qual defini-la. No entanto, as propriedades acima são bem úteis de se ter em mente e serão usadas na demonstração de outras proposições. As propriedades análogas à reflexividade e transitividade de uma relação de ordem são bem triviais. A antissimetria, por outro lado, é bem difícil, tanto que é um conhecido teorema, o Teorema de Cantor-Schröder-Bernstein. Ainda, é possível demonstrar que $\leq$ se comporta como uma relação total; ou seja, todo conjunto pode ser comparado. Vamos demonstrar primeiro as propriedades triviais. Em seguida, demonstraremos separadamente as outras duas.

⊢ **Proposição 4.2.** *Sejam X, Y e Z conjuntos não vazios. Então*

1. $\#(X) \leq \#(X)$;
2. $\#(X) \leq \#(Y)$ *e* $\#(Y) \leq \#(X) \Rightarrow \#(X) = \#(Y)$;
3. $\#(X) \leq \#(Y)$ *e* $\#(Y) \leq \#(Z) \Rightarrow \#(X) \leq \#(Z)$;
4. $\#(X) \leq \#(Y)$ *ou* $\#(Y) \leq \#(X)$.

□ *Demonstração.* 1. Claramente, a função identidade é uma bijeção de X em X, logo é uma injeção de X em X e, portanto, $\#(X) \leq \#(X)$.

2. Teorema de Cantor-Schröder-Bernstein.
3. $\#(X) \leq \#(Y)$ e $\#(Y) \leq \#(Z)$, então existem funções injetivas $C_1 : X \to Y$ e $C_2 : Y \to Z$. Mas então $C_2 \circ C_1 : X \to Z$ é uma função injetiva de X em Z e, portanto, $\#(X) \leq \#(Z)$.

■

MOSTRAR QUE INFINITO EQUIVALE A
X tal que $\#(X) \geq \aleph_0$.

4.2 Operações

:⊢ **Definição 4.5.** Sejam X e Y conjuntos não vazios. Definimos as seguintes "operações"entre cardinais:

1. $\#(X) + \#(Y) := \#(X + Y)$;
2. $\#(X) \times \#(Y) := \#(X \times Y)$;
3. $\#(X)^{\#(Y)} := \#\left(X^Y\right)$.

4.2.1 Cardinalidade de soma (ou união disjunta)

⊢ **Proposição 4.3.** *Seja $(C_i)_{i\in I}$ uma família não vazia de conjuntos disjuntos dois a dois. Então*

$$\#\left(\bigsqcup_{i\in I} C_i\right) = \#\left(\bigcup_{i\in I} C_i\right).$$

□ *Demonstração.* Consideremos a função

$$\begin{aligned} f\colon \bigsqcup_{i\in I} C_i &\longrightarrow \bigcup_{i\in I} C_i \\ (c, i) &\longmapsto c. \end{aligned}$$

Mostremos que f é bijeção. (Injetividade) Sejam $(c_1, i_1), (c_2, i_2) \in \bigsqcup_{i\in I} C_i$ tais que $c_1 = c_2$. Como os C_i são disjuntos dois a dois, existe único $i \in I$ tal que $c_1 = c_2 \in C_i$. Logo $i_1 = i_2 = i$ e, portanto, $(c_1, i_1) = (c_2, i_2)$. (Sobrejetividade) Seja $c \in \bigcup_{i\in I} C_i$. Então existe $i \in I$ tal que $c \in C_i$. ■

⊢ **Proposição 4.4.** *Seja $(C_i)_{i\in I}$ uma família não vazia de conjuntos de mesma cardinalidade. Então*

$$\#\left(\bigsqcup_{i\in I} C_i\right) = \#(I) \times \#(C)$$

para algum C de $(C_i)_{i\in I}$.

□ *Demonstração.* Como $I \neq \emptyset$, seja $j \in I$ e defina $C := C_j$. Como todos os conjuntos de $(C_i)_{i\in I}$ têm a mesma cardinalidade, para todo $i \in I$, seja $f_i : C_i \to C$ bijeção. Considere a função

$$\begin{aligned} f\colon \bigsqcup_{i\in I} C_i &\longrightarrow I \times C \\ (c, i) &\longmapsto (i, f_i(c)). \end{aligned}$$

Mostremos que f é bijeção. (Injetividade) Sejam $(c_1, i_1), (c_2, i_2) \in \bigsqcup_{i\in I} C_i$ tais que $(i_1, f_{i_1}(c_1)) = (i_2, f_{i_2}(c_2))$. Então $i_1 = i_2$ e $f_{i_1}(c_1) = f_{i_2}(c_2)$, o que implica que $f_{i_1} = f_{i_2}$ e, portanto, $c_1 = c_2$, já que f_{i_1} é injetiva. Logo $(c_1, i_1) = (c_2, i_2)$. (Sobrejetividade) Seja $(i, c) \in I \times C$. Como f_i é sobrejetiva, existe $c' \in C_i$ tal que $f_i(c') = c$. Portanto $f(c', i) = (i, f_i(c')) = (i, c)$.

Da sobrejetividade de f e da definição de produto de cardinais, segue que

$$\#\left(\bigsqcup_{i\in I} C_i\right) = \#(I \times C) = \#(I) \times \#(C).$$

■

⊢ **Teorema 4.5.** *Seja $(C_i)_{i\in I}$ uma família não vazia de conjuntos. Se existem $\min_{i\in I} \#(C_i)$ e $\max_{i\in I} \#(C_i)$, então*

$$\#(I) \times \min_{i\in I} \#(C_i) \leq \#\left(\bigsqcup_{i\in I} C_i\right) \leq \#(I) \times \max_{i\in I} \#(C_i).$$

□ *Demonstração.* Mostremos a primeira desigualdade. Seja $j \in I$ tal que

$$\#(C_j) := \min_{i\in I} \#(C_i)$$

e $C := C_j$. Para cada $i \in I$, existe função injetiva $f_i : C \to C_i$. Considere a função

$$\begin{aligned} f\colon I \times C &\longrightarrow \bigsqcup_{i\in I} C_i \\ (i, c) &\longmapsto (f_i(c), i). \end{aligned}$$

Mostremos que f é injetiva. Sejam $(i_1, c_1), (i_2, c_2) \in I \times C$ tais que $(f_{i_1}(c_1), i_1) = (f_{i_2}(c_2), i_2)$. Então $i_1 = i_2$ e $f_{i_1} = f_{i_2}$. Como f_{i_1} é injetiva, temos que $c_1 = c_2$, logo $(i_1, c_1) = (i_2, c_2)$.

Mostremos agora a segunda desigualdade. Seja $j \in I$ tal que $\#(C_j) := \max_{i\in I} \#(C_i)$ e $C := C_j$. Para cada $i \in I$, existe função sobrejetiva $f_i : C_i \to C$. Considere a função

$$\begin{aligned} f\colon \bigsqcup_{i\in I} C_i &\longrightarrow I \times C \\ (c, i) &\longmapsto (i, f_i(c)). \end{aligned}$$

Mostremos que f é sobrejetiva. Seja $(i, c) \in I \times C$. Como f_i é sobrejetiva, existe $c' \in C_i$ tal que $f_i(c') = c$. Portanto $f(c', i) = (i, f_i(c')) = (i, c)$. ■

Parte 2

Álgebra

Capítulo 5

Estruturas básicas

A *Álgebra* estuda objetos matemáticos conhecidos como *estruturas algébricas*. As definições desses objetos variam e podem ser tomadas de modo a serem mais ou menos gerais. No entanto, esses objetos em geral são n-listas cujas entradas são conjuntos e funções. Uma das definições que podem ser tomadas é a de que essas estruturas são listas em que a primeira entrada é um conjunto e as demais são funções. Em geral, essas funções são *operações n-árias*, funções da n-ésima potência de um conjunto nele mesmo. Não definiremos aqui esses objetos com detalhes, nos restringindo somente a casos específicos. Ao leitor fica a oportunidade de perceber as semelhanças entre as definições e generalizá-las, ou mesmo de procurar mais a respeito.

5.1 Conjuntos numéricos

Antes de apresentar as estruturas algébricas abstratas, analisaremos o construção dos números naturais e dos números inteiros.

5.1.1 Números naturais

:⊢ **Definição 5.1.** Um *modelo de números naturais* é uma tripla $\boldsymbol{N} = (N, 0, \mathbin{+\!\!\!+})$ em que

1. N é um conjunto, o *conjunto de números naturais*;
2. $0 \in N$, o *zero* de $\boldsymbol{N}$;
3. $\mathbin{+\!\!\!+} \colon N \to N$ é uma função injetiva tal que $\mathbin{+\!\!\!+}^{-1}(\{0\}) = \emptyset$, a função *sucessor*;
4. (Axioma da Indução) Para todo conjunto $I \subseteq N$, se $0 \in I$ e $\mathbin{+\!\!\!+}(I) \subseteq I$, então $I = N$.

O *um* de $\boldsymbol{N}$ é o elemento $1 := \mathbin{+\!\!\!+}(0)$.

Pela teoria de conjuntos, é possível definir um conjunto infinito $\boldsymbol{N}$ que satisfaz os axiomas de um modelo de números naturais. A construção considera $0 := \emptyset$, $1 := \{0\}$, e, de modo geral, $\vdash\!\!\dashv(n) := n \cup \{n\} = \{0, 1, \cdots, n\}$. Claramente a construção é feita com mais cuidado, mas a partir dessa construção podemos realmente achar um modelo de números naturais. A partir de agora, consideraremos que esse conjunto existe.

$\vdash$ **Proposição 5.1.** *Seja $\boldsymbol{N}$ um modelo de números naturais. Então, para todo $n \in N \setminus \{0\}$, existe $m \in N$ tal que $n = \vdash\!\!\dashv(m)$.*

□ *Demonstração.* Seja $I := \{n \in N : n = 0 \text{ ou } \exists m \in N \quad n = \vdash\!\!\dashv(m)\}$. Primeiro, notemos que $0 \in I$. Agora, seja $n \in I$. Então $\vdash\!\!\dashv(n) \in I$, pois $n \in N$ e $\vdash\!\!\dashv(n) = \vdash\!\!\dashv(n)$. Logo $I = N$. Assim, se $n \in N \setminus \{0\}$, segue que existe $m \in N$ tal que $n = \vdash\!\!\dashv(n)$. ■

Essa proposição mostra que $\vdash\!\!\dashv$ é sobrejetiva em $N \setminus \{0\}$ e, portanto, que $\vdash\!\!\dashv$ é uma bijeção entre N e $N \setminus \{0\}$, o que mostra que N é um conjunto infinito. No entanto, vale lembrar que a definição de conjunto infinito depende do conjunto dos números naturais.

5.1.1.1 Adição

$\vdash$ **Teorema 5.2.** *Seja $\boldsymbol{N}$ um modelo de números naturais. Existe uma única função*

$$\begin{aligned} +\colon N \times N &\longrightarrow N \\ (n_1, n_2) &\longmapsto n_1 + n_2 \end{aligned}$$

que satisfaz

1. *(A1)* $\forall n \in N \qquad n + 0 = n$;
2. *(A2)* $\forall n_1, n_2 \in N \qquad n_1 + \vdash\!\!\dashv(n_2) = \vdash\!\!\dashv(n_1 + n_2)$.

□ *Demonstração.* Primeiro mostraremos que essa função $+$ está bem definida. Para isso, devemos mostrar que, para todo $n_1, n_2 \in N$, existe único $n_3 \in N$ tal que $n_1 + n_2 = n_3$ satisfazendo $(A_1), (A_2)$. Consideremos o conjunto $I := \{n \in N : \exists! n_3 \in N \quad n_1 + n = n_3\}$. Primeiro, notemos que $0 \in I$, pois $n + 0 = n$ e, portanto, n_3 é único. Agora, seja $n \in I$. Então existe único $n_3 \in N$ tal que $n_1 + n = n_3$ e, como $\vdash\!\!\dashv$ é função, $\vdash\!\!\dashv(n_3) = \vdash\!\!\dashv(n_1 + n) \in N$ é único e tomando $n_1 + \vdash\!\!\dashv(n) = \vdash\!\!\dashv(n_1 + n)$, concluímos que $\vdash\!\!\dashv(n) \in I$ e, portanto, $I = N$. Logo $+$ está bem definida. Agora, mostremos que $+$ é única. Sejam $+_1, +_2 : N \times N \to N$ funções satisfazendo $(A_1), (A_2)$, $n_1 \in N$ e $I := \{n \in N : n_1 +_1 n = n_1 +_2 n\}$. Primeiro, notemos que $0 \in I$, pois $n_1 +_1 0 = n = n_1 +_2 0$. Agora, seja $n \in I$. Então

$$n_1 +_1 \vdash\!\!\dashv(n) = \vdash\!\!\dashv(n_1 +_1 n) = \vdash\!\!\dashv(n_1 +_2 n) = n_1 +_2 \vdash\!\!\dashv(n),$$

o que implica que $\vdash\!\!\dashv(n) \in I$ e, portanto, que $I = N$. Logo $+_1 = +_2$. ■

:⊢ **Definição 5.2.** Seja $\boldsymbol{N}$ um modelo de números naturais. A função $+$ é a *adição nos números naturais* e, dados $n_1, n_2 \in N$, o número $n_1 + n_2 \in N$ é a *soma de* n_1 *e* n_2.

⊢ **Teorema 5.3** (Associatividade da adição)**.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n_1, n_2, n_3 \in N \qquad (n_1 + n_2) + n_3 = n_1 + (n_2 + n_3).$$

□ *Demonstração.* Sejam $n_1, n_2 \in N$ e $I := \{n_3 \in N : (n_1+n_2)+n_3 = n_1+(n_2+n_3)\}$. Notemos que $0 \in I$, pois

$$\begin{aligned} (n_1 + n_2) + 0 &= n_1 + n_2 && \text{(A1)} \\ &= n_1 + (n_2 + 0). && \text{(A1)} \end{aligned}$$

Agora, seja $n \in I$. Então

$$\begin{aligned} (n_1 + n_2) + ⧺(n) &= ⧺((n_1 + n_2) + n) && \text{(A2)} \\ &= ⧺(n_1 + (n_2 + n)) && (n \in I) \\ &= n_1 + ⧺(n_2 + n) && \text{(A2)} \\ &= n_1 + (n_2 + ⧺(n)), && \text{(A2)} \end{aligned}$$

o que implica $⧺(n) \in I$. Logo $I = N$. ■

⊢ **Teorema 5.4.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n \in N \qquad ⧺(n) = n + 1.$$

□ *Demonstração.* Seja $n \in N$. Então

$$⧺(n) = ⧺(n + 0) = n + ⧺(0) = n + 1.$$

■

⊢ **Lema 5.5.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

1. $\forall n \in N \qquad 0 + n = n;$
2. $\forall n \in N \qquad 1 + n = n + 1.$

□ *Demonstração.* Demonstraremos ambas afirmações por indução em n.

1. Seja $I := \{n \in N : 0 + n = n\}$. Primeiro notemos que $0 \in I$, pois $0 + 0 = 0$. Agora, seja $n \in I$. Então

$$0 + ⧺(n) = 0 + (n + 1) = (0 + n) + 1 = n + 1 = ⧺(n),$$

o que implica que $⧺(n) \in I$ e, portanto, $I = N$.

2. Seja $I := \{n \in N : 1 + n = n + 1\}$. Primeiro notemos que $0 \in I$, pois $1 + 0 = 1 = 0 + 1$. Agora, seja $n \in I$. Então

$$1 + \mathrm{H}(n) = 1 + (n + 1) = (1 + n) + 1 = (n + 1) + 1 = \mathrm{H}(n) + 1,$$

o que implica que $\mathrm{H}(n) \in I$ e, portanto, $I = N$.

■

⊢ **Teorema 5.6** (Comutatividade da adição)**.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n_1, n_2 \in N \qquad n_1 + n_2 = n_2 + n_1.$$

□ *Demonstração.* Demonstraremos a afirmação por indução. Seja $n_1 \in N$ e $I := \{n \in N : n_1 + n = n + n_1\}$. Primeiro notemos que $0 \in I$, pois

$$n_1 + 0 = n_1 = 0 + n_1.$$

Agora, seja $n \in I$. Então

$$\begin{aligned} n_1 + \mathrm{H}(n) &= n_1 + (n + 1) \\ &= (n_1 + n) + 1 \\ &= (n + n_1) + 1 \\ &= n + (n_1 + 1) \\ &= n + (1 + n_1) \\ &= (n + 1) + n_1 \\ &= \mathrm{H}(n) + n_1, \end{aligned}$$

o que implica que $\mathrm{H}(n) \in I$ e, portanto, $I = N$. ■

5.1.1.2 Multiplicação

⊢ **Teorema 5.7.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Existe uma única função*

$$\begin{aligned} \times \colon N \times N &\longrightarrow N \\ (n_1, n_2) &\longmapsto n_1 \times n_2 \end{aligned}$$

que satisfaz

1. *(M1)* $\forall n \in N \qquad n \times 0 = 0$*;*
2. *(M2)* $\forall n_1, n_2 \in N \qquad n_1 \times \mathrm{H}(n_2) = (n_1 \times n_2) + n_1$.

□ *Demonstração.* Primeiro devemos mostrar que a função $\times$ está bem definida. Para isso, devemos mostrar que, para todo $n_1, n_2 \in N$, existe único $n_3 \in N$ tal que $n_1 \times n_2 = n_3$. Consideremos $I := \{n \in N : \exists! n_3 \in N \quad n_1 \times n = n_3\}$. Primeiro, notemos que $0 \in I$, pois $n_1 \times 0 = 0$ e, portanto, n_3 existe e é único. Agora, seja $n \in I$. Então existe único $n_3 \in N$ tal que $n_1 \times n = n_3$ e, como $+$ é função, $n_3 + n_1 = n_1 \times n + n$ é único e tomando $n_1 \times \mathrm{H}(n) = n_1 \times n + n_1$, concluímos que $\mathrm{H}(n) \in I$ e, portanto, $I = N$. Logo $\times$ está bem definida. Agora, devemos mostrar que $\times$ é única. Sejam $\times_1, \times_2 : N \times N \to N$ funções satisfazendo $(M_1), (M_2)$, $n_1 \in N$ e $I := \{n \in N : n_1 \times_1 n = n_1 \times_2 n\}$. Primeiro, notemos que $0 \in I$, pois $n_1 \times_1 0 = 0 = n_1 \times_2 0$. Agora, seja $n \in I$. Então

$$n_1 \times_1 \mathrm{H}(n) = n_1 \times_1 n + n_1 = n_1 \times_2 n + n_1 = n_1 \times_2 \mathrm{H}(n),$$

o que implica que $\mathrm{H}(n) \in I$ e, portanto, que $I = N$. Logo $\times_1 = \times_2$. ■

:⊢ **Definição 5.3.** Seja $\boldsymbol{N}$ um modelo de números naturais. A função $\times$ é a *multiplicação nos números naturais* e, dados $n_1, n_2 \in N$, o número $n_1 \times n_2 \in N$ é o *produto de* n_1 *e* n_2.

⊢ **Teorema 5.8** (Distributividade)**.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

1. $\forall n, m, k \in N \qquad n \times (m + k) = (n \times m) + (n \times k)$;
2. $\forall n, m, k \in N \qquad (n + m) \times k = (n \times k) + (m \times k)$.

□ *Demonstração.* 1. Sejam $n, m \in N$ e $I := \{k \in N : n \times (m + k) = (n \times m) + (n \times k)\}$. Primeiro, notemos que $0 \in I$, pois

$$\begin{aligned} n \times (m + 0) &= n \times m && (A_1) \\ &= n \times m + 0 && (A_1) \\ &= (n \times m) + (n \times 0). && (M_1) \end{aligned}$$

Agora, seja $k \in I$. Então

$$\begin{aligned} n \times (m + \mathrm{H}(k)) &= n \times \mathrm{H}(m + k) && (A_2) \\ &= (n \times (m + k)) + n && (M_2) \\ &= ((n \times m) + (n \times k)) + n && (k \in I) \\ &= (n \times m) + ((n \times k) + n) && (5.3) \\ &= (n \times m) + (n \times \mathrm{H}(k)), && (M_2) \end{aligned}$$

o que implica que $\mathrm{H}(k) \in I$ e, portanto, que $I = N$.

2. Sejam $n, m \in N$ e $I := \{k \in N : (n+m) \times k = (n \times k) + (m \times k)\}$. Primeiro, notemos que $0 \in I$, pois

$$\begin{aligned} (n+m) \times 0 &= 0 && (M_1) \\ &= 0 + 0 && (A_1) \\ &= (n \times 0) + (m \times 0). && (M_1) \end{aligned}$$

Agora, seja $k \in I$. Então

$$\begin{aligned} (n+m) \times \vdash\!\!\dashv(k) &= ((n+m) \times k) + (n+m) && (M_2) \\ &= ((n \times k) + (m \times k)) + (n+m) && (k \in I) \\ &= ((n \times k) + n) + ((m \times k) + m) && (5.3) \\ &= (n \times \vdash\!\!\dashv(k)) + (m \times \vdash\!\!\dashv(k)), && (M_2) \end{aligned}$$

o que implica que $\vdash\!\!\dashv(k) \in I$ e, portanto, que $I = N$.

■

⊢ **Teorema 5.9** (Associatividade da multiplicação)**.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n, m, k \in N \qquad (n \times m) \times k = n \times (m \times k).$$

□ *Demonstração.* Sejam $n, m \in N$ e $I := \{k \in N : (n \times m) \times k = n \times (m \times k)\}$. Primeiro, notemos que $0 \in I$, pois

$$(n \times m) \times 0 = 0 = n \times 0 = n \times (m \times 0) \qquad (M_1)$$

Agora, seja $k \in I$. Então

$$\begin{aligned} (n \times m) \times \vdash\!\!\dashv(k) &= ((n \times m) \times k) + (n \times m) && (M_2) \\ &= (n \times (m \times k)) + (n \times m) && (k \in I) \\ &= n \times ((m \times k) + m) && (5.8) \\ &= n \times (m \times \vdash\!\!\dashv(k)), && (M_2) \end{aligned}$$

o que implica que $\vdash\!\!\dashv(k) \in I$ e, portanto, que $I = N$. ■

⊢ **Lema 5.10.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

1. $\forall n \in N \qquad 0 \times n = 0$;
2. $\forall n \in N \qquad n \times 1 = n = 1 \times n$.

□ *Demonstração.* 1. Vamos mostrar por indução em n. Seja $I := \{n \in N : 0 \times n = 0\}$. Primeiro, notemos que $0 \in I$, pois $0 \times 0 = 0$. Agora, seja $n \in I$. Então

$$0 \times \vdash\!\!\dashv(n) = (0 \times n) + 0 = 0 + 0 = 0,$$

o que mostra que $\vdash\!\!\dashv(n) \in N$ e, portanto, $I = N$.

2. Seja $n \in N$. Então

$$n \times 1 = (n \times 0) + n = 0 + n = n.$$

Mostraremos a segunda igualdade por indução em n. Seja $I := \{n \in N : 1 \times n = n\}$. Primeiro, notemos que $0 \in I$, pois $1 \times 0 = 0$. Agora, seja $n \in I$. Então

$$1 \times \vdash\!\!\dashv(n) = (1 \times n) + 1 = n + 1 = \vdash\!\!\dashv(n),$$

o que implica que $\vdash\!\!\dashv(n) \in I$ e, portanto, que $I = N$.

■

⊢ **Teorema 5.11.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n, m \in N \qquad n \times m = m \times n.$$

□ *Demonstração.* Sejam $n \in N$ e $I := \{m \in N : n \times m = m \times n\}$. Primeiro, notemos que $0 \in I$, pois

$$\begin{aligned} n \times 0 &= 0 && (M_1) \\ &= 0 \times n. && (5.10) \end{aligned}$$

Agora, seja $m \in I$. Então

$$\begin{aligned} n \times \vdash\!\!\dashv(m) &= (n \times m) + n && (M_2) \\ &= (m \times n) + n && (m \in I) \\ &= (m \times n) + (1 \times n) && (5.10) \\ &= (m + 1) \times n && (5.8) \\ &= \vdash\!\!\dashv(m) \times n, && (5.4) \end{aligned}$$

o que implica que $\vdash\!\!\dashv(m) \in I$ e, portanto, que $I = N$. ■

5.1.1.3 Ordenação

⊢ **Lema 5.12.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

1. $\forall n, m, k \in N \qquad n + k = m + k \quad \Longrightarrow \quad n = m;$
2. $\forall n, m \in N \qquad n + m = 0 \quad \Longrightarrow \quad n = m = 0.$

□ *Demonstração.* 1. Seja $I := \{k \in N : \forall n, m \in N \quad n + k = m + k \quad \Longrightarrow n = m\}$. Primeiro, notemos que $0 \in I$, pois, para todos $n, m \in N$, se $n+0 = m+0$, então $n = m$. Agora, seja $k \in I$ e $n, m \in N$. Se $n + \vdash\!\!\!\dashv(k) = m + \vdash\!\!\!\dashv(k)$, então $\vdash\!\!\!\dashv(n + k) = \vdash\!\!\!\dashv(m + k)$ e, como $\vdash\!\!\!\dashv$ é injetiva, $n + k = m + k$, o que implica que $n = m$ e, assim, temos que $\vdash\!\!\!\dashv(k) \in I$ e, portanto, $I = N$.

2. Suponhamos, por absurdo, que $n \neq 0$ ou $m \neq 0$. Notemos que $n+m = m+n$; então, sem perda de generalidade, seja $m \neq 0$. Então existe $k \in N$ tal que $m = \vdash\!\!\!\dashv(k)$ e segue que $n + m = n + \vdash\!\!\!\dashv(k) = \vdash\!\!\!\dashv(n + k) = 0$, o que é absurdo, pois $\vdash\!\!\!\dashv^{-1}(\{0\}) = \emptyset$. Logo $n = m = 0$.

■

:⊢ **Definição 5.4.** Seja $\boldsymbol{N}$ um modelo dos números naturais. A relação binária $\leq$ em N é definida por

$$n \leq m \qquad \Longleftrightarrow \qquad \exists d \in N \quad n + d = m.$$

⊢ **Proposição 5.13.** *Seja $\boldsymbol{N}$ um modelo dos números naturais. A relação binária $\leq$ em N é uma relação de ordem total.*

□ *Demonstração.* Primeiro, notemos que $\leq$ é reflexiva, pois, pra todo $n \in N$, $n + 0 = n$, o que implica que $n \leq n$. Segundo, notemos que $\leq$ é antissimétrica. Sejam $n, m \in N$ tais que $n \leq m$ e $m \leq n$; então existem $d_1, d_2 \in N$ tais que $n + d_1 = m$ e $m + d_2 = n$ e, portanto, que $n + m = n + m + d_1 + d_2$, o que implica $d_1 + d_2 = 0$ e, portanto, que $d_1 = d_2 = 0$. Assim $n = m$. Terceiro, mostremos que $\leq$ é transitiva. Sejam $m, n, k \in N$ tais que $n \leq m$ e $m \leq k$. Então existem $d_1, d_2 \in N$ tais que $n + d_1 = m$ e $m + d_2 = k$. Assim, $n + d_1 + d_2 = k$, logo $n \leq k$. Isso termina a demonstração de que $\leq$ é uma ordem parcial. Por fim, devemos mostrar que a ordem parcial $\leq$ é total. Sejam $n \in N$ e $I := \{m \in N : n \leq m \text{ ou } m \leq n\}$. Primeiro, notemos que $0 \in I$, pois $0 + n = n$, logo $0 \leq n$. Agora, seja $m \in I$. Se $nleqm$, existe $d \in N$ tal que $n+d = m$, e segue que, como $n+d+1 = m+1 = \vdash\!\!\!\dashv(m)$, $n \leq \vdash\!\!\!\dashv(m)$. Se $m \leq n$, existe $d \in N$ tal que $m+d = n$. Consideramos dois casos: se $d = 0$, então $n+1 = m+1 = \vdash\!\!\!\dashv(m)$, logo $n \leq \vdash\!\!\!\dashv(m)$; se $d \neq 0$, existe $k \in N$ tal que $d = \vdash\!\!\!\dashv(k) = k + 1$, o que implica $n = m + d = m + k + 1 = m + 1 + k = \vdash\!\!\!\dashv(m) + k$ e, portanto, $\vdash\!\!\!\dashv(m) \leq n$. Assim, concluímos que $\vdash\!\!\!\dashv(m) \in I$ e, portanto, que $I = N$. Assim, fica provado que $\leq$ é uma ordem total. ■

Dessa forma, a relação binária $<$ fica definida como a ordem estrita associada a $\leq$.

⊢ **Teorema 5.14** (Boa ordenação)**.** *Seja $\boldsymbol{N}$ um modelo de números naturais. Então $(\boldsymbol{N}, \leq)$ é bem ordenado.*

□ *Demonstração.* Seja $C \subseteq N$ um conjunto que não tem menor elemento. Devemos mostrar que $C = \emptyset$. Notemos que $0 \notin C$ porque, para todo $n \in C$, $0 \leq n$, o que implicaria que $0 = \min C$. Consideremos $I := \{m \in N : \forall n \in C \quad m < n\}$. Inicialmente, ressaltemos que $C \cap I = \emptyset$, pois, se existe $m \in I \cap C$, então, como $m \in I$, para todo $n \in C$, $m < n$ e, como $m \in C$, segue que $m < m$, o que é absurdo. Então notemos que $0 \in I$, pois $0 \leq n$ para todo $n \in C$ e $0 \notin C$. Agora, seja $m \in I$. Então, para todo $n \in C$, $m < n$, o que implica que existe $d \in N \setminus \{0\}$ tal que $m + d = n$. Então segue que existe $k \in N$ tal que $d = \mathrm{H}(k) = k + 1$ e segue que $\mathrm{H}(m) + k = m + k + 1 = n$; ou seja, $\mathrm{H}(m) \leq n$. Agora notemos que $\mathrm{H}(m) \notin C$, pois, caso contrário, $\mathrm{H}(m) = \min C$. Portanto, para todo $n \in C$, $\mathrm{H}(m) < n$, o que mostra que $\mathrm{H}(m) \in I$ e, por sua vez, que $I = N$. Como $C \subseteq N$, segue que $C \cap N = C$. Mas então $\emptyset = C \cap I = C \cap N = C$. ■

⊢ **Teorema 5.15** (Indução completa)**.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Para todo conjunto* $I \subseteq N$*, se* $0 \in I$ *e*

$$\{m \in N : m < n\} \subseteq I \quad \Longrightarrow \quad \mathrm{H}(n) \in I,$$

então $I = N$*.*

□ *Demonstração.* Seja $I \subseteq N$ e suponha que $0 \in I$ e $\{m \in N : m < n\} \subseteq I \quad \Longrightarrow \mathrm{H}(n) \in I$. Então ■

⊢ **Lema 5.16.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. Então*

$$\forall n_1, n_2, m_1, m_2 \in N \qquad \begin{cases} n_1 \leq m_1 \\ n_2 \leq m_2 \end{cases} \quad \Longrightarrow \quad \begin{cases} n_1 + n_2 \leq m_1 + m_2 \\ n_1 \times n_2 \leq m_1 \times m_2. \end{cases}$$

□ *Demonstração.* Para $i \in \{1, 2\}$, como $n_i \leq m_i$, existe $d_i \in N$ tal que $n_i + d_i = m_i$. Assim, segue que $n_1 + d_1 + n_2 + d_2 = m_1 + m_2$ e, portanto, $n_1 + n_2 \leq m_1 + m_2$. Ainda, segue que

$$m_1 \times m_2 = (n_1 + d_1) \times (n_2 + d_2) = (n_1 \times n_2) + (n_1 \times d_2) + (d_1 \times n_1) + (d_1 \times d_2)$$

e, portanto, $n_1 \times n_2 \leq m_1 \times m_2$. ■

5.1.2 Números inteiros

⊢ **Proposição 5.17.** *Seja* $\boldsymbol{N}$ *um modelo de números naturais. A relação binária* $\sim$ *em* $N \times N$ *definida por*

$$\forall n_1, n_2, m_1, m_2 \qquad (n_1, n_2) \sim (m_1, m_2) \quad \Longleftrightarrow \quad n_1 + m_2 = n_2 + m_1$$

é uma relação de equivalência.

□ *Demonstração.* Sejam $(n_1, n_2), (m_1, m_2), (k_1, k_2) \in N \times N$. Primeiro, notemos que $n_1 + n_2 = n_2 + n_1$, o que mostra que $(n_1, n_2) \sim (n_1, n_2)$. Segundo, notemos que, se $(n_1, n_2) \sim (m_1, m_2)$, então $n_1 + m_2 = n_2 + m_1$, o que implica que $m_1 + n_2 = m_2 + n_1$ e, portanto, que $(m_1, m_2) \sim (n_1, n_2)$. Terceiro, notemos que, se $(n_1, n_2) \sim (m_1, m_2)$ e $(m_1, m_2) \sim (k_1, k_2)$, então $n_1 + m_2 = n_2 + m_1$ e $m_1 + k_2 = m_2 + k_1$, o que implica que $n_1 + m_2 + m_1 + k_2 = n_2 + m_1 + m_2 + k_1$ e, portanto, que $n_1 + k_2 = n_2 + k_1$, logo $(n_1, n_2) \sim (k_1, k_2)$. ■

:⊢ **Definição 5.5.** Seja $\boldsymbol{N}$ um modelo de números naturais com a equivalência $\sim$. O *modelo de números inteiros* associado a $\boldsymbol{N}$ é o par $\boldsymbol{Z} = (\boldsymbol{N}, Z)$, em que Z é o conjunto

$$Z := N \times N \big/ \sim,$$

o *conjunto dos números inteiros.*

⊢ **Proposição 5.18.** *Seja $\boldsymbol{Z}$ um modelo de números inteiros. Para todo $z \in Z$, existe único $d \in N$ tal que $z = [(n + d, n)]$ ou $z = [(n, n + d)]$.*

□ *Demonstração.* Seja $z \in Z$. Então $z = [(n_1, n_2)]$. Notemos que $n_1 \leq n_2$ ou $n_1 \geq n_2$. Agora, devemos notar que isso está bem definido para qualquer representante de z. Sejam $(n_1, n_2), (n_1', n_2') \in z$. Então $n_1 + n_2' = n_2 + n_1'$. Sem perda de generalidade, consideremos que $n_1 \geq n_2$. Nesse caso, existe $d \in N$ tal que $n_1 = n_2 + d$. Mas isso implica que $n_2 + d + n_2' = n_2 + n_1'$ e, portanto, que $n_1' = n_2' + d$ e, então $n_1' \geq n_2'$. Do mesmo modo, supondo $n_1' \geq n_2'$ achamos que $n_1 \geq n_2$. Ainda, o valor d é o mesmo em ambos os casos. Assim, se $n_1 \geq n_2$, temos que $z = [(n + d, n)]$ e, caso contrário, que $z = [(n, n + d)]$. A unicidade de d é óbvia pois, se existem d_1, d_2 tais que $n_1 = n_2 + d_1$ e $n_1 = n_2 + d_2$, então segue que $n_2 + d_1 = n_2 + d_2$ e, portanto, que $d_1 = d_2$. ■

Pela proposição anterior, um número inteiro de $\boldsymbol{Z}$ é unicamente representado pelo elemento $d \in N$ e sua posição no par ordenado. Por isso, se $z = [(n + d, n)]$, identificamos z com d e, se $z = [(n, n + d)]$, identificamos z com $-d$.

5.1.2.1 Adição e subtração

:⊢ **Definição 5.6.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. O *zero* de $\boldsymbol{Z}$ é o elemento $0 := [(n, n)]$.

:⊢ **Definição 5.7.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. A *adição nos números inteiros* é a função

$$\begin{aligned} +\colon Z \times Z &\longrightarrow Z \\ ([(n_1, n_2)], [(m_1, m_2)]) &\longmapsto [(n_1 + m_1, n_2 + m_2)]. \end{aligned}$$

Dados $n, m \in Z$, o número $n + m$ é a *soma de n e m.*

$\vdash$ **Teorema 5.19.** *Seja $\boldsymbol{Z}$ um modelo de números inteiros. A função $+$ está bem definida.*

$\square$ *Demonstração.* Sejam $n, m \in Z$ e $(n_1, n_2), (n_1', n_2') \in n$, $(m_1, m_2), (m_1', m_2') \in m$. Então $n + m$ pode ser calculado por

$$\begin{aligned}[(n_1, n_2)] + [(m_1, m_2)] &= [(n_1 + m_1, n_2 + m_2)] \\ [(n_1', n_2')] + [(m_1', m_2')] &= [(n_1' + m_1', n_2' + m_2')].\end{aligned}$$

Como $n_1 + n_2' = n_2 + n_1'$ e $m_1 + m_2' = m_2 + m_1'$, segue que

$$n_1 + n_2' + m_1 + m_2' = n_2 + n_1' + m_2 + m_1'$$

e, portanto, $(n_1 + m_1, n_2 + m_2) \sim (n_1' + m_1', n_2' + m_2')$, o que mostra que a soma $n + m$ está bem definida. ■

$\vdash$ **Proposição 5.20.** *Seja $\boldsymbol{Z}$ um modelo de números inteiros. Então*

1. $\forall n \in Z \qquad n + 0 = n;$
2. $\forall n, m, k \in Z \qquad (n + m) + k = n + (m + k);$
3. $\forall n, m \in Z \qquad n + m = m + n.$

$\square$ *Demonstração.* Sejam $n, m, k \in Z$ e $(n_1, n_2) \in n, (m_1, m_2) \in m, (k_1, k_2) \in k$.

1. Como $(0, 0) \in 0$, então $(n_1, n_2) + (0, 0) = (n_1, n_2)$, logo $n + 0 = n$.
2. Notemos que

$$\begin{aligned}((n_1, n_2) + (m_1, m_2)) + (k_1, k_2) &= (n_1 + m_1, n_2 + m_2) + (k_1, k_2) \\ &= (n_1 + m_1 + k_1, n_2 + m_2 + k_2) \\ &= (n_1, n_2) + (m_1 + k_1, m_2 + k_2) \\ &= (n_1, n_2) + ((m_1, m_2) + (k_1, k_2)),\end{aligned}$$

   logo $(n + m) + k = n + (m + k)$.
3. Notemos que

$$\begin{aligned}(n_1, n_2) + (m_1, m_2) &= (n_1 + m_1, n_2 + m_2) \\ &= (m_1 + n_1, m_2 + n_2) \\ &= (m_1, m_2) + (n_1, n_2),\end{aligned}$$

   logo $n + m = m + n$.

■

$:\vdash$ **Definição 5.8.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. A função *negativo* em $\boldsymbol{Z}$ é a função

$$\begin{aligned} - : Z &\longrightarrow Z \\ [(n_1, n_2)] &\longmapsto [(n_2, n_1)]. \end{aligned}$$

5.1.2.2 Multiplicação

A partir desta seção, usaremos a notação nm em vez de $n \times m$ para facilitar os cálculos.

:⊢ **Definição 5.9.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. O *um* de $\boldsymbol{Z}$ é o elemento $1 := [(n+1, n)]$.

:⊢ **Definição 5.10.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. A *multiplicação nos números inteiros* é a função

$$\begin{aligned} \times \colon Z \times Z &\longrightarrow Z \\ ([(n_1, n_2)], [(m_1, m_2)]) &\longmapsto [(n_1 m_1 + n_2 m_2, n_2 m_1 + n_1 m_2)]. \end{aligned}$$

Dados $n, m \in Z$, o número $n \times m$ é o *produto de n e m*.

⊢ **Teorema 5.21.** *Seja $\boldsymbol{Z}$ um modelo de números inteiros. A função $\times$ está bem definida.*

□ *Demonstração.* Sejam $n, m \in Z$ e $(n_1, n_2), (n_1', n_2') \in n$, $(m_1, m_2), (m_1', m_2') \in m$. Então $n \times m$ pode ser calculado por

$$\begin{aligned} [(n_1, n_2)] \times [(m_1, m_2)] &= [(n_1 m_1 + n_2 m_2, n_2 m_1 + n_1 m_2)] \\ [(n_1', n_2')] \times [(m_1', m_2')] &= [(n_1' m_1' + n_2' m_2', n_2' m_1' + n_1' m_2')]. \end{aligned}$$

Como $n_1 + n_2' = n_2 + n_1'$ e $m_1 + m_2' = m_2 + m_1'$, segue que

$$\begin{aligned} &(n_1 m_1 + n_2 m_2 + n_2' m_1' + n_1' m_2') + (n_1 m_2' + n_2 m_1' + n_1 m_1' + n_2 m_2') \\ &= n_1(m_1 + m_2') + n_2(m_2 + m_1') + (n_2' + n_1) m_1' + (n_1' + n_2) m_2' \\ &= n_1(m_1' + m_2) + n_2(m_2' + m_1) + (n_2 + n_1') m_1' + (n_1 + n_2') m_2' \\ &= (n_2 m_1 + n_1 m_2 + n_1' m_1' + n_2' m_2') + (n_1 m_2' + n_2 m_1' + n_1 m_1' + n_2 m_2'), \end{aligned}$$

o que implica que

$$n_1 m_1 + n_2 m_2 + n_2' m_1' + n_1' m_2' = n_2 m_1 + n_1 m_2 + n_1' m_1' + n_2' m_2'$$

e, portanto, $(n_1 m_1 + n_2 m_2, n_2 m_1 + n_1 m_2) \sim (n_1' m_1' + n_2' m_2', n_2' m_1' + n_1' m_2')$, o que mostra que o produto $n \times m$ está bem definido. ■

5.1.2.3 Ordenação

:⊢ **Definição 5.11.** Seja $\boldsymbol{Z}$ um modelo de números inteiros. A relação binária $\leq$ em N é definida por

$$[(n_1, n_2)] \leq [(m_1, m_2)] \quad \Longleftrightarrow \quad n_1 + m_2 \leq n_2 + m_1.$$

▷ **Exercício 5.1.** *Seja $\boldsymbol{Z}$ um modelo de números inteiros. A relação binária $\leq$ em N está bem definida e é uma relação de ordem total.*

5.2 Grupoides

Grupoides[1] são talvez a estrutura algébrica mais simples que se pode estudar. Eles consistem em um conjunto e uma operação binária nesse conjunto, sem que nenhuma propriedade adicional seja esperada dessa operação binária. O nome 'grupoide' vem por sua semelhança aos grupos, estruturas algébricas muito importantes que serão estudadas detalhadamente mais à frente. O sufixo '-oide' vem do grego e indica semelhança.

5.2.1 Operação binária e grupoide

Uma operação binária é uma lei de composição de 2 elementos de um conjunto que gera algum elemento do próprio conjunto.

:⊢ **Definição 5.12.** Seja X um conjunto. Uma *operação binária* em X é uma função

$$\begin{aligned} *\colon X \times X &\longrightarrow X \\ (x, x') &\longmapsto x * x'. \end{aligned}$$

A estrutura algébrica que consiste em um conjunto e uma estrutura binária é o objeto dessa seção. Nesse contexto, propriedades da operação binária podem ser melhor apresentadas e estudadas.

:⊢ **Definição 5.13.** Um *grupoide* é um par $\boldsymbol{G} = (G, *)$ em que G é um conjunto e $*\colon G \times G \longrightarrow G$ é uma operação binária em G.

► **Exemplo 5.1.** *Considere os seguintes exemplos.*

1. $(\mathbb{N}, +)$ *e* $(\mathbb{Z}, +)$ *são grupoides, em que* $+$ *é a adição dos números inteiros;*
2. $(\mathbb{N}, \times)$ *e* $(\mathbb{Z}, \times)$ *são grupoides, em que* $\times$ *é a multiplicação dos números inteiros;*
3. $(\mathbb{N}, \wedge)$ *e* $(\mathbb{Z}, \wedge)$ *são grupoides, em que* $\wedge$ *é o mínimo dos números inteiros;*
4. $(\mathbb{N}, \vee)$ *e* $(\mathbb{Z}, \vee)$ *são grupoides, em que* $\vee$ *é o máximo dos números inteiros.*

► **Exemplo 5.2.** *O jogo de pedra-papel-tesoura define um grupoide cuja operação é comutativa, mas não associativa. Representamos pedra (rocha) por* r*, papel por* p *e tesoura por* t*. A operação entre dois elementos distintos resulta no elemento que vence e entre um elemento com ele mesmo resulta nele mesmo. As regras de operação estão resumidas na tabela 5.1.*

[1] Essas estruturas são às vezes chamadas na literatura de 'magmas' para distingui-las de objetos da teoria de categorias também chamados 'grupoides'. Neste livro, o termo nunca será usado no sentido categórico.

$*$	r	p	t
r	r	p	r
p	p	p	t
t	r	t	t

TABELA 5.1: Tabela de pedra-papel-tesoura (r, p e t).

Por completude, definimos aqui a operação binária entre subconjuntos de um grupoide.

:⊢ **Definição 5.14.** Sejam $\boldsymbol{G}$ um grupoide e $S, S' \subseteq G$ subconjuntos. Denota-se

$$S * S' := \{s * s' \mid (s, s') \in S \times S'\}.$$

Para todo $x \in G$, denotam-se $x * S := \{x\} * S$ e $S * x := S * \{x\}$.

5.2.2 Associatividade

Existem 2 maneiras de operar 3 elementos $x_0, x_1, x_2 \in G$ de um grupoide mantendo a ordem dos índices:

$$(x_0 * x_1) * x_2,$$
$$x_0 * (x_1 * x_2).$$

Vamos no referir a cada maneira de escolher os parênteses nessas operações como uma *associação* dos elementos x_0, x_1, x_2. A *associatividade* é propriedade que afirma que ambas associações de 3 elementos produzem o mesmo elemento.

:⊢ **Definição 5.15** (Associatividade)**.** Seja X um conjunto. Uma operação binária *associativa* é uma operação binária $*\colon X \times X \longrightarrow X$ que satisfaz, para todos $x, x', x'' \in X$,

$$(x * x') * x'' = x * (x' * x'').$$

:⊢ **Definição 5.16.** Um *semigrupo* (ou *grupoide associativo*) é um grupoide $\boldsymbol{G} = (G, *)$ cuja operação binária $*$ é associativa.

► **Exemplo 5.3.** $(\mathbb{N}, +)$, $(\mathbb{N}, \times)$, $(\mathbb{N}, \wedge)$ *e* $(\mathbb{N}, \vee)$ *são semigrupos.*

Podemos generalizar essa propriedade considerando 4 elementos e todas as possíveis maneiras de operá-los. Existem 5 maneiras de operar 4 elementos x_0, x_1,

x_2, $x_3 \in G$ mantendo a ordem dos índices:

$$\begin{aligned}
&((x_0 * x_1) * x_2) * x_3,\\
&(x_0 * (x_1 * x_2)) * x_3,\\
&(x_0 * x_1) * (x_2 * x_3),\\
&x_0 * ((x_1 * x_2) * x_3),\\
&x_0 * (x_1 * (x_2 * x_3)).
\end{aligned}$$

Uma possível propriedade de associatividade generalizada seria estabelecer que essas 5 expressões são todas iguais. No entanto, como é de se esperar, a associatividade implica a associatividade generalizada, tanto para 4 elementos como para qualquer outra quantidade de elementos. Para demonstrar isso, devemos introduzir uma codificação para cada umas das associações de n elementos de um grupoide através da operação binária. Podemos estabelecer uma relação bijetiva entre um tipo específico de sequências finitas de 0s e 1s e associações de elementos. Chamaremos tais sequências de *sequências de associação*.

Notemos que, escrevendo cada uma das associações de 3 elementos usando a notação de função para $*$, obtemos

$$\begin{aligned}
&*\,(*(x_0, x_1), x_2)\\
&*\,(x_0, *(x_1, x_2))
\end{aligned}$$

e, escrevendo cada uma das associações de 4 elementos usando a notação de função para $*$, obtemos

$$\begin{aligned}
&*\,(*(*(x_0, x_1), x_2), x_3)\\
&*\,(*(x_0, *(x_1, x_2)), x_3)\\
&*\,(*(x_0, x_1), *(x_2, x_3))\\
&*\,(x_0, *(*(x_1, x_2), x_3))\\
&*\,(x_0, *(x_1, *(x_2, x_3))).
\end{aligned}$$

Se nos atentarmos à ordem em que os símbolos '$*$' e ',' aparecem nas expressões, obtemos as sequências

$$\begin{aligned}
&(0, 0, 1, 1)\\
&(0, 1, 0, 1)
\end{aligned}$$

e

$$\begin{gathered}(0,0,0,1,1,1)\\(0,0,1,0,1,1)\\(0,0,1,1,0,1)\\(0,1,0,0,1,1)\\(0,1,0,1,0,1).\end{gathered}$$

Podemos estabelecer uma regra indutiva de como gerar as próximas sequências a partir das anteriores da seguinte forma. Notemos que à associação vazia de 1 elemento $x_0 \in G$ podemos atribuir a sequência vazia $\emptyset \in \{0,1\}^0$ e, à única associação de 2 elementos $x_0, x_1 \in G$ — a saber $x_0 * x_1$ ou $*(x_0, x_1)$ — podemos atribuir a sequência $(0,1) \in \{0,1\}^2$, o que segue o mesmo padrão explicitado nos casos de 3 e 4 elementos. Então notemos que, com a notação de função, o primeiro símbolo a aparecer na expressão de uma associação a é sempre '$*$', e, como a operação $*$ é uma operação binária, devemos ter associações a_0 e a_1 de modo que a associação a será escrita $*(a_0, a_1)$. Isso significa que, dadas sequências $s^{(0)}$ e $s^{(1)}$ atribuídas às associações a_0 e a_1, podemos atribuir a sequência $s = (0, s^{(0)}, 1, s^{(1)})$ à associação a. Essa regra descreve indutivamente como atribuir uma sequência de 0s e 1s a cada associação diferente de elementos do grupoide.

O que acabou de ser apresentado foi uma motivação informal da relação entre a escolha de parênteses ao escrever uma expressão que envolve a operação entre elementos de um grupoide e uma sequência de 0s e 1s que a codifica. Definimos a seguir formalmente uma sequência de associação para então, a partir dessas sequências, definir uma operação entre vários elementos do grupoide.

:⊢ **Definição 5.17.** Define-se indutivamente para todo $n \in \mathbb{N} \setminus \{0\}$: Uma *sequência de n-associação* é

1. $(n = 1)$ a sequência vazia $\emptyset \in \{0,1\}^0$;
2. $(n > 1)$ uma sequência finita $s \in \{0,1\}^{2(n-1)}$ para a qual únicos $n_0, n_1 \in \mathbb{N} \setminus \{0\}$ e únicas sequências $s^{(0)} \in \{0,1\}^{2(n_0-1)}$ de n_0-associação e $s^{(1)} \in \{0,1\}^{2(n_1-1)}$ de n_1-associação satisfazem

$$s = (0, s^{(0)}, 1, s^{(1)}).$$

A sequência $s^{(0)}$ é a *subsequência formadora à esquerda* de s e a sequência $s^{(1)}$ é a *subsequência formadora à direita* de s.

Como consequência da definição, para toda sequência de n-associação $s = (0, s^{(0)}, 1, s^{(1)})$, vale que $1 + 2(n_0 - 1) + 1 + 2(n_1 - 1) + 1 = 2(n - 1)$, portanto

$$n_0 + n_1 = n.$$

Exibimos na tabela 5.2 as sequências de n-associação para n pequeno, ressaltando em vermelho e azul as subsequências $s^{(0)}$ e $s^{(1)}$ respectivas.

n	s	$(0, s^{(0)}, 1, s^{(1)})$
1	$\emptyset$	$\emptyset$
2	$(0,1)$	$(0,\emptyset,1,\emptyset)$
3	$(0,0,1,1)$	$(0,0,1,1,\emptyset)$
	$(0,1,0,1)$	$(0,\emptyset,1,0,1)$
4	$(0,0,0,1,1,1)$	$(0,0,0,1,1,1,\emptyset)$
	$(0,0,1,0,1,1)$	$(0,0,1,0,1,1,\emptyset)$
	$(0,0,1,1,0,1)$	$(0,0,1,1,0,1)$
	$(0,1,0,0,1,1)$	$(0,\emptyset,1,0,0,1,1)$
	$(0,1,0,1,0,1)$	$(0,\emptyset,1,0,1,0,1)$

TABELA 5.2: Sequências de n-associação $s \in \{0,1\}^{2(n-1)}$ e suas subsequências de associação formadoras para $n \in \{1,2,3,4\}$.

Pode-se mostrar que as sequências de associação são aquelas que satisfazem uma certa propriedade sobre a quantidade de entradas 1. Para isso, relembremos que denotamos, para todo $n \in \mathbb{N}$, $[n] = \{0, \ldots, n-1\}$ e lembremos também que uma sequência finita de 0s e 1s é uma função $s\colon [m] \longrightarrow \{0,1\}$, ou seja, um elemento $s \in \{0,1\}^m$. O conjunto dos índices i de s tais que $s_i = 1$ é

$$s^{-1}(1) = \{i \in [m] \mid s_i = 1\}\,.$$

▷ **Exercício 5.2.** *Seja $n \in \mathbb{N}$. Uma sequência $s \in \{0,1\}^{2(n-1)}$ é uma sequência de n-associação se, e somente se,*

1. $\#\left(s^{-1}(1)\right) = n-1$;
2. *Para todo $m \in [n-1]$,*
   - 2.1. $\#\left(s^{-1}(1) \cap [2m]\right) \leq m$;
   - 2.2. $\#\left(s^{-1}(1) \cap [2m+1]\right) \leq m$.

O exercício diz que uma sequência de n-associação deve ser composta por $n-1$ entradas 0 e $n-1$ entradas 1 de modo que suas primeiras entradas nunca consistam em mais entradas 1 do que entradas 0.

:⊢ **Definição 5.18.** Define-se indutivamente para todo $n \in \mathbb{N} \smallsetminus \{0\}$:
A *sequência de n-associação à esquerda* é

$$a_n := \begin{cases} \emptyset, & n = 1 \\ (0, a_{n-1}, 1, \emptyset), & n > 1 \end{cases}$$

A *sequência de n-associação à direita* é

$$z_n := \begin{cases} \emptyset, & n = 1 \\ (0, \emptyset, 1, z_{n-1}), & n > 1. \end{cases}$$

Essas sequências têm a forma para $n > 1$

$$a_n = ((0)^{n-1}, (1)^{n-1}) = (0, \ldots, 0, 1 \ldots, 1)$$

e

$$z_n = (0, 1)^{n-1} = (0, 1, \ldots, 0, 1).$$

Usando as sequências de associação, vamos definir indutivamente, como nos exemplos motivadores apresentados anteriormente, o operatório $\ast^s$ de n elementos do grupoide relativo a cada sequência de n-associação s.

:⊢ **Definição 5.19.** Sejam $\boldsymbol{G} = (G, *)$ um grupoide, $n \in \mathbb{N} \setminus \{0\}$, $(x_i)_{i\in[n]} \in G^n$ e $s \in \{0, 1\}^{2(n-1)}$ uma sequência de n-associação. O *s-operatório* de $(x_i)_{i\in[n]}$ é definido indutivamente por

$$\ast^s (x_i)_{i\in[n]} := \begin{cases} x_0, & s = \emptyset; \\ \left(\ast^{s^{(0)}} (x_i)_{i\in[n_0]}\right) * \left(\ast^{s^{(1)}} (x_{n_0+i})_{i\in[n_1]}\right), & s = (0, s^{(0)}, 1, s^{(1)}). \end{cases}$$

A tabela 5.3 mostra a relação entre as sequências de n-associação e as respectivas associações de n elementos para n pequeno.

n	$s = (0, s^{(0)}, 1, s^{(1)})$	$\ast^s (x_i)_{i\in[n]}$
1	$\emptyset$	x_0
2	$(0, \emptyset, 1, \emptyset)$	$x_0 * x_1$
3	$(0, 0, 1, 1, \emptyset)$	$(x_0 * x_1) * x_2$
	$(0, \emptyset, 1, 0, 1)$	$x_0 * (x_1 * x_2)$
4	$(0, 0, 0, 1, 1, 1, \emptyset)$	$((x_0 * x_1) * x_2) * x_3$
	$(0, 0, 1, 0, 1, 1, \emptyset)$	$(x_0 * (x_1 * x_2)) * x_3$
	$(0, 0, 1, 1, 0, 1)$	$(x_0 * x_1) * (x_2 * x_3)$
	$(0, \emptyset, 1, 0, 0, 1, 1)$	$x_0 * ((x_1 * x_2) * x_3)$
	$(0, \emptyset, 1, 0, 1, 0, 1)$	$x_0 * (x_1 * (x_2 * x_3))$

TABELA 5.3: Sequências de n-associação e suas respectivas associações de n elementos para $n \in \{1, 2, 3, 4\}$.

As sequências de associação a_n e z_n são as sequências que descrevem, sob certo ponto de vista, a maneira mais simples de como associar elementos. Os operatórios relativos a elas têm a forma

$$\begin{aligned}\ast^{a_n}(x_i)_{i\in[n]} &= (((x_0 * x_1)\cdots) * x_{n-1}),\\ \ast^{z_n}(x_i)_{i\in[n]} &= (x_0 * (\cdots(x_{n-2} * x_{n-1}))).\end{aligned}$$

Quando operamos elementos desse tipo um pelo outro, a associatividade permite transformar um essa operação em um operatório de um tipo só. A proposição a seguir se trata disso, e vai ser utilizada como lema para a demonstração da associatividade generalizada.

$\vdash$ **Proposição 5.22.** *Sejam* $\boldsymbol{G} = (G, *)$ *um semigrupo (grupoide associativo),* $n_0, n_1 \in \mathbb{N} \setminus \{0\}$ *e* $(x_i)_{i\in[n_0+n_1]} \in G^{n_0+n_1}$*. Então*

$$\ast^{a_{n_0+n_1}}(x_i)_{i\in[n_0+n_1]} = \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]}\right).$$

$\square$ *Demonstração.* Vamos demonstrar a proposição por indução em n_1. Para $n_1 = 1$,

$$\ast^{z_1}(x_{n_0+i})_{i\in[1]} = x_{n_0}$$

e, por definição da sequência a_{n_0+1} e do a_{n_0+1}-operatório,

$$\ast^{a_{n_0+1}}(x_i)_{i\in[n_0+1]} = \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * x_{n_0}$$

portanto

$$\begin{aligned}\ast^{a_{n_0+1}}(x_i)_{i\in[n_0+1]} &= \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * x_{n_0}\\ &= \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{z_1}(x_{n_0+i})_{i\in[1]}\right).\end{aligned}$$

Para o passo indutivo, suponhamos que a proposição vale para todo $m \in \{1, \ldots, n_1 - 1\}$ e mostremos que vale também para n_1. Pela hipótese do passo indutivo para $n_0 + 1$ e $n_1 - 1$,

$$\begin{aligned}\ast^{a_{n_0+n_1}}(x_i)_{i\in[n_0+n_1]} &= \ast^{a_{(n_0+1)+(n_1-1)}}(x_i)_{i\in[(n_0+1)+(n_1-1)]}\\ &= \left(\ast^{a_{n_0+1}}(x_i)_{i\in[n_0+1]}\right) * \left(\ast^{z_{n_1-1}}(x_{n_0+1+i})_{i\in[n_1-1]}\right).\end{aligned}$$

Mas, por definição da sequência z_{n_1} e do z_{n_1}-operatório, vale que

$$\ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]} = x_{n_0} * \left(\ast^{z_{n_1-1}}(x_{n_0+1+i})_{i\in[n_1-1]}\right).$$

Segue então da associatividade de $*$ que

$$\begin{aligned}\ast^{a_{n_0+n_1}}(x_i)_{i\in[n_0+n_1]} &= \left(\ast^{a_{n_0+1}}(x_i)_{i\in[n_0+1]}\right) * \left(\ast^{z_{n_1-1}}(x_{n_0+1+i})_{i\in[n_1-1]}\right)\\ &= \left(\left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * x_{n_0}\right) * \left(\ast^{z_{n_1-1}}(x_{n_0+1+i})_{i\in[n_1-1]}\right)\\ &= \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(x_{n_0} * \left(\ast^{z_{n_1-1}}(x_{n_0+1+i})_{i\in[n_1-1]}\right)\right)\\ &= \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]}\right).\end{aligned}$$

$\blacksquare$

Podemos finalmente enunciar a propriedade de associatividade generalizada que um grupoide associativo satisfaz.

⊢ **Proposição 5.23** (Associatividade generalizada)**.** *Sejam* $\boldsymbol{G} = (G, *)$ *um semigrupo (grupoide associativo),* $n \in \mathbb{N} \setminus \{0\}$*,* $(x_i)_{i\in[n]} \in G^n$ *e* $s, s' \in \{0,1\}^{2(n-1)}$ *sequências de* n*-associação. Então*

$$\ast^{s}(x_i)_{i\in[n]} = \ast^{s'}(x_i)_{i\in[n]}.$$

□ *Demonstração.* Vamos demonstrar uma versão simplificada da proposição para as sequências de associação à esquerda e à direita, a_n e z_n. Essa versão implicará a proposição. Mostraremos por indução em n que qualquer sequência de n-associação $s \in \{0,1\}^{2(n-1)}$ satisfaz

$$\ast^{s}(x_i)_{i\in[n]} = \ast^{a_n}(x_i)_{i\in[n]} = \ast^{z_n}(x_i)_{i\in[n]}. \tag{5.1}$$

Para isso, usaremos indução em n. Para $n = 1$, a equação 5.1 é válida já que a única sequência de 1-associação é $\emptyset = a_1 = z_1$. Suponhamos agora, para o passo indutivo, que a equação 5.1 valha para toda sequência de m-associação para $m \in \{1, \ldots, n-1\}$ e mostremos que a equação 5.1 vale também para todas sequências de n-associação. Seja $s \in \{0,1\}^{2(n-1)}$ uma sequência de n-associação e $s^{(0)} \in \{0,1\}^{2(n_0-1)}$ e $s^{(1)} \in \{0,1\}^{2(n_0-1)}$ suas subsequências de associação formadoras que satisfazem $s = (0, s^{(0)}, 1, s^{(1)})$. Como $n_0 + n_1 = n$ e $n_0, n_1 \in \mathbb{N} \setminus \{0\}$, então $n_0 < n$ e $n_1 < n$, e segue da hipótese de indução que

$$\begin{aligned}
\ast^{s^{(0)}}(x_i)_{i\in[n_0]} &= \ast^{a_{n_0}}(x_i)_{i\in[n_0]} \\
\ast^{s^{(1)}}(x_{n_0+i})_{i\in[n_1]} &= \ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]}.
\end{aligned}$$

Segue então da definição do operatório que

$$\begin{aligned}
\ast^{s}(x_i)_{i\in[n]} &= \left(\ast^{s^{(0)}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{s^{(1)}}(x_{n_0+i})_{i\in[n_1]}\right) \\
&= \left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]}\right).
\end{aligned}$$

Pela associatividade de $*$, segue da proposição 5.22 que

$$\left(\ast^{a_{n_0}}(x_i)_{i\in[n_0]}\right) * \left(\ast^{z_{n_1}}(x_{n_0+i})_{i\in[n_1]}\right) = \ast^{a_{n_0+n_1}}(x_i)_{i\in[n_0+n_1]},$$

o que implica

$$\ast^{s}(x_i)_{i\in[n]} = \ast^{a_{n_0+n_1}}(x_i)_{i\in[n_0+n_1]} = \ast^{a_n}(x_i)_{i\in[n]}.$$

Como isso vale para qualquer sequência de n-associação, segue que

$$\ast^{z_n}(x_i)_{i\in[n]} = \ast^{a_n}(x_i)_{i\in[n]}.$$

Isso conclui a demonstração da equação 5.1 e, finalmente, obtemos que, para todas sequência de n-associação $s, s' \in \{0,1\}^{2(n-1)}$ vale que

$$\ast^{s}(x_i)_{i\in[n]} = \ast^{a_n}(x_i)_{i\in[n]} = \ast^{s'}(x_i)_{i\in[n]}. \qquad \blacksquare$$

Essa proposição diz que, em um semigrupo (grupoide associativo), podemos colocar os parênteses como quisermos que o resultado será o mesmo. Sendo assim, definiremos o operatório de uma sequência finita de elementos do grupoide independente da escolha de parênteses, ou seja, da sequência de associação.

:⊢ **Definição 5.20.** Sejam $\boldsymbol{G} = (G, *)$ um semigrupo (grupoide associativo), $n \in \mathbb{N} \setminus \{0\}$ e $(x_i)_{i\in[n]} \in G^n$. O *operatório* de $(x_i)_{i\in[n]}$ é

$$\ast_{i\in[n]} x_i := \ast^{a_n}(x_i)_{i\in[n]} = (((x_0 * x_1)\cdots) * x_{n-1}).$$

Denota-se também $x_0 * \cdots * x_{n-1} := \ast_{i\in[n]} x_i$.

Seja $x \in G$. A *n-ésima potência* de x é

$$x^n := \ast_{i\in[n]} x.$$

Por causa da associatividade generalizada, a notação $x_0 * \cdots * x_{n-1}$ não gera ambiguidade. Essa definição considera que as operações vão sendo feitas à direita, mas uma mesma definição poderia ter sido feita para operações à esquerda.

O símbolo usado para a soma $+$ é o *somatório* $\sum$ e o símbolo usado para o produto $\times$ é o *produtório* $\prod$.

▷ **Exercício 5.3.** *Seja $\boldsymbol{G} = (G, *)$ um semigrupo (grupoide associativo). Para todos $n, n' \in \mathbb{N} \setminus \{0\}$ e todo $x \in G$,*

$$\begin{aligned} x^n * x^{n'} &= x^{n+n'} \\ (x^n)^{n'} &= x^{n\times n'}. \end{aligned}$$

5.2.3 Comutatividade

Existem 2 maneiras de operar 2 elementos $x_0, x_1 \in G$ de um grupoide sem se preocupar em manter a ordem dos índices:

$$\begin{aligned} &x_0 * x_1, \\ &x_1 * x_0. \end{aligned}$$

Vamos no referir a cada maneira de escolher a ordem dos índices nessas operações como uma *comutação* dos elementos x_0, x_1. A *comutatividade* é propriedade que afirma que ambas comutações de 2 elementos produzem o mesmo elemento.

:⊢ **Definição 5.21** (Comutatividade)**.** Seja X um conjunto. Uma operação binária *comutativa* é uma operação binária $*\colon X \times X \longrightarrow X$ que satisfaz, para todos $x, x' \in X$

$$x * x' = x' * x.$$

:⊢ **Definição 5.22.** Um *grupoide comutativo* é um grupoide $\boldsymbol{G} = (G, *)$ cuja operação binária $*$ é comutativa. Um semigrupo *comutativo* é um semigrupo cuja operação binária $*$ é comutativa.

Podemos generalizar essa propriedade considerando 3 elementos e todas as possíveis maneiras de operá-los sem se preocupar em manter a ordem dos índices. No entanto, se não supusermos a associatividade, essa tarefa será bem mais difícil, tendo que levar em conta, além da ordem dos índices, a associação específica escolhida (através de uma sequência de associação) para operar os elementos. Por isso, a partir de agora vamos supor associatividade.

Existem 6 maneiras de operar 3 elementos $x_0, x_1, x_2 \in G$ de um semigrupo (grupoide associativo) sem se preocupar em manter a ordem dos índices:

$$\begin{aligned}
&x_0 * x_1 * x_2,\\
&x_0 * x_2 * x_1,\\
&x_1 * x_0 * x_2,\\
&x_1 * x_2 * x_0,\\
&x_2 * x_0 * x_1,\\
&x_2 * x_1 * x_0.
\end{aligned}$$

Uma possível propriedade de comutatividade generalizada seria estabelecer que essas 6 expressões são todas iguais. No entanto, como é de se esperar, a comutatividade implica a comutatividade generalizada, tanto para 3 elementos como para qualquer outra quantidade de elementos. Para demonstrar isso, note que cada ordem de índices nas expressões acima é dada por uma bijeção do conjunto $\{0, 1, 2\}$ em si mesmo. Lembremos que o conjunto das bijeções de $[n]$ para $[n]$ é denotado $\overleftrightarrow{\mathcal{F}}(n)$.

⊢ **Proposição 5.24** (Comutatividade Generalizada)**.** *Sejam* $\boldsymbol{G} = (G, *)$ *um semigrupo comutativo e* $n \in \mathbb{N} \setminus \{0\}$*. Então, para toda bijeção* $\pi \in \overleftrightarrow{\mathcal{F}}(n)$*,*

$$\mathop{\ast}_{i \in [n]} x_{\pi(i)} = \mathop{\ast}_{i \in [n]} x_i.$$

□ *Demonstração.* A demonstração será por indução em n. Para $n = 1$, a única bijeção $\pi \in \overleftrightarrow{\mathcal{F}}(1)$ é a função identidade $\mathrm{I}\colon [1] \longrightarrow [1]$, portanto

$$\mathop{\ast}_{i\in[1]} x_{\mathrm{I}(i)} = \mathop{\ast}_{i\in[1]} x_i.$$

Para o passo indutivo, considere que vale para algum $n \in \mathbb{N} \smallsetminus \{0\}$ e seja $\pi \in \overleftrightarrow{\mathcal{F}}(n+1)$ uma bijeção. Definamos $k = \pi^{-1}(n)$ e a bijeção

$$\begin{aligned} \pi'\colon [n] &\longrightarrow [n] \\ i &\longmapsto \begin{cases} \pi(m) & i < k \\ \pi(m+1) & i > k. \end{cases} \end{aligned}$$

Da associatividade e da comutatividade de $*$ segue que

$$\begin{aligned} \mathop{\ast}_{i\in[n+1]} x_{\pi(i)} &= \mathop{\ast}_{i\in[k]} x_{\pi(i)} * x_{\pi(k)} * \mathop{\ast}_{i\in[n-k]} x_{\pi(i+k+1)} \\ &= \mathop{\ast}_{i\in[k]} x_{\pi(i)} * \mathop{\ast}_{i\in[n-k]} x_{\pi(i+k+1)} * x_{\pi(k)} \\ &= \mathop{\ast}_{i\in[k]} x_{\pi'(i)} * \mathop{\ast}_{i\in[n-k]} x_{\pi'(i+k)} * x_n \\ &= \mathop{\ast}_{i\in[n]} x_{\pi'(i)} * x_n. \end{aligned}$$

Finalmente, da hipótese de indução para π' segue que

$$\begin{aligned} \mathop{\ast}_{i\in[n+1]} x_{\pi(i)} &= \mathop{\ast}_{i\in[n]} x_{\pi'(i)} * x_n \\ &= \mathop{\ast}_{i\in[n]} x_i * x_n \\ &= \mathop{\ast}_{i\in[n+1]} x_i. \end{aligned}$$

■

▷ **Exercício 5.4.** *Seja* $\boldsymbol{G} = (G, *)$ *um semigrupo (grupoide associativo). Para todo* $n \in \mathbb{N} \smallsetminus \{0\}$ *e todos* $x, x' \in G$ *tais que* $x * x' = x' * x$,

$$(x * x')^n = x^n * (x')^n.$$

:⊢ **Definição 5.23** (Operatório de conjuntos)**.** Sejam $\boldsymbol{G} = (G, *)$ um semigrupo (grupoide associativo), $n \in \mathbb{N} \smallsetminus \{0\}$ e $(C_i)_{i\in[n]}$ uma família de subgrupos de G. Define-se

$$\mathop{\ast}_{i\in[n]} C_i := \left\{ \mathop{\ast}_{i\in[n]} x_i \;\middle|\; \forall_{i\in[n]} x_i \in C_i \right\}.$$

▷ **Exercício 5.5.** *Seja* $\boldsymbol{G} = (G, *)$ *um semigrupo (grupoide associativo).*

1. *Para todos subconjuntos* $S, S', S'' \subseteq G$,
$$(S * S') * S'' = S * (S' * S'').$$
2. *Se* $\boldsymbol{G}$ *é comutativo, para todos subconjuntos* $S, S' \subseteq G$ *vale*
$$S * S' = S' * S.$$

5.2.4 Subgrupoide

⊢ **Proposição 5.25** (Fechamento)**.** *Sejam* X *um conjunto,* $S \subseteq X$ *um subconjunto e* $*$ *uma operação binária em* X*. A restrição* $*|_{S\times S}$ *da operação binária* $*$ *a* $S \times S$ *é uma operação binária em* S *se, e somente se, para todos* $s, s' \in S$
$$s * s' \in S.$$

□ *Demonstração.* Basta notar que, como $S \subseteq X$, então $S \times S \subseteq X \times X$, e a proposição segue da proposição 3.4. ■

Denotamos $*|_{S\times S}$ por $*$ quando não há ambiguidade.

:⊢ **Definição 5.24.** Seja $\boldsymbol{G} = (G, *)$ um grupoide. Um *subgrupoide* de $\boldsymbol{G}$ é um grupoide $\boldsymbol{S} = (S, *_S)$ em que $S \subseteq G$ e $*_S = *|_{S\times S}$. Denota-se $\boldsymbol{S} \leq \boldsymbol{G}$. Um subgrupoide *próprio* de $\boldsymbol{G}$ é um grupoide $\boldsymbol{S} \leq \boldsymbol{G}$ em que S é um subconjunto próprio de G ($S \subset G$). Denota-se $\boldsymbol{S} < \boldsymbol{G}$.

⊢ **Proposição 5.26.** *Sejam* $\boldsymbol{G} = (G, *)$ *um grupoide e* $S \subseteq G$ *um conjunto tal que*

SM1. *(Fechamento) Para todos* $s, s' \in S$, $s * s' \in S$.

Então $\boldsymbol{S} = (S, *|_{S\times S}, 1)$ *é um grupoide. Ainda, se* $\boldsymbol{G}$ *é associativo, então* $\boldsymbol{S}$ *é associativo e, se* $\boldsymbol{G}$ *é comutativo, então* $\boldsymbol{S}$ *é comutativo.*

□ *Demonstração.* Denotemos $*_S := *|_{S\times S}$.

1. (Operação binária) Do fechamento, segue que $*_S$ é uma operação binária (5.25).
2. (Associatividade) Se $\boldsymbol{G}$ é associativo, segue que, para todos $s, s', s'' \in S$,
$$(s *_S s') *_S s'' = (s * s') * s'' = s * (s' * s'') = s *_S (s' *_S s'').$$
Logo $*_S$ é associativa.
3. (Comutatividade) Se $\boldsymbol{G}$ é comutativo, segue que, para todos $s, s' \in S$,
$$s *_S s' = s * s' = s' * s = s' *_S s.$$
Logo $*_S$ é comutativa. ■

5.2.5 Homomorfismo de grupoide

:⊢ **Definição 5.25.** Sejam $\boldsymbol{G} = (G, *)$ e $\boldsymbol{G}' = (G', *')$ grupoides. Um *homomorfismo de grupoide* de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é uma função $h\colon G \longrightarrow G'$ que satisfaz, para todos $x, x' \in G$

$$h(x * x') = h(x') *' h(x').$$

Denota-se $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$. O conjunto de todos homomorfismos de grupoide de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é denotado $\mathcal{H}(\boldsymbol{G}, \boldsymbol{G}')$.

Um *endomorfismo de grupoide* em $\boldsymbol{G}$ é um homomorfismo de grupoide de $\boldsymbol{G}$ para $\boldsymbol{G}$. O conjunto de todos endomorfismos de grupoide em $\boldsymbol{G}$ é denotado $\mathcal{H}(\boldsymbol{G})$.

A função identidade $\mathrm{I}\colon G \longrightarrow G$ é um homomorfismo de grupoide. Além disso, podemos gerar novos homomorfismos de grupoide compondo homomorfismos já conhecidos.

⊢ **Proposição 5.27** (Composição de homomorfismos)**.** *Sejam $\boldsymbol{G} = (G, *)$, $\boldsymbol{G}' = (G', *')$ e $\boldsymbol{G}'' = (G'', *'')$ grupoides e $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$ e $h'\colon \boldsymbol{G}' \longrightarrow \boldsymbol{G}''$ homomorfismos de grupoide. Então $h' \circ h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}''$ é homomorfismo de grupoide.*

□ *Demonstração.* Para todos $x, x' \in G$,

$$\begin{aligned}(h' \circ h)(x * x') &= h'(h(x * x')) \\ &= h'(h(x) *' h(x')) \\ &= h'(h(x)) *'' h'(h(x')) \\ &= (h' \circ h)(x) *'' (h' \circ h)(x').\end{aligned}$$

■

:⊢ **Definição 5.26.** Sejam $\boldsymbol{G} = (G, *)$ e $\boldsymbol{G}' = (G', *')$ grupoides. Um *isomorfismo de grupoide* de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é um homomorfismo de grupoide $h\colon G \longrightarrow G'$ tal que, para algum homomorfismo de grupoide $h'\colon G' \longrightarrow G$ vale

$$\begin{aligned}h' \circ h &= \mathrm{I}_G \\ h \circ h' &= \mathrm{I}_{G'}.\end{aligned}$$

O conjunto de todos isomorfismos de grupoide de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é denotado $\overleftrightarrow{\mathcal{H}}(\boldsymbol{G}, \boldsymbol{G}')$.

Um *automorfismo de grupoide* em $\boldsymbol{G}$ é um isomorfismo de grupoide de $\boldsymbol{G}$ para $\boldsymbol{G}$. O conjunto de todos automorfismos de grupoide em $\boldsymbol{G}$ é denotado $\overleftrightarrow{\mathcal{H}}(\boldsymbol{G})$.

▷ **Exercício 5.6.** *Sejam $\boldsymbol{G} = (G, *)$ e $\boldsymbol{G}' = (G', *')$ grupoides e $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$ homomorfismo de grupoide. Então h é um isomorfismo se, e somente se, h é bijetiva.*

5.3 Grupídeos

5.3.1 Identidades e grupídeo

Em um grupoide, pode-se ter um elemento que é neutro com relação à operação binária no sentido que qualquer elemento do grupoide é levado em si mesmo quando operado com esse elemento neutro. Esses elementos são chamados de 'identidades'.

:⊢ **Definição 5.27** (Identidade)**.** Seja $\boldsymbol{G} = (G, *)$ um grupoide.

- Uma *identidade à esquerda* de $\boldsymbol{G}$ é um elemento $i \in G$ que satisfaz, para todo $x \in G$,
$$i * x = x;$$
- Uma *identidade à direita* de $\boldsymbol{G}$ é um elemento $i \in G$ que satisfaz, para todo $x \in G$,
$$x * i = x;$$
- Uma *identidade* de $\boldsymbol{G}$ é um elemento $i \in G$ que satisfaz, para todo $x \in G$,
$$i * x = x = x * i.$$

► **Exemplo 5.4.** *Considere os seguintes exemplos.*

1. 0 *é uma identidade de* $(\mathbb{N}, +)$ *e* $(\mathbb{Z}, +)$*;*
2. 1 *é uma identidade de* $(\mathbb{N}, \times)$ *e* $(\mathbb{Z}, \times)$*;*
3. ∞ *é uma identidade de* $(\overset{\infty}{\mathbb{N}}, \wedge)$ *e* $(\overset{\infty}{\mathbb{Z}}, \wedge)$*, em que* $\overset{\infty}{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ *e* $\overset{\infty}{\mathbb{Z}} = \mathbb{Z} \cup \{\infty, -\infty\}$*;*
4. 0 *é uma identidade de* $(\mathbb{N}, \vee)$*;*
5. $-\infty$ *é uma identidade de* $(\overset{\infty}{\mathbb{Z}}, \vee)$*.*

As identidade são elementos importantes de qualquer grupoide e pode-se mostrar que, se existem identidades aos dois lados, elas são a mesma e a identidade nesse grupoide é única. Em particular, isso implica que identidades à esquerda ou à direita são únicas se a operação binária for comutativa.

⊢ **Proposição 5.28.** *Seja* $\boldsymbol{G} = (G, *)$ *um grupoide. Se* i *é uma identidade à esquerda e* i' *é uma identidade à direita de* $\boldsymbol{G}$*, então* $i = i'$ *e* i *é a única identidade de* $\boldsymbol{G}$*.*

□ *Demonstração.* Como i' é identidade à direita e i é identidade à esquerda de $\boldsymbol{G}$, segue que
$$i = i * i' = i'.$$
A unicidade segue pois qualquer identidade de $\boldsymbol{G}$ é identidade à esquerda. ■

:⊢ **Definição 5.28.** Um *grupídeo* é uma tripla $\boldsymbol{G} = (G, *, 1)$ em que $(G, *)$ é um grupoide e 1 é uma identidade de $(G, *)$.

A motivação para o nome 'grupídeo' vem da biologia, mas está ligada à estrutura algébrica em si. Na biologia, existe um conjunto de espécies chamado de superfamília dos Hominoides[2], que são primatas próximos aos humanos. Dentro dessa superfamília, há um subconjunto de espécies chamado de família dos Hominídeos[3], que contém orangotangos, gorilas, chipanzés e humanos.

Pode-se adotar por analogia o nome 'grupídeo' para um tipo específico de grupoide, aqueles que têm um elemento identidade. A nomenclatura é especialmente conveniente pois na palavra 'grup**íd**eo', as letras 'id' remetem ao fato de que a estrutura se trata de um grupoide com **id**entidade.

▷ **Exercício 5.7.** *Seja* $\boldsymbol{G} = (G, *, 1)$ *um grupídeo. Então*

$$G * G = G.$$

Uma propriedade interessante A seguinte proposição é conhecida como Eckmann–Hilton.

⊢ **Proposição 5.29.** *Sejam* G *um conjunto e* $(G, *, 1)$ *e* $(G, *', 1')$ *grupídeos (sobre o mesmo* G*) tais que*

- *(Lei de troca) para todos* $g_0, g_1, g_2, g_3 \in G$*,*

$$(g_0 * g_1) *' (g_2 * g_3) = (g_0 *' g_2) * (g_1 *' g_3).$$

Então $(G, *, 1) = (G, *', 1')$ *e* $*$ *é associativa e comutativa.*

□ *Demonstração.* A demonstração é direta.

- $(1 = 1')$ Da identidade e da lei da troca segue que

$$1 = 1 * 1 = (1' *' 1) * (1 *' 1') = (1' * 1) *' (1 * 1') = 1' *' 1' = 1'.$$

- $(* = *')$ Para todos $g, g' \in G$, da identidade, da lei da troca e de $1 = 1'$ segue que

$$g * g' = (g *' 1) * (1 *' g') = (g * 1) *' (1 * g) = g *' g'.$$

- (Associatividade) Para todos $g, g', g'' \in G$, da identidade, da lei da troca e de $(G, *, 1) = (G, *', 1')$ segue que

$$(g * g') * g'' = (g * g') * (1 * g'') = (g * 1) * (g' * g'') = g * (g' * g'').$$

- (Comutatividade) Para todos $g, g' \in G$, da identidade, da lei da troca e de $(G, *, 1) = (G, *', 1')$ segue que

$$g * g' = (1 * g) * (g' * 1) = (1 * g') * (g * 1) = g' * g.$$ ■

[2] O nome em latim é *Hominoidea* https://pt.wikipedia.org/wiki/Hominoidea.
[3] O nome em latim é *Hominidae* https://pt.wikipedia.org/wiki/Hominoidea.

5.3.2 Monoide (grupídeo associativo)

:⊢ **Definição 5.29.** Um *monoide* (ou *grupídeo associativo*) é um grupídeo $\boldsymbol{G} = (G, *, 1)$ cuja operação binária $*$ é associativa. Um grupídeo *comutativo* é um grupídeo cuja operação binária $*$ é comutativa.

Equivalentemente, um monoide é uma tripla $\boldsymbol{G} = (G, *, 1)$ em que $(G, *)$ é um semigrupo e 1 é identidade de $(G, *)$.

► **Exemplo 5.5.** *Considere os seguintes exemplos.*

1. $(\mathbb{N}, +)$ *e* $(\mathbb{Z}, +)$ *são monoides comutativos;*
2. $(\mathbb{N}, \times, 1)$ *e* $(\mathbb{Z}, \times, 1)$ *são monoides comutativos;*
3. $(\overset{\infty}{\mathbb{N}}, \wedge, \infty)$ *e* $(\overset{\infty}{\mathbb{Z}}, \wedge, \infty)$ *são monoides comutativos;*
4. $(\mathbb{N}, \vee, 0)$ *é monoide comutativo;*
5. $(\mathbb{Z} \cup \{-\infty, +\infty\}, \vee, -\infty)$ *é monoide comutativo.*

Notação. Denotaremos a identidade de um monoide $\boldsymbol{G}$ por 1_G quando houver ambiguidade.

:⊢ **Definição 5.30.** Seja $\boldsymbol{G} = (G, *, 1)$ um monoide (grupídeo associativo). O *operatório vazio* é

$$\mathop{\ast}_{i \in [0]} x_i := 1.$$

Seja $x \in G$. A 0-*ésima potência* de x é

$$x^0 = 1.$$

▷ **Exercício 5.8.** *Seja* $\boldsymbol{G} = (G, *, 1)$ *um monoide (grupídeo associativo).*

1. *Para todos* $n, n' \in \mathbb{N}$ *e todo* $x \in G$,

$$x^n * x^{n'} = x^{n+n'}$$
$$(x^n)^{n'} = x^{n \times n'}.$$

2. *Para todo* $n \in \mathbb{N}$ *e todos* $x, x' \in G$ *tais que* $x * x' = x' * x$,

$$(x * x')^n = x^n * (x')^n.$$

5.3.3 Subgrupídeo

:⊢ **Definição 5.31.** Seja $\boldsymbol{G} = (G, *, 1)$ um grupídeo. Um *subgrupídeo* de $\boldsymbol{G}$ é um grupídeo $\boldsymbol{S} = (S, *_S, 1_S)$ em que $S \subseteq G$, $*_S = *|_{S \times S}$ e $1_S = 1$. Denota-se $\boldsymbol{S} \leq \boldsymbol{G}$. Um subgrupídeo *próprio* de $\boldsymbol{G}$ é um grupídeo $\boldsymbol{S} \leq \boldsymbol{G}$ em que S é um subconjunto próprio de G ($S \subset G$). Denota-se $\boldsymbol{S} < \boldsymbol{G}$.

⊢ **Proposição 5.30.** *Sejam* $\boldsymbol{G} = (G, *, 1)$ *um grupídeo e* $S \subseteq G$ *um conjunto tal que*

SM1. *(Fechamento) Para todos* $s, s' \in S$, $s * s' \in S$;
SM2. *(Identidade)* $1 \in S$.

Então $\boldsymbol{S} = (S, *|_{S\times S}, 1)$ *é um grupídeo. Ainda, se* $\boldsymbol{G}$ *é associativo, então* $\boldsymbol{S}$ *é associativo e, se* $\boldsymbol{G}$ *é comutativo, então* $\boldsymbol{S}$ *é comutativo.*

□ *Demonstração.* Denotemos $*_S := *_{S\times S}$.

1. (Operação binária) Segue de (5.26).
2. (Identidade) Seja $s \in S$. Como $1 \in S$, da identidade de $*$ segue que

$$1 *_S s = 1 * s = s = s * 1 = s *_S 1.$$

   Logo 1 é identidade de $(S, *_S)$.
3. (Associatividade) Segue de (5.26).
4. (Comutatividade) Segue de (5.26). ■

Vale observar que somente sabendo que $\boldsymbol{S}$ é um grupídeo, isto é, que um subconjunto S de um grupídeo $\boldsymbol{G}$ com a operação do grupídeo restrita a esse subconjunto formam um grupídeo com algum elemento identidade, não podemos garantir que a identidade de $\boldsymbol{S}$ seja a mesma que a de $\boldsymbol{G}$.

5.3.4 Homomorfismos de grupídeo

:⊢ **Definição 5.32.** Sejam $\boldsymbol{G} = (G, *, 1)$ e $\boldsymbol{G}' = (G', *', 1')$ grupídeos. Um *homomorfismo de grupídeo* de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é uma função $h\colon G \longrightarrow G'$ que satisfaz

1. h é um homomorfismo de grupoide de $(G, *)$ para $(G', *')$:
   1.1. Para todos $g_0, g_1 \in G$, $h(g_0 * g_1) = h(g_0) *' h(g_1)$;
2. $h(1) = 1'$.

Denota-se $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$. O conjunto de todos homomorfismos de grupídeo de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é denotado $\mathcal{H}(\boldsymbol{G}, \boldsymbol{G}')$.

Um *endomorfismo de grupídeo* em $\boldsymbol{G}$ é um isomorfismo de grupídeo de $\boldsymbol{G}$ para $\boldsymbol{G}$. O conjunto de todos endomorfismos de grupídeo em $\boldsymbol{G}$ é denotado $\mathcal{H}(\boldsymbol{G})$.

Podemos notar que precisamos garantir que a função h leve a identidade de um grupídeo para a identidade de outro, já que isso não seria verdade se função fosse somente um homomorfismo de grupoide. No entanto, mesmo sem a segunda propriedade, um homomorfismo de grupoide entre grupos garante que a imagem da identidade do primeiro é a identidade do conjunto imagem. Veremos mais adiante que um homomorfismo de grupos é simplesmente um homomorfismo de grupoide, pois ele é suficiente para preservar a estrutura algébrica de grupos.

⊢ **Proposição 5.31** (Composição de homomorfismos)**.** *Sejam* $\boldsymbol{G} = (G, *, 1)$, $\boldsymbol{G}' = (G', *', 1')$ *e* $\boldsymbol{G}'' = (G'', *'', 1'')$ *grupídeos e* $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$ *e* $h'\colon \boldsymbol{G}' \longrightarrow \boldsymbol{G}''$ *homomorfismos de grupídeo. Então* $h' \circ h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}''$ *é homomorfismo de grupídeo.*

□ *Demonstração.*
1. Como homomorfismos de grupídeo são homomorfismos de grupoide, o resultado segue da proposição na seção de grupoides que afirma que composição de homomorfismos é homomorfismo (5.27).
2. Para mostrar que a identidade é preservada, basta notar que

$$(h' \circ h)(1) = h'(h(1)) = h'(1') = 1''. \qquad \blacksquare$$

:⊢ **Definição 5.33.** Sejam $\boldsymbol{G} = (G, *, 1)$ e $\boldsymbol{G}' = (G', *', 1')$ grupídeos. Um *isomorfismo de grupídeo* de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é um homomorfismo de grupídeo $h\colon G \longrightarrow G'$ tal que, para algum homomorfismo de grupídeo $h'\colon G' \longrightarrow G$ vale

$$\begin{aligned} h' \circ h &= \mathrm{I}_G \\ h \circ h' &= \mathrm{I}_{G'}. \end{aligned}$$

O conjunto de todos isomorfismos de grupídeo de $\boldsymbol{G}$ para $\boldsymbol{G}'$ é denotado $\overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{G}, \boldsymbol{G}')$.

Um *automorfismo de grupídeo* em $\boldsymbol{G}$ é um isomorfismo de grupídeo de $\boldsymbol{G}$ para $\boldsymbol{G}$. O conjunto de todos automorfismos de grupídeo em $\boldsymbol{G}$ é denotado $\overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{G})$.

▷ **Exercício 5.9.** *Sejam* $\boldsymbol{G} = (G, *, 1)$ *e* $\boldsymbol{G}' = (G', *', 1')$ *grupídeos e* $h\colon \boldsymbol{G} \longrightarrow \boldsymbol{G}'$ *um homomorfismo de grupídeo. Então* h *é um isomorfismo de grupídeo se, e somente se,* h *é bijetiva.*

5.4 Grupíneo

5.4.1 Inversos e grupíneo

:⊢ **Definição 5.34** (Inverso)**.** Sejam $\boldsymbol{G} = (G, *, 1)$ um grupídeo e $g \in G$.

- Um *inverso à esquerda* de g em $\boldsymbol{G}$ é um elemento $g' \in G$ que satisfaz

$$g' * g = 1;$$

- Um *inverso à direita* de g em $\boldsymbol{G}$ é um elemento $g' \in G$ que satisfaz

$$g * g' = 1;$$

- Um *inverso* de g em $\boldsymbol{G}$ é um elemento $g' \in G$ que satisfaz

$$g' * g = 1 = g * g'.$$

⊢ **Proposição 5.32.** *Sejam* $\boldsymbol{G} = (G, *, 1)$ *um grupídeo e* $g \in G$.

1. *Se* g *tem inverso à direita* g' *em* $\boldsymbol{G}$*, então* g *é inverso à esquerda de* g' *em* $\boldsymbol{G}$*;*
2. *Se* g *tem inverso à esquerda* g' *em* $\boldsymbol{G}$*, então* g *é inverso à direita de* g' *em* $\boldsymbol{G}$*;*
3. *Se* g *tem inverso* g' *em* $\boldsymbol{G}$*, então* g *é inverso de* g' *em* $\boldsymbol{G}$.

□ *Demonstração.* 1. Como g' é inverso à direita de g,

$$g * g' = 1,$$

o que mostra que g é inverso à esquerda de g'.

2. Como $g\diagdown$ é inverso à esquerda de g,

$$g' * g = 1,$$

o que mostra que g é inverso à direita de g'.

3. Segue dos itens anteriores. ■

⊢ **Proposição 5.33.** *Seja* $\boldsymbol{G} = (G, *, 1)$ *um monoide (grupídeo associativo).*

1. *Para todos* $g_0, g_1 \in G$ *que têm inversos à esquerda* g_0' *e* g_1'*, respectivamente,* $g_0 * g_1$ *tem inverso à esquerda* $g_1' * g_0'$*;*
2. *Para todos* $g_0, g_1 \in G$ *que têm inversos à direita* g_0' *e* g_1'*, respectivamente,* $g_0 * g_1$ *tem inverso à direita* $g_1' * g_0'$.

□ *Demonstração.* 1. Segue da associatividade que

$$(g_1' * g_0') * (g_0 * g_1) = g_1' * g_0' * g_0 * g_1 = g_1' * 1 * g_1 = g_1' * g_1 = 1.$$

2. Segue da associatividade que

$$(g_0 * g_1) * (g_1' * g_0') = g_0 * g_1 * g_1' * g_0' = g_0 * 1 * g_0' = g_0 * g_0' = 1. \quad ■$$

⊢ **Proposição 5.34.** *Seja* $\boldsymbol{G} = (G, *, 1)$ *um monoide (grupídeo associativo). Se* $g \in G$ *tem inverso à esquerda* g' *e inverso à direita* g'' *em* $\boldsymbol{G}$*, então* $g' = g''$ *e* g' *é o único inverso de* g *em* $\boldsymbol{G}$.

□ *Demonstração.* Como g' é inverso à esquerda e g'' é inverso à direita de g, segue da associatividade que

$$g' = g' * 1 = g' * (g * g'') = (g' * g) * g'' = 1 * g'' = g''.$$

A unicidade segue pois qualquer inverso de g em $\boldsymbol{G}$ é inverso à esquerda. ■

Notação. A unicidade do inverso nos permite denotar o inverso de um elemento $g \in G$ de algum modo fixo. Quando a operação binária é a adição (denotada $+$), denotamos o inverso por $-g$; quando é a multiplicação (denotada $\times$, $\cdot$, $\circ$), denotamos o inverso por g^{-1} (ou possivelmente por $\diagup g$, g^{-}). Para todo $g \in G$ com inverso $\diagup g$ e todo $g' \in G$, denota-se $g' \diagup g := g' * (\diagup g)$ por simplicidade. O conjunto de elementos invertíveis de G é denotado $G^{\diagup}$.

$\vdash$ **Proposição 5.35.** *Seja $(G, *, 1)$ um monoide (grupídeo associativo). Se $g \in G$ tem inverso $\diagup g$ em $\boldsymbol{G}$, então $\diagup g$ tem inverso e*

$$\diagup(\diagup g) = g.$$

$\square$ *Demonstração.* Como $\diagup g$ é inverso de g, então g é inverso de $\diagup g$ e, da unicidade do inverso em monoides segue que $\diagup(\diagup g) = g$. $\blacksquare$

$:\vdash$ **Definição 5.35** (Inversão)**.** Sejam $\boldsymbol{G} = (G, *, 1)$ um grupídeo.

- Uma *inversão à esquerda* de $\boldsymbol{G}$ é uma operação unária

$$\begin{aligned} \diagdown : G &\longrightarrow G \\ g &\longmapsto g\diagdown \end{aligned}$$

  tal que, para todo $g \in G$, $g\diagdown$ é o inverso à esquerda de g em $\boldsymbol{G}$. Para todos $g_0, g_1 \in G$, denota-se

$$g_0 \diagdown g_1 := (g_0\diagdown) * g_1;$$

- Uma *inversão à direita* de $\boldsymbol{G}$ é uma operação unária

$$\begin{aligned} \diagup : G &\longrightarrow G \\ g &\longmapsto \diagup g \end{aligned}$$

  tal que, para todo $g \in G$, $\diagup g$ é o inverso à direita de g em $\boldsymbol{G}$. Para todos $g_0, g_1 \in G$, denota-se

$$g_1 \diagup g_0 := (g_1\diagup) * g_0;$$

- Uma *inversão* de $\boldsymbol{G}$ é uma operação unária

$$\begin{aligned} \diagup : G &\longrightarrow G \\ g &\longmapsto g^{-1} \end{aligned}$$

  tal que, para todo $g \in G$, g^{-1} é o inverso de g em $\boldsymbol{G}$.

$:\vdash$ **Definição 5.36.** Um *grupíneo* é uma quádrupla $\boldsymbol{G} = (G, \times, \diagup, 1)$ em que $(G, \times, 1)$ é um grupídeo e $\diagup : G \longrightarrow G$ é uma inversão de $(G, \times, 1)$.

Como no caso de grupídeos, a motivação para o nome 'grupíneo' vem da biologia, mas está ligada à estrutura algébrica em si. Dentro da família dos Hominídeos, há um subconjunto de espécies chamado de subfamília dos Homininéos[4], que contém gorilas, chipanzés e humanos.

Da mesma forma, pode-se adotar por analogia o nome 'grupíneo' para um tipo específico de grupídeo, aqueles que têm uma inversão. A nomenclatura é novamente conveniente pois na palavra 'grup**íne**o', as letras 'in' remetem ao fato de que a estrutura se trata de um grupídeo com **in**versão.

Não estudaremos grupíneos aqui, pois os grupíneos associativos, chamados grupos, são estruturas muito importantes às quais dedicaremos um capítulo específico. No entanto, deixamos o seguinte exercício. Uma interação interessante entre identidade à direita e inversões à direita é que se ambas coexistem, então elas serão, respectivamente, uma identidade e uma inversão.

▷ **Exercício 5.10.** *Sejam* $\boldsymbol{G} = (G, *)$ *um semigrupo (grupoide associativo),* 1 *uma identidade à direita (ou à esquerda) de* $\boldsymbol{G}$ *e* $\nearrow : G \longrightarrow G$ *uma inversão à direita (ou à esquerda) de* $(G, *, 1)$*. Então* $(G, *, 1, \nearrow)$ *é um grupíneo.*

⊢ **Proposição 5.36** (Cancelamento)**.** *Seja* $\boldsymbol{G} = (G, *, 1)$ *um monoide (grupídeo associativo).*

1. *(Cancelamento à esquerda) Para todos* $g, g_0, g_1 \in G$ *tais que* g *tem inverso à esquerda* g'*, se* $g * g_0 = g * g_1$ *então* $g_0 = g_1$*;*
2. *(Cancelamento à direita) Para todos* $g, g_0, g_1 \in G$ *tais que* g *tem inverso à direita* g'*, se* $g_0 * g = g_1 * g$ *então* $g_0 = g_1$*.*

□ *Demonstração.* 1. Segue da associatividade que

$$\begin{aligned} g_0 &= 1 * g_0 \\ &= (g' * g) * g_0 \\ &= g' * (g * g_0) \\ &= g' * (g * g_1) \\ &= (g' * g) * g_1 \\ &= 1 * g_1 \\ &= g_1. \end{aligned}$$

[4] O nome em latim é *Homininae* https://pt.wikipedia.org/wiki/Hominidae#Subfam%C3%ADlia_Homininae.

2. Segue da associatividade que

$$\begin{aligned} g_0 &= g_0 * 1 \\ &= g_0 * (g * g') \\ &= (g_0 * g) * g' \\ &= (g_1 * g) * g' \\ &= g_1 * (g * g') \\ &= g_1 * 1 \\ &= g_1. \end{aligned}$$

■

Capítulo 6

Grupos

6.1 Conceitos básicos

6.1.1 Grupo e subgrupo

:⊢ **Definição 6.1.** Um *grupo* é uma quádrupla $\boldsymbol{G} = (G, \times, \diagup, 1)$ em que G é um conjunto, $\times\colon G \times G \longrightarrow G$ é uma operação binária associativa, $1 \in G$ é identidade de $(G, \times)$ e $\diagup\colon G \longrightarrow G$ é inversa[1] de $(G, \times, 1)$. Um grupo *comutativo* é um grupo cuja operação binária $\times$ é comutativa.

Isto é, as operações de um grupo satisfazem

G1. (Associatividade) Para todos $g, g', g'' \in G$,

$$(g \times g') \times g'' = g \times (g' \times g'');$$

G2. (Identidade) Para todo $g \in G$,

$$1 \times g = g = g \times 1;$$

G3. (Invertibilidade) Para todo $g \in G$, (denotando $g^{-1} := \diagup(g)$)

$$g^{-1} \times g = 1 = g \times g^{-1}.$$

e, no caso de um grupo comutativo,

G4. (Comutatividade) Para todos $g, g' \in G$,

$$g \times g' = g' \times g.$$

[1] O uso do símbolo '$\diagup$' para representar a função inversa de $\times$ não é usual, mas é bastante útil termos um símbolo para denotar essa função e o uso de uma barra como sinal de divisão é bastante comum. Pode-se pensar que, assim como o símbolo '$\times$' é o símbolo '$+$' rotacionado um oitavo de volta, o símbolo '$\diagup$' é o símbolo '$-$' rotacionado do mesmo tanto.

Notação. O símbolo da operação binária $\times$ será omitido sempre que possível, denotando-se $gg' := g \times g$. Além disso, a barra da inversa será substituída pelo expoente -1, denotando-se $g^{-1} := \diagup(g)$.

Quando o grupo é comutativo, em geral denotam-se a operação binária por $+$, a operação unária por $-$ e a constante por 0, de modo que o grupo seja denotado $(G, +, -, 0)$.

Em vista das definições introdutórias do capítulo anterior, um grupo é um grupíneo associativo, ou seja, um monoide $\boldsymbol{G} = (G, \times, 1)$ cujos elementos têm inverso sob $\times$ e um grupo comutativo é um grupo cuja operação binária $\times$ é comutativa.

► **Exemplo 6.1.** $(\mathbb{Z}, +, -, 0)$ *é um grupo.*

Como definido no capítulo anterior, a n-ésima potência de um elemento $g \in G$ é

$$g^n = \bigtimes_{i \in [n]} g.$$

No caso de um grupo, a inversa g^{-1} do elemento está definida, portanto podemos definir potências negativas de um elemento do grupo.

:⊢ **Definição 6.2.** Sejam $\boldsymbol{G} = (G, \times, \diagup, 1)$ um grupo, $n \in \mathbb{N} \setminus \{0\}$ e $g \in G$. A *n-ésima potência inversa* de g é

$$g^{-n} := (g^{-1})^n.$$

▷ **Exercício 6.1.** *Seja $\boldsymbol{G}$ um grupo.*

1. *Para todos $i, i' \in \mathbb{Z}$ e todo $g \in G$,*

$$g^i g^{i'} = g^{i+i'}$$
$$(g^i)^{i'} = g^{ii'}.$$

2. *Para todo $i \in \mathbb{Z}$ e todos $g, g' \in G$ tais que $gg' = g'g$,*

$$(gg')^i = g^i (g')^i.$$

3. *Para todos $g_0, \ldots, g_{n-1} \in G$,*

$$(g_0 \cdots g_{n-1})^{-1} = g_{n-1}^{-1} \cdots g_0^{-1}$$

▷ **Exercício 6.2** (Cancelamento)**.** *Seja $\boldsymbol{G}$ um grupo.*

1. *Para todos $g, g_0, g_0 \in G$, se $gg_0 = gg_1$ então $g_0 = g_1$.*

2. *Para todos $g, g_0, g_1 \in G$, se $g_0 g = g_1 g$ então $g_0 = g_1$.*

:⊢ **Definição 6.3.** Seja $\boldsymbol{G} = (G, \times, \diagup, 1)$ um grupo. Um *subgrupo* de $\boldsymbol{G}$ é um grupo $\boldsymbol{S} = (S, \times_S, \diagup_S, 1_S)$ em que $S \subseteq G$, $\times_S = \times|_{S\times S}$, $\diagup = \diagup|_S$ e $1_S = 1$. Denota-se $\boldsymbol{S} \leq \boldsymbol{G}$. Um subgrupo *próprio* de $\boldsymbol{G}$ é um subgrupo $\boldsymbol{S} \leq \boldsymbol{G}$ tal que S é um conjunto próprio de G ($S \subset G$). Denota-se $\boldsymbol{S} < \boldsymbol{G}$.

⊢ **Proposição 6.1.** *Sejam $\boldsymbol{G}$ um grupo e $S \subseteq G$ um conjunto que satisfaz*

SG1. *(Não-vacuidade) $S \neq \emptyset$;*
SG2. *(Fechamento) Para todos $s, s' \in S$, $ss' \in S$;*
SG3. *(Invertibilidade) Para todo $s \in S$, $s^{-1} \in S$.*

Então $\boldsymbol{S} = (S, \times|_{S\times S}, \diagup|_S, 1)$ é um grupo. Ainda, se $\boldsymbol{G}$ é comutativo, então $\boldsymbol{S}$ é comutativo.

□ *Demonstração.* Denotemos $\times_S := \times|_{S\times S}$.

1. (Operação binária) Segue de 5.30.
2. (Associatividade) Segue de 5.30.
3. (Identidade) Como $S \neq \emptyset$, existe $s \in S$, portanto $s^{-1} \in S$. Isso implica que $1 = s \times s^{-1} \in S$ e então segue de 5.30.
4. (Invertibilidade) Seja $s \in S$. Pela terceira propriedade de S, segue que o inverso de s em $\boldsymbol{G}$ pertence a S. Então
$$s^{-1} \times_S s = s^{-1} \times s = e = s \times s^{-1} = s \times_S s^{-1},$$
portanto s^{-1} é o inverso de s em $\boldsymbol{S}$.
5. (Comutatividade) Segue de 5.30.

■

⊢ **Proposição 6.2.** *Sejam $\boldsymbol{G}$ um grupo e $\mathcal{G}$ o conjunto dos subgrupos de $\boldsymbol{G}$. Então $(\mathcal{G}, \leq)$ é um conjunto parcialmente ordenado.*

□ *Demonstração.* Claramente, $\boldsymbol{G} \in \mathcal{G}$, portanto $\mathcal{G}$ não é vazio. Mostremos que $\leq$ é uma ordem parcial em $\mathcal{G}$.

- (Reflexividade) Seja $\boldsymbol{S} \in \mathcal{G}$. Então $\boldsymbol{S} \leq \boldsymbol{S}$, pois $\boldsymbol{S}$ é um grupo, $S \subseteq S$ e $\times|_{S\times S} = \times|_{S\times S}$.
- (Antissimetria) Sejam $\boldsymbol{S_1}, \boldsymbol{S_2} \in \mathcal{G}$ tais que $\boldsymbol{S_1} \leq \boldsymbol{S_2}$ e $\boldsymbol{S_2} \leq \boldsymbol{S_1}$. Então $S_1 \subseteq S_2$ e $S_2 \subseteq S_1$, o que implica $S_1 = S_2$, e portanto $\times|_{S_1\times S_1} = \times|_{S_2\times S_2}$, o que implica $\boldsymbol{S_1} = \boldsymbol{S_2}$.
- (Transitividade) Sejam $\boldsymbol{S_1}, \boldsymbol{S_2}, \boldsymbol{S_3} \in \mathcal{G}$ tais que $\boldsymbol{S_1} \leq \boldsymbol{S_2}$ e $\boldsymbol{S_2} \leq \boldsymbol{S_3}$. Então $S_1 \subseteq S_2$ e $S_2 \subseteq S_3$, o que implica $S_1 \subseteq S_3$ e, portanto, $\times_{S_1} = \times|_{S_3\times S_3}$.

■

▷ **Exercício 6.3.** *Seja* $\boldsymbol{G}$ *um grupo. Então* $\{\mathbf{1}\}$ *e* $\boldsymbol{G}$ *são subgrupos de* $\boldsymbol{G}$.

⊢ **Proposição 6.3.** *Sejam* $\boldsymbol{G}$ *um grupo,* $(\boldsymbol{S}_i)_{i\in I}$ *uma família de subgrupos de* $\boldsymbol{G}$ *e*

$$S := \bigcap_{i\in I} S_i.$$

Então $\boldsymbol{S}$ *é um subgrupo de* $\boldsymbol{G}$.

□ *Demonstração.* (SG1) Para todo $i \in I$, $\boldsymbol{S_i} \leq \boldsymbol{G}$, portanto $e \in N_i$. Logo $e \in N$. (SG2) Sejam $s_1, s_2 \in S$. Para todo $i \in I$, $s_1, s_2 \in S_i$. Como $\boldsymbol{S_i} \leq \boldsymbol{G}$, segue que $s_1 s_2 \in S_i$, o que implica que $s_1 s_2 \in S$. (SG3) Seja $s \in S$. Para todo $i \in I$, $s \in S_i$ e, como $\boldsymbol{S_i} \leq \boldsymbol{G}$, segue que $s^{-1} \in S_i$, o que implica que $s^{-1} \in S$. ■

⊢ **Proposição 6.4.** *Sejam* $\boldsymbol{G}$ *um grupo,* $(\boldsymbol{S}_i)_{i\in I}$ *uma família superiormente dirigida de subgrupos de* $\boldsymbol{G}$ *(para todos* $i_1, i_2 \in I$*, existe* $i \in I$ *tal que* $S_{i_1} \subseteq S_i$ *e* $S_{i_2} \subseteq S_i$*) e*

$$S := \bigcup_{i\in I} S_i.$$

Então $\boldsymbol{S}$ *é um subgrupo de* $\boldsymbol{G}$.

□ *Demonstração.* (Não-vacuidade) Para todo $i \in I$, $\boldsymbol{S_i} \leq \boldsymbol{G}$, portanto $e \in N_i$. Logo $e \in N$. (Fechamento) Sejam $s_1, s_2 \in S$. Existem $i_1, i_2 \in I$ tais que $s_1 \in S_{i_1}$ e $s_2 \in S_{i_2}$ e, pela propriedade, existe $i \in I$ tal que $S_{i_1} \subseteq S_i$ e $S_{i_2} \subseteq S_i$. Então $s_1, s_2 \in S_i$. Como $\boldsymbol{S_i} \leq \boldsymbol{G}$, segue que $s_1 s_2 \in S_i$, o que implica que $s_1 s_2 \in S$. (Invertibilidade) Seja $s \in S$. Existe $i \in I$ tal que $s \in N_i$ e, como $\boldsymbol{S_i} \leq \boldsymbol{G}$, segue que $s^{-1} \in S_i$, o que implica que $s^{-1} \in S$. ■

A propriedade definida acima é equivalente a dizer que a família $(S_i)_{i\in I}$ é um conjunto dirigido com respeito a $\subseteq$.

Lembremos que o conjunto dos elementos invertíveis de um monoide $\boldsymbol{M}$ é denotado por $M^{\diagup}$.

⊢ **Proposição 6.5.** *Seja* $\boldsymbol{M} = (M, \times, 1)$ *um monoide. Então existe uma operação unária* $\diagup : M^{\diagup} \longrightarrow M^{\diagup}$ *tal que* $\boldsymbol{M}^{\diagup} = (M^{\diagup}, \times|_{M^{\diagup}\times M^{\diagup}}, \diagup, 1)$ *é um grupo.*

□ *Demonstração.* Sejam $m_1, m_2 \in M^{\diagup}$. Então existem $m_1^{-1}, m_2^{-1} \in M$ tais que

$$m_1 m_1^{-1} = 1 \quad \text{e} \quad m_2 m_2^{-1} = 1.$$

Portanto

$$(m_1 m_2)(m_2^{-1} m_1^{-1}) = m_1 (m_2 m_2^{-1}) m_1^{-1} = m_1 m_1^{-1} = 1,$$

o que mostra que $m_1 m_2 \in M^{\diagup}$. Ainda, note que $1 \in M^{\diagup}$, pois $1 \times 1 = 1$. Como $M^{\diagup}$ é contém a identidade e é fechado sob $\times$, segue que $(M^{\diagup}, \times, 1)$ é um monoide (5.30). Por fim, $\boldsymbol{M}^{\diagup}$ é um grupo pois, por definição, todo elemento tem inverso e ele é único, o que define a operação $\diagup$. ■

6.1.2 Coclasses e índice de subgrupo

:⊢ **Definição 6.4.** Sejam $\boldsymbol{G}$ um grupo, $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo e $g \in G$. A *coclasse à esquerda* de $\boldsymbol{S}$ em $\boldsymbol{G}$ com *representante* g é o conjunto

$$gS := \{gs : s \in S\}.$$

A *coclasse à direita* de $\boldsymbol{S}$ em $\boldsymbol{G}$ com *representante* g é o conjunto

$$Sg := \{sg : s \in S\}.$$

Coclasses também são conhecidas como classes laterais. As definições de coclasses à esquerda ou à direita são análogas e, por consequência, toda definição ou proposição envolvendo uma das duas tem uma definição ou proposição dual envolvendo a outra. Por esse motivo, durante este capítulo consideraremos sempre coclasses à esquerda.

⊢ **Proposição 6.6.** *Sejam $\boldsymbol{G}$ um grupos, $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo e $g_1 g_2 \in G$. Então*

1. $g_1(g_2S) = (g_1g_2)S$ *e* $(Sg_1)g_2 = S(g_1g_2)$
2. $g_1S = g_2S \Leftrightarrow g_2^{-1}g_1 \in S$ *e* $Sg_1 = Sg_2 \Leftrightarrow g_2^{-1}g_1 \in S$.

□ *Demonstração.* 1. ($\subseteq$) Seja $g \in g_1(g_2S)$. Existem $g' \in g_2S$ tal que $g = g_1g'$ e, portanto, existe $s \in S$ tal que $g' = g_2s$. Mas então, pela associatividade, $g = g_1(g_2s) = (g_1g_2)s$, e segue que $g \in (g_1g_2)S$. ($\supseteq$) Seja $g \in (g_1g_2)S$. Então existe $s \in S$ tal que $g = (g_1g_2)s$ e, da associatividade, segue que $g = g_1(g_2s)$, logo $g \in g_1(g_2S)$. A outra igualdade é análoga.

2. ($\Leftarrow$) Seja $g \in g_1S = g_2S$. Então existem $s, s' \in S$ tais que $g = g_1s = g_2s'$, logo $g_2^{-1}g_1 = s's^{-1}$. Como $\boldsymbol{S}$ é subgrupo, segue que $g_2^{-1}g_1 = s's^{-1} \in S$.
($\Rightarrow$) ($\subseteq$) Se $g_2^{-1}g_1 \in S$, existe $s \in S$ tal que $g_2^{-1}g_1 = s$, portanto $g_1 = g_2s$, logo $g_1 \in g_2S$. Assim, dado $g \in g_1S$, existe $s' \in S$ tal que $g = g_1s'$. Como $\boldsymbol{S}$ é subgrupo, $ss' \in S$, logo $g = g_2ss' \in g_2S$. ($\supseteq$) Da mesma forma, como $\boldsymbol{S}$ é subgrupo, $g_1^{-1}g_2 = (g_2^{-1}g_1)^{-1} = s^{-1} \in S$, o que implica que $g_2 = g_1s^{-1} \in g_1S$. Assim, dado $g \in g_2S$, existe $s' \in S$ tal que $g = g_2s'$. Como $\boldsymbol{S}$ é subgrupo, $s^{-1}s' \in S$, logo $g = g_1s^{-1}s' \in g_2S$.

■

Essa proposição nos permite denotar os conjuntos acima simplesmente por g_1g_2S e Sg_1g_2, respectivamente.

A proposição a seguir mostra que as cardinalidades das coclasses de um subgrupo são sempre iguais.

⊢ **Proposição 6.7.** *Sejam $\boldsymbol{G}$ um grupo, $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo e $g_1, g_2 \in G$. Então*

$$\#(g_1S) = \#(g_2S).$$

□ *Demonstração.* Considere a relação

$$\begin{aligned} f\colon g_1S &\longrightarrow g_2S \\ g &\longmapsto g_2g_1^{-1}g. \end{aligned}$$

Vamos mostrar que f é função; isto é, está bem definida, que $g_1^{-1}g \in S$. Primeiro, note que, se $g \in g_1S$, existe $s \in S$ tal que $g = g_1s$. Então segue que $f(s) = g_2g_1^{-1}g = g_2g_1^{-1}g_1s = g_2s \in g_2S$, o que mostra que f está bem definida. Agora, note que a função

$$\begin{aligned} f^{-1}\colon g_2S &\longrightarrow g_1S \\ g &\longmapsto g_1g_2^{-1}g, \end{aligned}$$

que está bem definida pelo mesmo argumento de cima, é a inversa de f, pois $f^{-1} \circ f = 1_{g_1S}$ e $f \circ f^{-1} = 1_{g_2S}$. Isso mostra que f é uma bijeção. Portanto $\#(g_1S) = \#(g_2S)$.

■

⊢ **Proposição 6.8.** *Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo. A relação binária $\sim$ em G definida por*

$$g_1 \sim g_2 \quad \Longleftrightarrow \quad g_2^{-1}g_1 \in S$$

é uma relação de equivalência e suas classes de equivalência são as coclasses à esquerda (à direita) de $\boldsymbol{S}$ em $\boldsymbol{G}$.

□ *Demonstração.* Primeiro vamos demonstrar as três propriedades de relação de equivalência. (Reflexividade) Seja $g \in G$. Então $g^{-1}g = e \in S$, pois S é subgrupo. Logo $g \sim g$. (Simetria) Sejam $g_1, g_2 \in G$. Se $g_2^{-1}g_1 \in S$, como S é subgrupo, então $g_1^{-1}g_2 = (g_2^{-1}g_1)^{-1} \in S$. Logo $g_2 \sim g_1$. (Transitividade) Sejam $g_1, g_2, g_3 \in G$. Se $g_1 \sim g_2$ e $g_2 \sim g_3$, então $g_2^{-1}g_1 \in S$ e $g_3^{-1}g_2 \in S$. Como S é subgrupo, segue que $g_3^{-1}g_1 = g_3^{-1}g_2g_2^{-1}g_1 \in S$. Logo $g_1 \sim g_3$.

Agora, seja $g \in G$. Vamos mostrar que $[g] = gS$. Seja $s \in [g]$. Então $s \sim g$, o que implica que $g^{-1}s \in S$, que por sua vez implica que existe $s' \in S$ tal que $s' = g^{-1}s$ e, portanto, $s = gs'$. Logo $s \in gS$. Agora, seja $s \in gS$. Então existe $s' \in S$ tal que $s = gs'$, o que implica $g^{-1}s = s'$, que por sua ve implica $g^{-1}s \in S$ e, portanto, $s \sim g$. Logo $s \in [g]$. ■

Como $\sim$ é relação de equivalência, particiona G, e essas partições são as coclasses de $\boldsymbol{S}$ em $\boldsymbol{G}$. Assim, podemos considerar o conjunto G/S das classes de equivalências como o conjunto das coclasses de S em $\boldsymbol{G}$. É importante notar que esse conjunto ainda não possui estrutura de grupo. Isso será possível mais à frente, mas não para qualquer subgrupo, somente subgrupos que chamamos de normais.

:⊢ **Definição 6.5.** Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo. O *conjunto quociente* de $\boldsymbol{G}$ por $\boldsymbol{S}$ é o conjunto

$$G/S := G/\sim.$$

:⊢ **Definição 6.6.** Seja $\boldsymbol{G}$ um grupo e $\boldsymbol{S} \leq \boldsymbol{G}$ subgrupo. O *índice* de $\boldsymbol{S}$ em $\boldsymbol{G}$ é número cardinal

$$[G : S] := \#\left(G/S\right).$$

⊢ **Proposição 6.9.** *Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo. Então*

$$\#(G) = [G : S] \times \#(S).$$

□ *Demonstração.* O conjunto G/S é uma partição de G, pois é um conjunto quociente (3.19). Isso implica que G/S é uma família de conjuntos não vazios, disjuntos dois a dois, e que $G = \bigcup_{[g]\in G/S} gS$. Da terceira condição temos que

$$\#(G) = \#\left(\bigcup_{[g]\in G/S} gS\right).$$

Da segunda condição, temos por 4.3 que

$$\#\left(\bigcup_{[g]\in G/S} gS\right) = \#\left(\bigsqcup_{[g]\in G/S} gS\right).$$

Por fim, da primeira condição e do fato de que as coclasses de S têm a mesma cardinalidade, concluímos por 4.4 que

$$\#\left(\bigsqcup_{[g]\in G/S} gS\right) = \#\left(G/S\right) \times \#(S).$$

Disso segue que $\#(G) = [G : S] \times \#(S)$. ■

▷ **Exercício 6.4.** *Sejam $\boldsymbol{G}$ um grupo, $\boldsymbol{S} \leq \boldsymbol{G}$ um subgrupo e $\boldsymbol{S'} \leq \boldsymbol{S}$ um subgrupo. Então*

$$[G : S'] = [G : S][S : S'].$$

6.1.3 Subgrupo normal e grupo quociente

:⊢ **Definição 6.7.** Seja $\boldsymbol{G}$ um grupo. Um subgrupo *normal* de $\boldsymbol{G}$ é um subgrupo $\boldsymbol{N} \leq \boldsymbol{G}$ que satisfaz

SGN1. (Normalidade) Para todos $g \in G$ e $n \in N$, $gng^{-1} \in N$.

Denota-se $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$. Um subgrupo normal *próprio* de $\boldsymbol{G}$ é um subgrupo $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ que é um conjunto próprio de G: $N \subset G$. Denota-se $\boldsymbol{N} \triangleleft \boldsymbol{G}$.

$\vdash$ **Proposição 6.10.** *Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{N} \leq \boldsymbol{G}$ um subgrupo. São equivalentes:*

1. $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$;
2. *Para todo $g \in G$, $N = gNg^{-1}$;*
3. *para todo $g \in G$, $gN = Ng$;*
4. *Para todos $g_1, g_2 \in G$, $g_1 g_2 N = g_2 g_1 N$.*

$\vdash$ **Proposição 6.11.** *Seja $\boldsymbol{G}$ um grupo. Então $\{\boldsymbol{e}\}$ e $\boldsymbol{G}$ são subgrupos normais de $\boldsymbol{G}$.*

$\vdash$ **Proposição 6.12.** *Sejam $\boldsymbol{G}$ um grupo, $(\boldsymbol{N}_i)_{i \in I}$ uma família de subgrupos normais de $\boldsymbol{G}$ e*

$$N := \bigcap_{i \in I} N_i.$$

Então $\boldsymbol{N}$ é um subgrupo normal de $\boldsymbol{G}$.

$\square$ *Demonstração.* (Subgrupo) Como $(\boldsymbol{N}_i)_{i \in I}$ uma família de subgrupos de $\boldsymbol{G}$, segue que $\boldsymbol{N}$ é um subgrupo de $\boldsymbol{G}$ (6.3). (SGN1.) Sejam $g \in G$ e $n \in N$. Para todo $i \in I$, $n \in N_i$, portanto $gng^{-1} \in N_i$, o que implica $gng^{-1} \in N$. ■

$\vdash$ **Proposição 6.13.** *Sejam $\boldsymbol{G}$ um grupo, $(\boldsymbol{N}_i)_{i \in I}$ uma família superiormente dirigida de subgrupos normais de $\boldsymbol{G}$ (para todos $i_1, i_2 \in I$, existe $i \in I$ tal que $N_{i_1} \subseteq N_i$ e $N_{i_2} \subseteq N_i$) e*

$$N := \bigcup_{i \in I} N_i.$$

Então $\boldsymbol{N}$ é um subgrupo normal de $\boldsymbol{G}$.

$\square$ *Demonstração.* (Subgrupo) Como $(\boldsymbol{N}_i)_{i \in I}$ uma família superiormente dirigida de subgrupos de $\boldsymbol{G}$, segue que $\boldsymbol{N}$ é um subgrupo de $\boldsymbol{G}$ (6.4). (Normalidade) Sejam $g \in G$ e $n \in N$. Existe $i \in I$ tal que $n \in N_i$, portanto $gng^{-1} \in N_i$, o que implica $gng^{-1} \in N$. ■

$:\vdash$ **Definição 6.8.** Seja $\boldsymbol{G}$ um grupo e $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ um subgrupo normal. O *grupo quociente* de $\boldsymbol{G}$ por $\boldsymbol{N}$ é a quádrupla $\boldsymbol{G}/\boldsymbol{N} = (G/N, \times, 1N, ^{-1})$, em que G/N é o conjunto quociente de G pela relação de equivalência induzida por N,

$$\begin{aligned} \times \colon {}^{G}\!/_{N} \times {}^{G}\!/_{N} &\longrightarrow {}^{G}\!/_{N} \\ (g_1 N, g_2 N) &\longmapsto g_1 g_2 N \end{aligned}$$

e $(gN)^{-1} = g^{-1}N$.

Uma notação um pouco mais cuidadosa ressaltaria que as operações binárias de $\boldsymbol{G}$ e de $\boldsymbol{G/N}$ não são a mesma e, se denotarmos a primeira como $\times$ e a segunda como $\star$, teríamos a definição acima nos dando $g_1N \star g_2N := (g_1 \times g_2)N$. No entanto, como $\times$ de G/N está sempre relacionada a $\times$ de G, mantemos a mesma notação para ambas e a mesma convenção de omiti-la quando possível. Vale notar, também, que pela associatividade da $\times$ de $\boldsymbol{G}$, temos que $g_1g_2N = g_1(g_2N) = (g_1g_2)N$. O mesmo vale para a inversa $\diagup$.

$\vdash$ **Proposição 6.14.** *Seja $\boldsymbol{G}$ um grupo e $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ um subgrupo normal. Então $\boldsymbol{G/N}$ é um grupo.*

$\square$ *Demonstração.* Para simplificar as contas, usaremos a notação $[g] = gN$ quando conveniente. (Operação Binária) Devemos mostrar que a função definida acima está bem definida. Sejam $g_1, g_1', g_2, g_2' \in G$ tais que $g_1N = g_1'N$ e $g_2N = g_2'N$. Então

$$g_1g_2N = g_1Ng_2 = g_1'Ng_2 = g_1'Ng_2 = g_1'g_2N = g_1'g_2'N.$$

(Associatividade) Sejam $g_1, g_2, g_3 \in G$. Da associatividade da $\times$ de $\boldsymbol{G}$ segue que

$$([g_1][g_2])[g_3] = [g_1g_2][g_3] = [g_1g_2g_3] = [g_1][g_2g_3] = [g_1]([g_2][g_3]).$$

(identidade) Seja $g \in G$. Então

$$[1][g] = [1g] = [g] = [g1] = [g][1].$$

(Invertibilidade) Seja $g \in G$. Então

$$[g^{-1}][g] = [g^{-1}g] = [e] = [gg^{-1}] = [g][g^{-1}].$$

$\blacksquare$

6.1.4 Homomorfismo de grupo

$\vdash$ **Definição 6.9.** Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos. Um *homomorfismo de grupos* de $\boldsymbol{G_1}$ para $\boldsymbol{G_2}$ é uma função $h : G_1 \to G_2$ que satisfaz, para todos $g_1, g_2 \in G_1$,

$$h(g_1g_2) = h(g_1)h(g_2).$$

Denota-se $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$. O conjunto dos homomorfismos de $\boldsymbol{G_1}$ para $\boldsymbol{G_2}$ é denotado $\mathcal{H}(\boldsymbol{G_1}, \boldsymbol{G_2})$.

Note que a propriedade de homomorfismos de semigrupos e de grupos é a mesma.

⊢ **Proposição 6.15** (Homomorfismos preservam a estrutura algébrica)**.** *Sejam* $\boldsymbol{G_1}$ *e* $\boldsymbol{G_2}$ *grupos* $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ *um homomorfismo de grupos. Então*

1. $h(1_1) = 1_2$;
2. *Para todo* $g \in G_1$, $h(g)^{-1} = h(g^{-1})$.

□ *Demonstração.* 1.

$$\begin{aligned} h(1_1) &= h(1_1)1_2 \\ &= h(1_1)h(1_1)h(1_1)^{-1} \\ &= h(1_1 1_1)h(1_1)^{-1} \\ &= h(1_1)h(1_1)^{-1} \\ &= 1_2. \end{aligned}$$

2.

$$\begin{aligned} h(g)^{-1} &= h(g)^{-1}1_2 \\ &= h(g)^{-1}h(1_1) \\ &= h(g)^{-1}h(gg^{-1}) \\ &= h(g)^{-1}h(g)h(g^{-1}) \\ &= 1_2 h(g^{-1}) \\ &= h(g^{-1}). \end{aligned}$$

■

Essa proposição mostra que, como mencionado na seção de monoides, um homomorfismo de grupos é, de fato, um homomorfismo de monoides que preserva a inversa.

⊢ **Proposição 6.16** (Composição de homomorfismos)**.** *Sejam* $\boldsymbol{G_1}$, $\boldsymbol{G_2}$ *e* $\boldsymbol{G_3}$ *grupos e* $h_1 : \boldsymbol{G_1} \to \boldsymbol{G_2}$ *e* $h_2 : \boldsymbol{G_2} \to \boldsymbol{G_3}$ *homomorfismos de grupos. Então* $(h_2 \circ h_1) : \boldsymbol{G_1} \to \boldsymbol{G_3}$ *é um homomorfismo de grupos.*

□ *Demonstração.* Como um homomorfismo de grupos é um homomorfismo de semigrupos, o resultado segue da proposição na seção de semigrupos que afirma que composição de homomorfismos é homomorfismo (5.27). ■

⊢ **Proposição 6.17.** *Sejam* $\boldsymbol{G}$ *um grupo e* $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ *um subgrupo normal. A projeção canônica* $\pi : G \to G/N$, *definida por*

$$\begin{aligned} \pi \colon G &\longrightarrow {}^{G}\!/_{N} \\ g &\longmapsto gN, \end{aligned}$$

é um homomorfismo de grupos sobrejetivo.

□ *Demonstração.* Sejam $g_1, g_2 \in G$. Então, da definição de produto em $\boldsymbol{G/N}$, segue que

$$\pi(g_1 g_2) = g_1 g_2 N = g_1 N g_2 N = \pi(g_1)\pi(g_2).$$

(Sobrejetividade) Seja $gN \in G/N$. Então, $g \in G$, temos que $h(g) = gN$. ■

⊢ **Proposição 6.18.** *Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos e $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ um homomorfismo de grupos. Se $\boldsymbol{S} \leq \boldsymbol{G_2}$ é um subgrupo, então $\boldsymbol{h^{-1}(S)} \leq \boldsymbol{G_1}$ é um subgrupo, e se $\boldsymbol{N} \trianglelefteq \boldsymbol{G_2}$ é um subgrupo normal, então $\boldsymbol{h^{-1}(N)} \trianglelefteq \boldsymbol{G_1}$ é um subgrupo normal.*

□ *Demonstração.* (SG1) Como $e_2 \in S$ e h é homomorfismo segue que $h(e_1) = e_2$, portanto $e_1 \in h^{-1}(S)$. (SG2) Sejam $s_1, s_2 \in h^{-1}(S)$. Então $h(s_1), h(s_2) \in S$ e, como $\boldsymbol{S}$ é subgrupo, $h(s_1)h(s_2) \in S$. Logo, como h é homomorfismo, $h(s_1 s_2) = h(s_1)h(s_2) \in S$ e, portanto, $s_1 s_2 \in h^{-1}(S)$. (SG3) Seja $s \in h^{-1}(S)$. Então $h(s) \in S$ e, como $\boldsymbol{S}$ é subgrupo, $h(s)^{-1} \in S$. Como h é homomorfismo, segue que $h(s^{-1}) = h(s)^{-1} \in S$ e, portanto, $s^{-1} \in h^{-1}(S)$. (SGN1.) Sejam $g \in G_1$ e $n \in h^{-1}(N)$. Então $h(g) \in G_2$ e $h(n) \in N$. Como h é homomorfismo, segue que $h(gng^{-1}) = h(g)h(n)h(g)^{-1}$ e, como $\boldsymbol{N}$ é subgrupo normal, segue que $h(g)h(n)h(g)^{-1} \in N$. Logo $gng^{-1} \in h^{-1}(N)$. ■

⊢ **Proposição 6.19.** *Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos e $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ um homomorfismo de grupos. Se $\boldsymbol{S} \leq \boldsymbol{G_1}$, então $\boldsymbol{h(S)} \leq \boldsymbol{G_2}$, e se h é sobrejetivo e $\boldsymbol{N} \trianglelefteq \boldsymbol{G_1}$ um subgrupo normal, então $\boldsymbol{h(N)} \trianglelefteq \boldsymbol{G_2}$.*

□ *Demonstração.* (SG1) Como $e_1 \in S$, segue que $e_2 = h(e_1) \in h(S)$. (SG2) Sejam $s_1, s_2 \in h(S)$. Então existem $s_1', s_2' \in S$ tais que $h(s_1') = s_1$ e $h(s_2') = s_2$. Como $\boldsymbol{S}$ é subgrupo, segue que $s_1' s_2' \in S$ e, como h é homomorfismo, segue que

$$s_1 s_2 = h(s_1')h(s_2') = h(s_2' s_2') \in h(S).$$

(SG3) Seja $s \in h(S)$. Então existe $s' \in S$ tal que $h(s') = s$. Como $\boldsymbol{S}$ é subgrupo, segue que $s^{-1} \in S$ e, portanto, $h(s)^{-1} = h(s^{-1}) \in h(S)$. (SGN1.) Sejam $g \in G_2$ e $n \in h(N)$. Existe $n' \in N$ tal que $h(n') = n$ e, como h é sobrejetivo, existe $g' \in G_1$ tal que $h(g') = g$. Como N é normal, $gng^{-1} \in N$ e, como h é homomorfismo,

$$gng^{-1} = h(g')h(n')h(g')^{-1} = h(g'n'g'^{-1}) \in h(N).$$

■

6.1.5 Núcleo, imagem e isomorfismo

:⊢ **Definição 6.10.** Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos e $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ um homomorfismo de grupos. O *núcleo* de h é o conjunto

$$\mathrm{nuc}(h) := h^{-1}(1_2) = \{g \in G_1 \mid h(g) = 1_2\}$$

e a *imagem* de h é o conjunto

$$\mathrm{im}(h) := h(G_1) = \{g_2 \in G_2 \mid \exists g_1 \in G_1, h(g_1) = g_2\}.$$

$\vdash$ **Proposição 6.20.** *Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos e $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ um homomorfismo de grupos. Então h é injetiva se, e somente se,* $\mathrm{nuc}(h) = \{1_1\}$.

$\square$ *Demonstração.* ($\Rightarrow$) Suponha que h é injetiva. Seja $n \in \mathrm{nuc}(h)$. Então $h(n) = 1_2$. Mas $h(1_1) = 1_2$ e, como h é injetiva, concluímos que $n = 1_1$.
($\Leftarrow$) Suponha que $\mathrm{nuc}(h) = \{1_1\}$. Sejam $g_1, g_2 \in G_1$. Se $h(g_1) = h(g_2)$, temos que $h(g_1 g_2^{-1}) = h(g_1)h(g_2)^{-1} = 1_2$, o que implica que $g_1 g_2^{-1} = 1_1$, pois $\mathrm{nuc}(h) = \{1_1\}$. Logo $g_1 = g_2$, e concluímos que h é injetiva. $\blacksquare$

$:\vdash$ **Definição 6.11.** Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos. Um *isomorfismo de grupos* é um homomorfismo de grupos invertível. O conjunto de todos esses isomorfismos é denotado por $\overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{G_1}, \boldsymbol{G_2})$.

$\vdash$ **Proposição 6.21.** *Sejam $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ grupos e $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ um isomorfismo de grupos. Então $h^{-1} : \boldsymbol{G_2} \to \boldsymbol{G_1}$ é um isomorfismo de grupos.*

$\square$ *Demonstração.* Como h é bijetiva, sua inversa h^{-1} também é bijetiva. Sejam $g_1 g_1 \in G_2$. Como h é bijetiva, existem $g_1', g_2' \in G_1$ tais que $h(g_1') = g_1$ e $h(g_2') = g_2$. Assim, como h é homomorfismo, segue que

$$\begin{aligned} h^{-1}(g_1 g_2) &= h^{-1}(h(g_1')h(g_2')) \\ &= h^{-1}(h(g_1' g_2')) \\ &= g_1' g_2' \\ &= h^{-1}(g_1)h^{-1}(g_2). \end{aligned}$$

$\blacksquare$

$:\vdash$ **Definição 6.12.** Grupos *isomorfos* são grupos $\boldsymbol{G_1}$ e $\boldsymbol{G_2}$ para os quais existe isomorfismo $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$. Denota-se $\boldsymbol{G_1} \simeq \boldsymbol{G_2}$.

$\vdash$ **Proposição 6.22.** *Sejam $\boldsymbol{G_1}$, $\boldsymbol{G_2}$ e $\boldsymbol{G_3}$ grupos. Então*

1. *(Reflexividade)* $\boldsymbol{G_1} \simeq \boldsymbol{G_1}$;
2. *(Antissimetria)* $\boldsymbol{G_1} \simeq \boldsymbol{G_2} \Rightarrow \boldsymbol{G_2} \simeq \boldsymbol{G_1}$;
3. *(Transitividade)* $\boldsymbol{G_1} \simeq \boldsymbol{G_2}$ *e* $\boldsymbol{G_2} \simeq \boldsymbol{G_3} \Rightarrow \boldsymbol{G_1} \simeq \boldsymbol{G_3}$.

$\square$ *Demonstração.*
1. A função I_{G_1} é um isomorfismo de grupos.
2. A inversa é um isomorfismo de grupos (6.21).
3. A composição de homomorfismo é homomorfismo e a composição de bijeções é bijeção.

$\blacksquare$

Não dizemos que $\simeq$ é uma relação de equivalência porque ela não está propriamente definida em um conjunto, já que não existe o conjunto de todos os grupos por não existir o conjunto de todos os conjuntos. No entanto, é claro que o que a proposição afirma é que ela satisfaz as propriedades de uma relação de equivalência.

6.1.6 Teoremas de isomorfismo

$\vdash$ **Teorema 6.23** (1º teorema de isomorfismo)**.** *Sejam* $\boldsymbol{G_1}$ *e* $\boldsymbol{G_2}$ *grupos e* $h : \boldsymbol{G_1} \to \boldsymbol{G_2}$ *um homomorfismo de grupos. Então* $\mathbf{nuc}\,(\boldsymbol{h}) \trianglelefteq \boldsymbol{G_1}$, $\mathbf{im}\,(\boldsymbol{h}) \leq \boldsymbol{G_2}$ *e*

$$\boldsymbol{G_1}\big/_{\mathbf{nuc}(\boldsymbol{h})} \simeq \mathbf{im}\,(\boldsymbol{h}).$$

$\square$ *Demonstração.* Como $\mathrm{nuc}(h) = h^{-1}(e_2)$ e $\boldsymbol{\{e_2\}} \trianglelefteq \boldsymbol{G_2}$ (6.11), a proposição segue da proposição 6.18. Como $\mathrm{im}(h) = h(G_1)$ e $\boldsymbol{G_1} \leq \boldsymbol{G_1}$ (6.11), a proposição segue da proposição 6.18. Por causa disso, $\boldsymbol{G_1} \diagup \mathbf{nuc}\,(\boldsymbol{h})$ e $\mathbf{im}\,(\boldsymbol{h})$ são grupos. Consideremos a função

$$\begin{aligned} \eta\colon G_1 \diagup \mathrm{nuc}(h) &\longrightarrow \mathrm{im}(h) \\ g\,\mathrm{nuc}(h) &\longmapsto h(g). \end{aligned}$$

Primeiro mostremos que η é função. Sejam $g_1, g_2 \in G_1$ tais que $g_1\,\mathrm{nuc}(h) = g_2\,\mathrm{nuc}(h)$. Então $g_2^{-1}g_1 \in \mathrm{nuc}(h)$ (6.6), o que implica que $h(g_2^{-1}g_1) = e$. Como h é homomorfismo, $e = h(g_2^{-1}g_1) = h(g_2)^{-1}h(g_1)$, portanto $h(g_1) = h(g_2)$. Isso implica que $\eta(g_1\,\mathrm{nuc}(h)) = \eta(g_1\,\mathrm{nuc}(h))$.

Agora, mostremos que η é isomorfismo de grupos. Para simplificar as contas, denotamos $[g] = g\,\mathrm{nuc}(h)$. Primeiro mostramos que η é homomorfismo. Sejam $g_1, g_2 \in G_1$. Então

$$\eta([g_1][g_2]) = \eta([g_1g_2]) = h(g_1g_2) = h(g_1)h(g_2) = \eta([g_1])\eta([g_2]).$$

Por fim, devemos mostrar que η é bijetivo. (Injetividade) Seja $[g] \in \mathrm{nuc}(\eta)$. Então $\eta([g]) = 1_2$, logo $h(a) = 1_2$. Mas isso implica que $g \in \mathrm{nuc}(h)$. Portanto $[g] = [1_1]$, e segue que $\mathrm{nuc}(\eta) = \{[1_1]\}$, o que é equivalente à injetividade (6.20). (Sobrejetividade) Para todo $g \in \mathrm{im}(h)$, existe $g' \in G_1$ tal que $g = h(g')$. Mas $h(g') = \eta(g'\,\mathrm{nuc}(h))$, e segue a sobrejetividade. ■

$\vdash$ **Teorema 6.24** (2º teorema de isomorfismo)**.** *Sejam* $\boldsymbol{G}$ *um grupo,* $\boldsymbol{S} \leq \boldsymbol{G}$ *um subgrupo e* $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ *um subgrupo normal. Então* $\boldsymbol{SN} \leq \boldsymbol{G}$, $\boldsymbol{S} \cap \boldsymbol{N} \trianglelefteq \boldsymbol{S}$ *e*

$$\boldsymbol{S}\big/_{\boldsymbol{S} \cap \boldsymbol{N}} \simeq \boldsymbol{SN}\big/_{\boldsymbol{N}}.$$

$\vdash$ **Teorema 6.25** (3º teorema de isomorfismo)**.** *Sejam* $\boldsymbol{G}$ *um grupo e* $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ *um subgrupo normal. Então*

1. *Se* $\boldsymbol{S} \leq \boldsymbol{G}$ *tal que* $N \subseteq S \subseteq G$*, então* $\boldsymbol{S}/\boldsymbol{N} \leq \boldsymbol{G}/\boldsymbol{N}$*. Por outro lado, se* $\boldsymbol{S}' \leq \boldsymbol{G}/\boldsymbol{N}$*, existe* $\boldsymbol{S} \leq \boldsymbol{G}$ *tal que* $\boldsymbol{S}' = \boldsymbol{S}/\boldsymbol{N}$*.*
2. *Se* $\boldsymbol{S} \trianglelefteq \boldsymbol{G}$ *tal que* $N \subseteq S \subseteq G$*, então* $\boldsymbol{S}/\boldsymbol{N} \trianglelefteq \boldsymbol{G}/\boldsymbol{N}$*. Por outro lado, se* $\boldsymbol{S}' \trianglelefteq \boldsymbol{G}/\boldsymbol{N}$*, existe* $\boldsymbol{S} \leq \boldsymbol{G}$ *tal que* $\boldsymbol{S}' = \boldsymbol{S}/\boldsymbol{N}$*.*
3. *Se* $\boldsymbol{N}' \trianglelefteq \boldsymbol{G}$ *tal que* $N \subseteq N' \subseteq G$*, então*

$$(\boldsymbol{G}/\boldsymbol{N})\big/(\boldsymbol{N}'/\boldsymbol{N}) \simeq \boldsymbol{G}\big/\boldsymbol{N}'.$$

6.2 Construções categóricas

6.2.1 Produto de grupos

:⊢ **Definição 6.13.** Seja $(\boldsymbol{G_i})_{i\in I} = (G_i, \times_i, \diagup_i, 1_i,)_{i\in I}$ uma família de grupos. O *produto* da família $(\boldsymbol{G_i})_{i\in I}$ é a quádrupla

$$\prod_{i\in I} \boldsymbol{G_i} := (G, \times, \diagup, 1),$$

em que $G = \prod_{i\in I} G_i$,

$$\begin{aligned} \times \colon G \times G &\longrightarrow G \\ (g_1, g_2) &\longmapsto ((g_1)_i \times_i (g_2)_i)_{i\in I}. \end{aligned}$$

$1 := (1_i)_{i\in I}$ e $g^{-1} := (g_i^{-1})_{i\in I}$.

⊢ **Proposição 6.26.** *Seja $(\boldsymbol{G_i})_{i\in I}$ uma família de grupos. Então o produto $\prod_{i\in I} \boldsymbol{G_i}$ é um grupo. Se para todo $i \in I$ $\boldsymbol{G_i}$ é comutativo, então $\prod_{i\in I} \boldsymbol{G_i}$ é comutativo.*

□ *Demonstração.* (Associatividade) Sejam $g, g', g'' \in G$. Então, da associatividade de cada $\times_i$,

$$(gg')g'' = ((g_i g_i')g_i'')_{i\in I} = (g_i(g_i' g_i'')_{i\in I} = g(g'g'').$$

(Identidade) Para todo $g = (g_i)_{i\in I} \in G$,

$$1g = (1_i g_i)_{i\in I} = (g_i)_{i\in I} = g = (g_i)_{i\in I} = (g_i 1_i)_{i\in I} = g1.$$

(Invertibilidade) Seja $g \in G$. Então

$$g^{-1}g = (g_i^{-1} g_i)_{i\in I} = (1_i)_{i\in I} = (g_i g_i^{-1})_{i\in I} = gg^{-1}.$$

(Comutatividade) Sejam $g, g' \in G$. Então, da comutatividade de cada $\times_i$,

$$gg' = (g_i g_i')_{i\in I} = (g_i' g_i)_{i\in I} = g'g.$$

■

⊢ **Proposição 6.27.** *Seja $(\boldsymbol{G_i})_{i\in I}$ uma família de grupos. Para todo $i \in I$, a projeção canônica $\pi_i : \prod_{i\in I} G_i \to G_i$ é um homomorfismo de grupos.*

□ *Demonstração.* Sejam $g, g' \in \prod_{i\in I} G_i$. Então

$$\pi_i(gg') = \pi_i((g_i g_i')_{i\in I}) = g_i g_i' = \pi_i(g)\pi_i(g').$$

■

$\vdash$ **Proposição 6.28** (Propriedade Universal)**.** *Sejam* $(\boldsymbol{G_i})_{i\in I}$ *uma família de grupos,* $\boldsymbol{X}$ *um grupo e, para todo* $i \in I$*,* $h_i : \boldsymbol{X} \to \boldsymbol{G_i}$ *um homomorfismo de grupos. Então existe um único homomorfismo de grupos* $h : \boldsymbol{X} \to \prod_{i\in I} \boldsymbol{G_i}$ *tal que, para todo* $i \in I$*,* $\pi_i \circ h = h_i$ *(o diagrama comuta).*

$$\begin{tikzcd}
 & \prod\limits_{i\in I} \boldsymbol{G_i} \arrow[d, "\pi_i"] \\
\boldsymbol{X} \arrow[ur, dashed, "h"] \arrow[r, "h_i"'] & \boldsymbol{G_i}
\end{tikzcd}$$

$\square$ *Demonstração.* Defina a função

$$\begin{aligned} h\colon X &\longrightarrow \prod_{i\in I} G_i \\ x &\longmapsto (h_i(x))_{i\in I}. \end{aligned}$$

Da propriedade universal para o produto de conjuntos, h é a única função tal que, para todo $i \in I$, $\pi_i \circ h = h_i$. Basta mostrar que h é homomorfismo de grupos. Por simplicidade, apenas a operação $\times$ em G será explicitada. Sejam $x_1, x_2 \in X$. Então, como h_i são homomorfismos de grupo,

$$h(x_1x_2) = (h_i(x_1x_2))_{i\in I} = (h_i(x_1)h_i(x_2))_{i\in I} = h(x_1)h(x_2).$$

$\blacksquare$

6.2.2 Grupo livre

$:\vdash$ **Definição 6.14.** Seja C um conjunto. O *conjunto de inversos formais* de C é o conjunto $C^{-1} := C \times \{-1\}$ e seus elementos são denotados $c^{-1} := (c, -1)$.

Uma *palavra* em C é uma sequência finita $(c_1, \ldots, c_n) \in C^n$. Denota-se $c_1 \cdots c_n$.

Seja $p = c_1 \cdots c_n$ uma palavra em C. A *palavra inversa* de p é a palavra $p^{-1} := c_n^{-1} \cdots c_1^{-1}$.

$:\vdash$ **Definição 6.15.** Seja C um conjunto e p_1, p_2 palavras em $C \times \{1, -1\}$. A relação de equivalência entre as palavras p_1 e p_2 é definida por

$$p_1 \sim p_2 \quad \Longleftrightarrow \quad p_1p_2^{-1} \rightsquigarrow e.$$

$$C^* := \bigcup_{n\in\mathbb{N}} (C \times \{1, -1\})^n$$

Define-se $C^0 = \{\emptyset\}$.

$$\langle C\rangle := C^* \diagup \sim$$

A inclusão é definida.

$$\begin{aligned} \iota\colon C &\longrightarrow \langle C\rangle \\ c &\longmapsto [c]. \end{aligned}$$

⊢ **Proposição 6.29** (Propriedade Universal)**.** *Seja C um conjunto, $\boldsymbol{X} = (X, \star)$ um grupo e $f : C \to X$ uma função. Então existe um único homomorfismo de grupos $h : \langle \boldsymbol{C}\rangle \to \boldsymbol{X}$ tal que $h \circ \iota = f$ (o diagrama comuta).*

$$\begin{tikzcd} \langle \boldsymbol{C}\rangle \arrow[dr, dashed, "h"] & \\ C \arrow[u, "\iota"] \arrow[r, "f"'] & \boldsymbol{X} \end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} h\colon \langle C\rangle &\longrightarrow X \\ [c_1 \cdots c_n] &\longmapsto f(c_1) \star \cdots \star f(c_n), \end{aligned}$$

de modo que $h([e]) = e_X$. Então, para todo $c \in C$,

$$h \circ \iota(c) = h(\iota(c)) = h([c]) = f(c).$$

Logo $h \circ \iota = f$. Para mostrar que é um homomorfismo de grupos, sejam $[c_1 \cdots c_n]$, $[d_1 \cdots d_m] \in \langle C\rangle$. Então

$$\begin{aligned} h([c_1 \cdots c_n][d_1 \cdots d_m]) &= h([c_1 \cdots c_n d_1 \cdots d_m]) \\ &= f(c_1) \star \cdots \star f(c_n) \star f(d_1) \star \cdots \star f(d_m) \\ &= h([c_1 \cdots c_n]) \star h([d_1 \cdots d_m]). \end{aligned}$$

Isso mostra a existência. Para mostrar a unicidade, seja $\overline{h} : \langle \boldsymbol{C} \rangle \to \boldsymbol{X}$ um homomorfismo de grupos tal que $\overline{h} \circ \iota = f$. Seja $[c_1 \cdots c_n] \in \langle C \rangle$. Como $[c_1 \cdots c_n] = [c_1] \cdots [c_n] = \iota(c_1) \cdots \iota(c_n)$, segue que

$$\begin{aligned}
\overline{h}([c_1 \cdots c_n]) &= \overline{h}(\iota(c_1) \cdots \iota(c_n)) \\
&= \overline{h}(\iota(c_1)) \star \cdots \star \overline{h}(\iota(c_n)) \\
&= f(c_1) \star \cdots \star f(c_n) \\
&= h([c_1 \cdots c_n]),
\end{aligned}$$

o que implica que $\overline{h} = h$. ■

6.2.3 Coproduto de grupos

:⊢ **Definição 6.16.** Seja $(\boldsymbol{G_i})_{i \in I}$ uma família de grupos. O *coproduto* da família $(\boldsymbol{G_i})_{i \in I}$ é o par

$$\bigsqcup_{i \in I} \boldsymbol{G_i} := (G, \times),$$

em que $G := \langle \bigsqcup_{i \in I} G_i \rangle$ é o grupo livre sobre o coproduto de conjuntos $\bigsqcup_{i \in I} G_i$ e

$$\begin{aligned}
\times \colon G \times G &\longrightarrow G \\
([p_1], [p_2]) &\longmapsto [p_1 p_2].
\end{aligned}$$

⊢ **Proposição 6.30** (Propriedade Universal)**.** *Sejam $(\boldsymbol{G_i})_{i \in I}$ uma família de grupos, $\boldsymbol{X} = (X, \star)$ um grupo e, para todo $i \in I$, $h_i : \boldsymbol{G_i} \to \boldsymbol{X}$ um homomorfismo de grupos. Então existe um único homomorfismo de grupos $h : \bigsqcup_{i \in I} \boldsymbol{G_i} \to \boldsymbol{X}$ tal que, para todo $i \in I$, $h \circ \iota_i = h_i$ (o diagrama comuta).*

$$\begin{tikzcd}
\bigsqcup_{i \in I} \boldsymbol{G_i} \arrow[dr, dashed, "h"] & \\
\boldsymbol{G_i} \arrow[u, "\iota_i"] \arrow[r, "h_i"'] & \boldsymbol{X}
\end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned}
h \colon \left\langle \bigsqcup_{i \in I} G_i \right\rangle &\longrightarrow X \\
[(g_1, i_1) \cdots (g_n, i_n)] &\longmapsto h_{i_1}(g_1) \star \cdots \star h_{i_n}(g_n).
\end{aligned}$$

Por simplicidade, seja $G := \langle \bigsqcup_{i \in I} G_i \rangle$. Para mostrar que h é homomorfismo, sejam $[(g_1, i_1) \cdots (g_n, i_n)]$ e $[(g'_1, i'_1) \cdots (g'_m, i'_m)] \in G$. Então

$$\begin{aligned}
&h([(g_1, i_1) \cdots (g_n, i_n)][(g'_1, i'_1) \cdots (g'_m, i'_m)]) \\
&= h([(g_1, i_1) \cdots (g_n, i_n)(g'_1, i'_1) \cdots (g'_m, i'_m)]) \\
&= h_{i_1}(g_1) \star \cdots \star h_{i_n}(g_n) \star h_{i'_1}(g'_1) \star \cdots \star h_{i'_n}(g'_n) \\
&= h([(g_1, i_1) \cdots (g_n, i_n)]) \star h([(g'_1, i'_1) \cdots (g'_m, i'_m)]).
\end{aligned}$$

■

6.3 Construções específicas

6.3.1 Translações e conjugações

:⊢ **Definição 6.17.** Sejam $\boldsymbol{G}$ um grupo e $g \in G$. A *conjugação* por g é a função

$$\begin{aligned} C_g \colon G &\longrightarrow G \\ h &\longmapsto ghg^{-1}, \end{aligned}$$

a *translação à direita* por g é a função

$$\begin{aligned} D_g \colon G &\longrightarrow G \\ h &\longmapsto hg \end{aligned}$$

e a *translação à esquerda* por g é a função

$$\begin{aligned} E_g \colon G &\longrightarrow G \\ h &\longmapsto gh. \end{aligned}$$

⊢ **Proposição 6.31.** *Sejam $\boldsymbol{G}$ um grupo e $g \in G$. As funções $C.$, D e E são bijeções.*

□ *Demonstração.* Basta mostrarmos para D_g e E_g, pois $C_g = E_g \circ D_{g^{-1}}$. Primeiro, mostremos que as translações são homomorfismos. Para todos $g, g', h \in G$,

$$E_g \circ E_g(h) = gg'h = E_{gg'}(h)$$

e

$$D_g \circ D_{g'}(h) = hg'g = D_{g'g}(h),$$

o que mostra que $E_{gg'} = E_g \circ E_g$

Isso mostra que as translações à esquerda e à direita têm inversa $E_{g^{-1}}$ e $D_{g^{-1}}$, respectivamente, pois

$$E_g \circ E_{g^{-1}} = E_{gg^{-1}} = \mathrm{I} = E_{g^{-1}g} = E_{g^{-1}} \circ E_g$$

e

$$D_g \circ D_{g^{-1}} = D_{gg^{-1}} = \mathrm{I} = D_{g^{-1}g} = D_{g^{-1}} \circ D_g.$$

■

⊢ **Proposição 6.32.** *Sejam $\boldsymbol{G}$ um grupo e $g \in G$. A conjugação C_g é isomorfismo de grupo.*

□ *Demonstração.* Como C_g é bijeção, basta mostrar que é homomorfismo de grupo. Para todos h, $h' \in G$,

$$C_g(hh') = ghh'g^{-1} = ghg^{-1}gh'g^{-1} = C_g(h)C_g(h').$$

■

A *conjugação* em $\boldsymbol{G}$ é a função

$$\begin{aligned} C\colon G &\longrightarrow \overset{\leftrightarrow}{\mathcal{H}}(G) \\ g &\longmapsto C_g \end{aligned}$$

6.3.2 Centro, automorfismos internos e externos

:⊢ **Definição 6.18.** Seja $\boldsymbol{G}$ um grupo. O *centro* de $\boldsymbol{G}$ é o conjunto

$$\mathrm{Z}(G) := \{g \in G \mid \forall_{h \in G}\ C_g(h) = h\}$$

Notemos que, como C é homomorfismo de grupos e

$$\begin{aligned} \mathrm{Z}(G) &= \{g \in G \mid \forall_{h \in G}\ C_g(h) = h\} \\ &= \{g \in G \mid C_g = \mathrm{I}\} \\ &= \operatorname{nuc} C, \end{aligned}$$

o centro de $\boldsymbol{G}$ é um subgrupo normal de $\boldsymbol{G}$.

:⊢ **Definição 6.19.** Seja $\boldsymbol{G}$ um grupo. Um *automorfismo interno* de $\boldsymbol{G}$ é um automorfismo da forma C_g para algum $g \in G$.

Pode-se mostrar que um automorfismo $h\colon G \longrightarrow G$ é interno (ou seja, é conjugação por algum elemento de G) se, e somente se, para todo supergrupo $\boldsymbol{H} \geq \boldsymbol{G}$, esse automorfismo pode ser estendido para $\boldsymbol{H}$. A ida é evidente mas a volta é um pouco mais complicada [**art:Schupp-CharacterizationInnerAutomorphisms**].

$$1 \to \mathrm{Z}(G) \to G \to \overset{\leftrightarrow}{\mathcal{H}}(G) \to Out(G) \to 1$$

6.3.3 Centralizador e normalizador

6.3.4 Produto semidireto

:⊢ **Definição 6.20.** Sejam $\boldsymbol{G}$ e $\boldsymbol{H}$ grupos e $A\colon \boldsymbol{G} \curvearrowright \boldsymbol{H}$ uma ação do grupo $\boldsymbol{G}$ sobre o grupo $\boldsymbol{H}$. O *produto semidireto* (à esquerda) de $\boldsymbol{H}$ por $\boldsymbol{G}$ com respeito a A é o grupo

$$\boldsymbol{G} \ltimes_A \boldsymbol{H} := (G \times H, \times_{G \ltimes_A H}, /_{G \ltimes_A H}, 1_{G \ltimes_A H})$$

em que

$$\begin{aligned} \times_{G\ltimes_A H}\colon (G\times H)\times(G\times H) &\longrightarrow G\times H \\ ((g,h),(g',h')) &\longmapsto (gg', hA_g(h')), \end{aligned}$$

$$\begin{aligned} /_{G\ltimes_A H}\colon G\times H &\longrightarrow G\times H \\ (g,h) &\longmapsto (g^{-1}, A_{g^{-1}}(h^{-1})) \end{aligned}$$

e

$$1_{G\ltimes_A H} := (1_G, 1_H).$$

Por simplicidade, denotamos $G\ltimes H$ quando a ação está subentendida no contexto.

O produto semidireto à direita é definido analogamente e denotado $\boldsymbol{H}\rtimes_A \boldsymbol{G}$, com

$$\begin{aligned} \times_{H\rtimes G}\colon (H\times G)\times(H\times G) &\longrightarrow H\times G \\ ((h,g),(h',g')) &\longmapsto (A_g(h')h, gg') \end{aligned}$$

e as mesmas inversa e identidade do caso à esquerda. Vale ressaltar que com uma ação do grupo $\boldsymbol{G}$ sobre o grupo $\boldsymbol{H}$ queremos dizer um homomorfismo de grupos

$$A\colon \boldsymbol{G}\to \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{H}),$$

em que $\overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{H})$ é o grupo dos automorfismos de grupo de $\boldsymbol{H}$.

▷ **Exercício 6.5.** *Sejam $\boldsymbol{G}$ e $\boldsymbol{H}$ grupos e $A\colon \boldsymbol{G}\curvearrowright\boldsymbol{H}$ uma ação do grupo $\boldsymbol{G}$ sobre o grupo $\boldsymbol{H}$.*

1. *O produto semidireto $\boldsymbol{G}\ltimes_A\boldsymbol{H}$ é um grupo.*
2. *$\boldsymbol{G}\simeq\boldsymbol{G}\ltimes\{\boldsymbol{1_H}\}\leq\boldsymbol{G}\ltimes\boldsymbol{H}$, e $\boldsymbol{G}\ltimes\{\boldsymbol{1_H}\}\trianglelefteq\boldsymbol{G}\ltimes\boldsymbol{H}$ se, e somente se, $\boldsymbol{G}\ltimes\boldsymbol{H}=\boldsymbol{G}\times\boldsymbol{H}$ (ou $A=\mathrm{I}_H$);*
3. *$\boldsymbol{H}\simeq\{\boldsymbol{1_G}\}\ltimes\boldsymbol{H}\trianglelefteq\boldsymbol{G}\ltimes\boldsymbol{H}$;*

► **Exemplo 6.2.** 1. *O produto direto de grupos $\boldsymbol{G}$ e $\boldsymbol{H}$ é um produto semidireto cuja ação é a função constante*

$$\begin{aligned} \mathrm{I}_H\colon G &\longrightarrow \overset{\leftrightarrow}{\mathcal{H}}(H) \\ g &\longmapsto \mathrm{I}_H\colon H\longrightarrow H \\ &\qquad\quad h\longmapsto h \end{aligned}$$

2. *Seja* $n \in \mathbb{N}$. $\mathrm{O}(n) \simeq \mathbb{S}^0 \ltimes \mathrm{SO}(n)$ *com ação dada por*

$$\begin{aligned} A\colon \mathbb{S}^0 &\longrightarrow \overleftrightarrow{\mathcal{H}}(\mathrm{SO}(n)) \\ u &\longmapsto C_{\mathrm{I}_{n-1}\oplus u}\colon \mathrm{SO}(n) \longrightarrow \mathrm{SO}(n) \\ &\qquad\qquad N \longmapsto (\mathrm{I}_{n-1} \oplus u) \circ N \circ (\mathrm{I}_{n-1} \oplus u)^{-1}. \end{aligned}$$

A transformação linear $\mathrm{I}_{n-1} \oplus i \in \mathrm{O}(n)$ *é uma reflexão de* $\mathbb{R}^n$ *que pode ser representada pela matrix*

$$\begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & u \end{pmatrix}.$$

6.3.5 Grupo simples e subgrupo normal maximal

:⊢ **Definição 6.21.** Um *grupo simples* é um grupo não-trivial $\boldsymbol{G}$ cujos únicos subgrupos normais são $\boldsymbol{\{1\}}$ e $\boldsymbol{G}$.

:⊢ **Definição 6.22.** Seja $\boldsymbol{G}$ um grupo. Um subgrupo normal *maximal* de $\boldsymbol{G}$ é um subgrupo normal próprio $\boldsymbol{M} \lhd \boldsymbol{G}$ que satisfaz

1. (Maximalidade) Para todo $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$,

$$M \subseteq N \Rightarrow N = M \quad \text{ou} \quad N = G.$$

⊢ **Proposição 6.33.** *Sejam* $\boldsymbol{G}$ *um grupo e* $\boldsymbol{M} \trianglelefteq \boldsymbol{G}$ *um subgrupo normal. Então* $\boldsymbol{M}$ *é maximal se, e somente se,* $\boldsymbol{G} \diagup \boldsymbol{M}$ *é simples.*

□ *Demonstração.* Consideremos a projeção canônica

$$\begin{aligned} \pi\colon G &\longrightarrow G\big/M \\ g &\longmapsto gM. \end{aligned}$$

(⇒) Suponhamos que $\boldsymbol{M}$ é maximal. Então $\boldsymbol{M}$ é um subgrupo próprio, o implica que $\boldsymbol{G} \diagup \boldsymbol{M}$ é não-trivial. Seja $\boldsymbol{N} \trianglelefteq \boldsymbol{G} \diagup \boldsymbol{M}$. Sabemos que $\boldsymbol{\pi^{-1}(N)} \trianglelefteq \boldsymbol{G}$ (6.18). Como $[e] \in N$, então $\pi^{-1}(1) \subseteq \pi^{-1}(N)$. Notando que $\pi^{-1}([1]) = \mathrm{nuc}(\pi) = M$, segue que $M \subseteq \pi^{-1}(N)$. Como $\boldsymbol{M}$ é maximal, segue que $\pi^{-1}(N) = N$ ou $\pi^{-1}(N) = G$. Notemos que $N = \pi(\pi^{-1}(N))$, pois π é sobrejetiva. No primeiro caso, $N = \pi(\pi^{-1}(N)) = \pi(M) = \{[1]\}$. No segundo caso, $N = \pi(\pi^{-1}(N)) = \pi(G) = G \diagup M$. Portanto $\boldsymbol{G} \diagup \boldsymbol{M}$ é simples.

(⇐) Suponhamos que $\boldsymbol{G} \diagup \boldsymbol{M}$ é simples. Seja $\boldsymbol{N} \trianglelefteq \boldsymbol{G}$ tal que $M \subseteq N$. Como π é homomorfismo de grupos sobrejetivo, segue que $\pi(N) \trianglelefteq \boldsymbol{G} \diagup \boldsymbol{M}$ (6.19). Como $\boldsymbol{G} \diagup \boldsymbol{M}$ é simples, então $\pi(N) = \{[e]\}$ ou $\pi(N) = G \diagup M$. No primeiro caso, $N = \mathrm{nuc}(\pi) = M$. No segundo caso, $N = \pi^{-1}(\pi(N)) = \pi^{-1}(N) = G$. Logo $\boldsymbol{M}$ é maximal.

■

?⊢ **Conjectura 6.1.** *Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{N} \lhd \boldsymbol{G}$ um subgrupo normal próprio. Então $\boldsymbol{G}$ tem subgrupo normal maximal.*

□ *Demonstração.* Usaremos o lema de Zorn. Seja $P \subseteq \wp(G)$ o conjunto de todos os subconjuntos $S \subset G$ tais que $\boldsymbol{S} \lhd \boldsymbol{G}$. Então $(P, \subseteq)$ é um conjunto parcialmente ordenado com a contenção de conjuntos usual. Agora, seja $(C)_{i \in I}$ uma cadeia de $(P, \subseteq)$. Consideremos o conjunto $C := \bigcup_{i \in I} C_i$. Como

Notemos que P não é vazio, pois $N \in P$. Seja

Então P tem elemento maximal. ■

6.3.6 Sequência subnormal

:⊢ **Definição 6.23.** Seja $\boldsymbol{G}$ um grupo. Uma *sequência subnormal* de $\boldsymbol{G}$ é uma sequência finita $(\boldsymbol{N_i})_{i \in [n]}$ de subgrupos de $\boldsymbol{G}$ que satisfaz

$$\boldsymbol{\{1\}} = \boldsymbol{N_0} \trianglelefteq \cdots \trianglelefteq \boldsymbol{N_{n-1}} = \boldsymbol{G}.$$

O grupo $\boldsymbol{N_{i+1}} \diagup \boldsymbol{N_i}$ é o i-ésimo *grupo fator* da sequência. Uma *sequência normal* é uma sequência subnormal em que, para todo $i \in [n]$, $\boldsymbol{N_i} \trianglelefteq \boldsymbol{G}$.

Uma sequência subnormal *estrita* de $\boldsymbol{G}$ é uma sequência subnormal $(\boldsymbol{N_i})_{i \in [n]}$ de $\boldsymbol{G}$ que satisfaz

$$\boldsymbol{\{1\}} = \boldsymbol{N_0} \lhd \cdots \lhd \boldsymbol{N_{n-1}} = \boldsymbol{G}.$$

O *comprimento* de uma sequência subnormal estrita é o número n.

6.3.7 Conjunto gerador

:⊢ **Definição 6.24.** Seja $\boldsymbol{G}$ um grupo e $S \subseteq G$ um conjunto. O grupo *gerado* por S é o grupo $\langle \boldsymbol{S} \rangle \leq \boldsymbol{G}$ em que

$$\langle S \rangle := \left\{ s_1 \cdots s_n \;\middle|\; n \in \mathbb{N},\ s_i \in S \ \text{ ou } \ s_i^{-1} \in S \right\}.$$

$$\langle S \rangle := \left\{ \mathop{\ast}_{i \in [n]} s_i \;\middle|\; n \in \mathbb{N},\ s_i \in S \ \text{ ou } \ s_i^{-1} \in S \right\}.$$

Um *conjunto gerador* de $\boldsymbol{G}$ é um conjunto $S \subseteq G$ tal que $\langle S \rangle = G$.

6.4 Ação de grupos

6.4.1 Grupo simétrico

Para falar de ação de grupo, primeiro demonstraremos que o conjunto de bijeções de um conjunto para si mesmo é um grupo com as operações de composição, inversa de função e função identidade.

:⊢ **Definição 6.25.** Seja C um conjunto. O *grupo simétrico* (ou *grupo de bijeções*) de C é a lista

$$\overleftrightarrow{\mathfrak{F}}(C) = (\overleftrightarrow{\mathcal{F}}(C), \circ, ^{-1}, \mathrm{I}),$$

em que $\overleftrightarrow{\mathcal{F}}(C)$ é o conjunto de bijeções de C para C, $\circ$ é a composição de funções, $^{-1}$ é a inversa de funções e I é a função identidade.

⊢ **Proposição 6.34.** *Seja C um conjunto. O grupo simétrico $\overleftrightarrow{\mathfrak{F}}(C)$ é um grupo.*

□ *Demonstração.* Se $C = \emptyset$, então $\overleftrightarrow{\mathcal{F}}(C) = \{\emptyset\}$. Assim, como $\emptyset \circ \emptyset = \emptyset$, segue que $\circ$ é operação binária em $\{\emptyset\}$. Ainda, segue que $\circ$ é associativa, pois $(\emptyset \circ \emptyset) \circ \emptyset = \emptyset \circ (\emptyset \circ \emptyset)$; tem elemento neutro $\emptyset$, pois $\emptyset \circ \emptyset = \emptyset$, e que todo elemento tem inverso, pois $\emptyset^{-1} = \emptyset$.

Suponhamos, então, que $C \neq \emptyset$ e sejam $p_1, p_2 \in \overleftrightarrow{\mathcal{F}}(C)$. Como p_1 e p_2 são bijeções, a função $p_2 \circ p_1 : C \to C$ é uma bijeção entre C e C (3.11 e 3.12) e, portanto, $p_2 \circ p_1 \in \overleftrightarrow{\mathcal{F}}(C)$. Isso mostra que $\circ$ é uma operação binária em $\overleftrightarrow{\mathcal{F}}(C)$.

1. A composição de funções é associativa, pois, para todos $p_1, p_2, p_3 \in \overleftrightarrow{\mathcal{F}}(C)$, $p_3 \circ (p_2 \circ p_1) = (p_3 \circ p_2) \circ p_1$ (3.5).
2. Ainda, notemos que I_C é o elemento neutro de $\overleftrightarrow{\mathfrak{F}}(C)$, pois, para todo $p \in \overleftrightarrow{\mathcal{F}}(C)$, vale $p \circ \mathrm{I}_C = \mathrm{I}_C \circ p = p$ (3.7).
3. Por fim, como p é uma bijeção, existe função inversa $p^{-1} : C \to C$ que é bijeção entre C e C (3.8 e 3.9); logo existe $p^{-1} \in \overleftrightarrow{\mathcal{F}}(C)$ tal que $p \circ p^{-1} = p^{-1} \circ p = \mathrm{I}_C$.

Portanto concluímos que $\overleftrightarrow{\mathfrak{F}}(C)$ é um grupo. ■

⊢ **Proposição 6.35.** *Sejam A e B conjuntos tais que $\#(A) = \#(B)$. Então*

$$\overleftrightarrow{\mathfrak{F}}(A) \simeq \overleftrightarrow{\mathfrak{F}}(B).$$

□ *Demonstração.* Seja $\phi : A \to B$ uma bijeção e considere a função

$$\begin{aligned} h \colon \overleftrightarrow{\mathcal{F}}(A) &\longrightarrow \overleftrightarrow{\mathcal{F}}(B) \\ p &\longmapsto \phi \circ p \circ \phi^{-1}. \end{aligned}$$

Primeiro notemos que h é homomorfismo de grupos. Sejam $p_1, p_2 \in \overleftrightarrow{\mathcal{F}}(A)$. Então

$$\begin{aligned} h(p_2 \circ p_1)(c) &= \phi \circ (p_2 \circ p_1) \circ \phi^{-1} \\ &= \phi \circ (p_2 \circ \phi^{-1} \circ \phi \circ p_1) \circ \phi^{-1} \\ &= (\phi \circ p_2 \circ \phi^{-1}) \circ (\phi \circ p_1 \circ \phi^{-1}) \\ &= h(p_2) \circ h(p_1). \end{aligned}$$

Portanto h é um homomorfismo de grupos entre $\overleftrightarrow{\boldsymbol{\mathcal{F}}}(A)$ e $\overleftrightarrow{\boldsymbol{\mathcal{F}}}(B)$.

Agora notemos que h é uma bijeção. A inversa de h é a função

$$\begin{aligned} h^{-1} \colon \overleftrightarrow{\mathcal{F}}(B) &\longrightarrow \overleftrightarrow{\mathcal{F}}(A) \\ p &\longmapsto \phi^{-1} \circ p \circ \phi, \end{aligned}$$

pois, para todo $p \in \overleftrightarrow{\mathcal{F}}(B)$,

$$\begin{aligned} (h \circ h^{-1})(p) &= h(h^{-1}(p)) \\ &= \phi \circ h^{-1}(p) \circ \phi^{-1} \\ &= \phi \circ (\phi^{-1} \circ p \circ \phi) \circ \phi^{-1} \\ &= p \\ &= 1_{\overleftrightarrow{\mathcal{F}}(B)}(p), \end{aligned}$$

o que mostra que $h \circ h^{-1} = 1_{\overleftrightarrow{\mathcal{F}}(B)}$, e, para todo $p \in \overleftrightarrow{\mathcal{F}}(A)$,

$$\begin{aligned} (h^{-1} \circ h)(p) &= h^{-1}(h(p)) \\ &= \phi^{-1} \circ h(p) \circ \phi \\ &= \phi^{-1} \circ (\phi \circ p \circ \phi^{-1}) \circ \phi \\ &= p \\ &= 1_{\overleftrightarrow{\mathcal{F}}(A)}(p), \end{aligned}$$

o que mostra que $h^{-1} \circ h = 1_{\overleftrightarrow{\mathcal{F}}(A)}$. Assim, está provado que h é isomorfismo entre $\overleftrightarrow{\boldsymbol{\mathcal{F}}}(A)$ e $\overleftrightarrow{\boldsymbol{\mathcal{F}}}(B)$. ■

Essa proposição mostra que podemos estudar somente os grupos simétricos dos números cardinais, pois isso será equivalente a estudar qualquer grupo simétrico. Em particular, para todo conjunto finito, podemos estudar seu grupo simétrico considerando somente o grupo simétrico $\mathfrak{S}_n$, em que n é o número de elementos do conjunto. A partir de agora, as proposições serão considerando esses grupos $\mathfrak{S}_n$.

⊢ **Teorema 6.36.** *Seja* $\boldsymbol{G}$ *um grupo. Então*

$$\boldsymbol{G} \lesssim \overleftrightarrow{\boldsymbol{\mathcal{F}}}(G).$$

□ *Demonstração.* Consideremos a função

$$\begin{aligned} h\colon G &\longrightarrow \overset{\leftrightarrow}{\mathcal{F}}(G) \\ g &\longmapsto h(g)\colon G \longrightarrow G \\ & \qquad\qquad x \longmapsto g \times x \end{aligned}$$

Primeiro, devemos mostrar que $h(g) \in \overset{\leftrightarrow}{\mathcal{F}}(G)$, para que h esteja bem definida. Para isso, notemos que $h(g)$ está bem definida, já que, para todo $x \in G$, $g \times x \in G$. Ainda, $h(g)$ é uma bijeção, pois $h(g)^{-1} = h(g^{-1})$, já que, para todo $x \in G$,

$$(h(g) \circ h(g)^{-1})(x) = h(g)((h(g)^{-1})(x)) = h(g)(g^{-1} \times x) = g \times g^{-1} \times x = x = 1_G,$$

o que mostra que $h(g) \circ h(g)^{-1} = 1_G$, e

$$(h(g)^{-1} \circ h(g))(x) = h(g)^{-1}((h(g)(x)) = h(g)^{-1}(g \times x) = g^{-1} \times g \times x = x = 1_G,$$

o que mostra que $h(g)^{-1} \circ h(g) = 1_G$. Isso mostra que $h(g)$ é uma bijeção e, portanto, $h(g) \in \overset{\leftrightarrow}{\mathcal{F}}(G)$.

Agora, notemos que h é um homomorfismo de grupos, pois, para todos $g_1, g_2 \in G$, segue que, para todo $x \in G$,

$$\begin{aligned} h(g_1 \times g_2)(x) &= (g_1 \times g_2) \times x \\ &= g_1 \times (g_2 \times x) \\ &= h(g_1)(g_2 \times x) \\ &= h(g_1)(h(g_2)(x)) \\ &= (h(g_1) \circ h(g_2))(x), \end{aligned}$$

o que mostra que $h(g_1 \times g_2) = h(g_1) \circ h(g_2)$. Por fim, notemos que h é injetiva, já que, se $g \in G$ é tal que $h(g) = id_G$, então, para todo $x \in G$,

$$g \times x = h(g)(x) = 1_G(x) = x,$$

o que mostra que $g = e_G$ e, portanto, que $\mathrm{nuc}(h) = \{e_G\}$. ■

Esse teorema é um teorema muito importante, pois ele mostra que, de certa forma, todo grupo é um subconjunto de permutações. Por causa disso que grupos são pensados como os objetos algébricos que modelam a simetria.

6.4.2 Ações

A noção de uma ação de grupo, de certa forma, generaliza uma estratégia usada na demonstração do teorema de que todo grupo é isomorfo a um subgrupo de seu grupo simétrico.

:⊢ **Definição 6.26.** Sejam $\boldsymbol{G}$ um grupo e X um conjunto. Uma *ação* de $\boldsymbol{G}$ em X é um homomorfismo de grupos

$$\begin{aligned} A\colon G &\longrightarrow \overset{\leftrightarrow}{\mathcal{F}}(X) \\ g &\longmapsto g\cdot\colon X \longrightarrow X \\ & \qquad\qquad x \longmapsto g\cdot x. \end{aligned}$$

Denota-se $A\colon \boldsymbol{G} \curvearrowright X$. Diz-se que o grupo $\boldsymbol{G}$ *age* no conjunto X, e denota-se $\boldsymbol{G} \curvearrowright X$, se, e somente se, existe ação de $\boldsymbol{G}$ em X.

A definição acima é uma definição muito simplificada pois ela depende de conceitos um pouco mais complexos, como de homomorfismo de grupos e de grupo de bijeções. A proposição abaixo mostra como essa definição é equivalente a uma definição mais explícita de ação de grupo que também é comumente usada.

⊢ **Proposição 6.37.** *Sejam X um conjunto e $\boldsymbol{G}$ um grupo. Então $\boldsymbol{G} \curvearrowright X$ se, e somente se, existe uma função*

$$\begin{aligned} \cdot\colon G \times X &\longrightarrow X \\ (g, x) &\longmapsto g \cdot x \end{aligned}$$

que satisfaz

1. *(Identidade) Para todo $x \in X$,*

$$e \cdot x = x;$$

2. *(Compatibilidade) Para todos $g_0, g_1 \in G$ e $x \in X$,*

$$(g_1 g_0) \cdot x = g_1 \cdot (g_0 \cdot x).$$

□ *Demonstração.* Se existe ação $A\colon \boldsymbol{G} \curvearrowright X$, basta definir $g \cdot x := A(g)(x)$ e segue que $e\cdot = 1$ e $(g_1 g_0)\cdot = (g_1\cdot) \circ (g_0\cdot)$. Reciprocamente, definimos a ação A a partir de $\cdot$ da mesma forma e, se $g_0, g_1 \in G$ e $x \in X$, segue que

$$A(g_1 g_0)(x) = (g_1 g_0) \cdot x = g_1 \cdot (g_0 \cdot x) = A(g_1)(A(g_0)(x)) = A(g_1) \circ A(g_0)(x),$$

logo $A(g_1 g_0) = A(g_1) \circ A(g_0)$. ■

Essa proposição é geralmente considerada a definição de uma ação *à esquerda* de $\boldsymbol{G}$ em X por causa da posição em que a composição ocorre. Analogamente, uma ação à direita pode ser definida, mas toda ação à direita pode ser traduzida em uma ação à esquerda, de modo que é suficiente estudar somente ações à esquerda. É comum, ainda, estudar ações em conjuntos X com alguma estrutura adicional, geralmente uma topologia. Nesse caso, exige-se que a ação seja uma função contínua, mas também é necessário que G tenha uma estrutura topológica, portanto isso não será definido com cuidado agora.

6.4.3 Órbitas e estabilizadores

:⊢ **Definição 6.27.** Sejam X um conjunto, $\boldsymbol{G}$ um grupo, $\boldsymbol{G} \curvearrowright X$ e $x \in X$. A *órbita* de x sob $\boldsymbol{G}$ é o conjunto

$$G \cdot x := \{g \cdot x \mid g \in G\}.$$

:⊢ **Definição 6.28.** Sejam X um conjunto, $\boldsymbol{G}$ um grupo, $A : \boldsymbol{G} \curvearrowright X$ e $x \in X$. O *estabilizador* de x é

$$G_x = \{g \in G \mid g \cdot x = x\} = A^{-1}(\{1\}) = \mathrm{nuc}(A).$$

O estabilizador de x é um grupo e por isso é chamado de *subgrupo estabilizador* de $\boldsymbol{G}$ com respeito a x, e também é conhecido como *grupo de isotropia* de x.

6.4.4 Permutações finitas

▷ **Exercício 6.6.** *Seja $n \in \mathbb{N}$. Então $|\mathfrak{S}_n| = n!$.*

:⊢ **Definição 6.29.** Seja $n \in \mathbb{N}$. Uma *permutação* de n objetos é um elemento $p \in \mathfrak{S}_n$, denotado por

$$p = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ p(1) & p(2) & \cdots & p(n-1) & p(n) \end{pmatrix}.$$

Notação. Seja $n \in \mathbb{N}$. A composição de duas permutações $p_1, p_2 \in \mathfrak{S}_n$, quando representadas na notação acima, é denotada

$$p_2 \circ p_1 = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ p_2(1) & p_2(2) & \cdots & p_2(n-1) & p_2(n) \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ p_1(1) & p_1(2) & \cdots & p_1(n-1) & p_1(n) \end{pmatrix}.$$

:⊢ **Definição 6.30.** Sejam $n \in \mathbb{N}$ e $p \in \mathfrak{S}_n$. A *matriz de permutação* de p é a matriz $[p] \in \mathbb{M}_n(\mathbb{Z})$ cujas entradas são dadas por

$$[p]_{i,j} = \delta_{i,p(j)} = \begin{cases} 1 & i = p(j) \\ 0 & i \neq p(j). \end{cases}$$

O *conjunto das matrizes de permutação* de $\mathfrak{S}_n$ é o conjunto

$$[\mathfrak{S}_n] := \{[p] \mid p \in \mathfrak{S}_n\}.$$

⊢ **Proposição 6.38.** *Seja $n \in \mathbb{N}$. Então o par $[\boldsymbol{\mathfrak{S}_n}] = ([\mathfrak{S}_n], \cdot)$, em que $\cdot$ é o produto de matrizes, é um grupo, e*

$$\boldsymbol{\mathfrak{S}_n} \simeq [\boldsymbol{\mathfrak{S}_n}].$$

□ *Demonstração.* Primeiro, notemos que, para todos $p, q \in \mathfrak{S}_n$,

$$[p][q]_{i,j} = \bigplus_{k=0}^{n-1} [p]_{i,k}[q]_{k,j} = \bigplus_{k=0}^{n-1} \delta_{i,p(k)}\delta_{k,p(j)}.$$

Mas o produto $\delta_{i,p(k)}\delta_{k,p(j)}$ é igual a 1 se, e somente se, $i = p(k)$ e $k = q(j)$. Como p é bijeção, a segunda condição é equivalente a $p(k) = pq(j)$, e isso mostra que as duas condições são equivalentes a $i = p(k) = pq(j)$. Como p é bijeção, para cada $i \in [n]$, $k = p^{-1}(i)$ é o único $k \in [n]$ tal que a condição é satisfeita, e segue que

$$[p][q]_{i,j} = \bigplus_{k=1}^{n} \delta_{i,p(k)}\delta_{k,p(j)} = \bigplus_{k=1}^{n} \delta_{i,pq(j)} = [pq]_{i,j}.$$

e, como $pq \in \mathfrak{S}_n$, então $[p][q] = [pq] \in [\mathfrak{S}_n]$. Isso mostra que o produto de matrizes é uma operação binária em $[\mathfrak{S}_n]$. Agora, disso segue que $[p][p^{-1}] = [pp^{-1}] = [id]$

Disso, segue que $[\mathfrak{S}_n]$ é um grupo, pois é subgrupo de $\mathbb{M}_n(\mathbb{Z})$. Por fim, consideremos a função

$$\begin{aligned} h\colon \mathfrak{S}_n &\longrightarrow [\mathfrak{S}_n] \\ p &\longmapsto [p]. \end{aligned}$$

Note que que h é homomorfismo, pois, para todos $p, q \in \mathfrak{S}_n$,

$$h(pq) = [pq] = [p][q] = h(p)h(q).$$

Ainda, h ■

:⊢ **Definição 6.31.** Sejam $n \in \mathbb{N}$, $p \in \mathfrak{S}_n$ e $m \in [n]$. A *órbita de m sob p* é o conjunto

$$\mathcal{O}_p(m) := \left\{ p^k(m) \;\middle|\; k \in \mathbb{Z} \right\}.$$

O *período* da órbita $\mathcal{O}_p(m)$ é o número $\#\left(\mathcal{O}_p(m)\right)$. Uma *órbita trivial* é uma órbita de período 1. Uma órbita de p é a órbita de um elemento $m \in [n]$ sob p.

O *conjunto de órbitas* de p é o conjunto

$$\mathcal{O}_p := \left\{ \mathcal{O}_p(m) \mid m \in [n] \right\}.$$

⊢ **Proposição 6.39.** *Sejam $n \in \mathbb{N}$ e $p \in \mathfrak{S}_n$. O conjunto $\mathcal{O}_p$ é uma partição de $[n]$.*

□ *Demonstração.* Primeiro, notemos que $m \in \mathcal{O}_p(m)$ e, portanto, $\emptyset \nsubseteq \mathcal{O}_p$. Ainda, $\bigcup_{m\in[n]} \mathcal{O}_p(m) = [n]$, já que, para todo $m \in [n]$, $m \in \mathcal{O}_p(m)$, o que mostra que $[n] \subseteq \bigcup_{m\in[n]}$ e, para todo $l \in \bigcup_{m\in[n]} \mathcal{O}_p(m)$, existe $m \in [n]$ tal que $l \in \mathcal{O}_p(m)$ e, portanto, existe $k \in \mathbb{N}$ tal que $l = p^k(m) \in [n]$, o que mostra que $\bigcup_{m\in[n]} \subseteq [n]$. Por fim, sejam $o_1, o_2 \in \mathcal{O}_p$. Então existem $m_1, m_2 \in [n]$ tais que $o_1 = \mathcal{O}_p(m_1)$ e $o_2 = \mathcal{O}_p(m_2)$. Se existe $l \in \mathcal{O}_p(m_1) \cap \mathcal{O}_p(m_2)$, então existem $k_1, k_2 \in \mathbb{Z}$ tais que $l = p^{k_1}(m_1) = p^{k_2}(m_2)$. Assim, segue que $m_1 = p^{k_2-k_1}(m_2)$ e, portanto, $m_1 \in \mathcal{O}_p(m_2)$. Mas isso implica que $\mathcal{O}_p(m_1) \subseteq \mathcal{O}_p(m_2)$; a inclusão contrária é análoga e concluímos que $\mathcal{O}_p(m_1) = \mathcal{O}_p(m_2)$. Logo $\mathcal{O}_p$ é uma partição de $[n]$. ■

6.4.4.1 Ciclos e transposições

:⊢ **Definição 6.32.** Sejam $n, k \in \mathbb{N}$. Um *ciclo* de $\mathfrak{S}_n$ é um elemento $c \in \mathfrak{S}_n$ para o qual existe $m \in [n]$ tal que, para todo $m' \in [n]$, $c(m') = m'$ ou existe $d \in [n]$ tal que $m' = c^d(m)$. O *comprimento* de um ciclo é a ordem desse ciclo. Um ciclo c cujo comprimento é k é denotado

$$c = \Big(m \; c(m) \; c^2(m) \; \cdots \; c^{k-2}(m) \; c^{k-1}(m)\Big).$$

⊢ **Proposição 6.40.** *Sejam $n \in \mathbb{N}$.*

1. *Se $c_1, c_2 \in \mathfrak{S}_n$ são ciclos disjuntos, então $c_2 \circ c_1 = c_1 \circ c_2$.*

⊢ **Proposição 6.41** (Fatoração de Permutação)**.** *Seja $n \in \mathbb{N}$ e $p \in \mathfrak{S}_n$. Então existem únicos ciclos $c_1, \ldots, c_k \in \mathfrak{S}_n$ disjuntos dois a dois tais que $p = c_1 \circ \cdots \circ c_k$.*

□ *Demonstração.* Seja $k := \#(\mathcal{O}_p)$. O conjunto $\mathcal{O}_p$ particiona $[n]$. Sejam $(o_k)_{i\in[k]}$ uma indexação de $\mathcal{O}_p$ e $m_1, \cdot, m_k \in [n]$ tais que $o_i = \mathcal{O}_p(m_i)$ para todo $i \in [k]$, e seja $k_i := \#(\mathcal{O}_p(m_i))$. Definamos $c_i := \Big(m_i \; \cdots \; p^{k_i-1}(m_i)\Big)$. Então segue que

$$p = \bigtimes_{i=1}^{k} \Big(m_i \; \cdots \; p^{k_i-1}(m_i)\Big) = \bigtimes_{i=1}^{k} c_i.$$

■

6.5 Grupo linear geral

:⊢ **Definição 6.33.** Sejam $\boldsymbol{C}$ um corpo e $n \in \mathbb{N}$. O *grupo linear geral* de $\boldsymbol{C}$ de ordem n é o conjunto

$$\mathrm{GL}_n(\boldsymbol{C}) := \{M \in \mathbb{M}_{n\times n}(\boldsymbol{C}) \mid \det M \neq 0\}.$$

Se $\boldsymbol{V}$ é um espaço vetorial finito de dimensão d sobre um corpo $\boldsymbol{C}$, o *grupo linear geral* de $\boldsymbol{V}$ é

$$\mathrm{GL}(\boldsymbol{V}) := \mathrm{GL}_n(\boldsymbol{C}).$$

Como $\det(AB) = \det(A)\det(B)$ e as matrizes invertíveis são as que têm determinante não nulo, segue que o conjunto acima forma um grupo com respeito ao produto de matrizes.

:⊢ **Definição 6.34.** Sejam $\boldsymbol{C}$ um corpo e $n \in \mathbb{N}$. O *grupo linear especial* de $\boldsymbol{C}$ de ordem n é o conjunto

$$\mathrm{SL}_n(\boldsymbol{C}) := \{M \in \mathrm{GL}(n, \boldsymbol{C}) \mid \det M = 1\}.$$

6.6 Representação de grupos

:⊢ **Definição 6.35.** Sejam $\boldsymbol{G}$ um grupo e $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Uma *representação* de $\boldsymbol{G}$ em $\boldsymbol{V}$ é um homomorfismo de grupos

$$\rho : \boldsymbol{G} \to \mathrm{GL}(\boldsymbol{V}).$$

O espaço $\boldsymbol{V}$ é o *espaço de representação* e a dimensão de $\boldsymbol{V}$ é a *dimensão da representação.*

:⊢ **Definição 6.36.** Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$.

6.7 Sequências Exatas

:⊢ **Definição 6.37.** Uma *sequência crescente finita* de homomorfismos de grupo é uma sequência de homomorfismos de grupo $(h_i)_{i\in[n]}$ tal que $n \in \mathbb{N}$ e, para todo $i \in [n]$, $\boldsymbol{G}_i$ e $\boldsymbol{G}_{i+1}$ são grupos e $h_i \colon \boldsymbol{G}_i \longrightarrow \boldsymbol{G}_{i+1}$. Denota-se

$$G_0 \xrightarrow{h_0} \cdots \xrightarrow{h_{i-1}} G_i \xrightarrow{h_i} G_{i+1} \xrightarrow{h_{i+1}} \cdots \xrightarrow{h_{n-1}} G_n.$$

Uma *sequência decrescente finita* de homomorfismos de grupo é uma sequência de homomorfismos de grupo $(h_i)_{i\in[n]}$ tal que $n \in \mathbb{N}$ e, para todo $i \in [n]$, $\boldsymbol{G}_i$ e $\boldsymbol{G}_{i+1}$ são grupos e $h_i \colon \boldsymbol{G}_{i+1} \longrightarrow \boldsymbol{G}_i$. Denota-se

$$G_n \xrightarrow{h_{n-1}} \cdots \xrightarrow{h_{i+1}} G_{i+1} \xrightarrow{h_i} G_i \xrightarrow{h_{i-1}} \cdots \xrightarrow{h_0} G_0.$$

Uma *sequência crescente infinita* de homomorfismos de grupo é uma sequência de homomorfismos de grupo $(h_i)_{i\in\mathbb{N}}$ tal que, para todo $i \in \mathbb{N}$, $\boldsymbol{G}_i$ é grupo e $h_i \colon \boldsymbol{G}_i \longrightarrow \boldsymbol{G}_{i+1}$. Denota-se

$$G_0 \xrightarrow{h_0} \cdots \xrightarrow{h_{i-1}} G_i \xrightarrow{h_i} G_{i+1} \xrightarrow{h_{i+1}} \cdots .$$

Uma *sequência decrescente infinita* de homomorfismos de grupo é uma sequência de homomorfismos de grupo $(h_i)_{i\in\mathbb{N}}$ tal que, para todo $i \in \mathbb{N}$, $\boldsymbol{G}_i$ é grupo e $h_i \colon \boldsymbol{G}_{i+1} \longrightarrow \boldsymbol{G}_i$. Denota-se

$$\cdots \xrightarrow{h_{i+1}} G_{i+1} \xrightarrow{h_i} G_i \xrightarrow{h_{i-1}} \cdots \xrightarrow{h_0} G_0.$$

Vamos usar em geral a notação crescente finita, mas sempre ficará claro como adaptar a demonstração para os outros casos.

:⊢ **Definição 6.38.** Uma *sequência exata* de homomorfismos de grupo é uma sequência de homomorfismo de grupos $(h_i)_{i\in[n]}$ tal que, para todo $i \in [n]$,

$$\mathrm{im}(h_i) = \mathrm{nuc}(h_{i+1}).$$

▷ **Exercício 6.7.** *1. A sequência* $\{1\} \xrightarrow{1} G \xrightarrow{h} G'$ *é exata se, e somente se, h é monomorfismo;*

2. A sequência $G' \xrightarrow{h} G \xrightarrow{1} \{1\}$ *é exata se, e somente se, h é epimorfismo;*

3. A sequência $\{1\} \xrightarrow{1} G \xrightarrow{h} G' \xrightarrow{1} \{1\}$ *é exata se, e somente se, h é isomorfismo;*

4. A sequência $\{1\} \xrightarrow{1} G \xrightarrow{h} G' \xrightarrow{h'} G'' \xrightarrow{1} \{1\}$ *é exata se, e somente se,*

$$G'' \simeq {}^{G'}\big/_{h(G)}.$$

:⊢ **Definição 6.39.** Uma *sequência exata curta* é uma sequência exata da forma

$$\{1\} \xrightarrow{1} G \xrightarrow{h} G' \xrightarrow{h'} G'' \xrightarrow{1} \{1\}.$$

⊢ **Proposição 6.42.** *A sequência* $G \xrightarrow{h} G' \xrightarrow{h'} G''$ *é exata se, e somente se,* $\{1\} \xrightarrow{1} \mathrm{im}(h) \xrightarrow{\iota} G' \xrightarrow{\mathrm{p}} G'/\mathrm{nuc}(h') \xrightarrow{1} \{1\}$ *é uma sequência exata curta.*

□ *Demonstração.* (⇒) Supondo que $\mathrm{im}(h) = \mathrm{nuc}(h')$, temos que

$$\iota(\mathrm{im}(h)) = \mathrm{im}(h) = \mathrm{nuc}(h'),$$

portanto

$$G'\big/_{\iota(\mathrm{im}(h))} \simeq G'\big/_{\mathrm{nuc}(h')},$$

o que implica que a sequência curta é exata.

(⇐) Supondo que a sequência curta é exata, temos que

$$G'\big/_{\mathrm{im}(h)} = G'\big/_{\iota(\mathrm{im}(h))} \simeq G'\big/_{\mathrm{nuc}(h')},$$

o que implica que $\mathrm{im}(h) = \mathrm{nuc}(h')$, logo que a sequência $G \xrightarrow{h} G' \xrightarrow{h'} G''$ é exata. ■

Capítulo 7

Anéis

7.1 Conceitos básicos

7.1.1 Anel e subanel

:⊢ **Definição 7.1.** Um *anel* é uma lista $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ tal que

1. $\boldsymbol{A}^+ := (A, +, -, 0)$ é um grupo comutativo, em que $+$ é a *adição*, $-$ é a *inversão aditiva* e 0 é a *nulidade* (ou o *zero*) de $\boldsymbol{A}$;
2. $\boldsymbol{A}^\times := (A, \times, 1)$ é um monoide comutativo, em que $\times$ é a *multiplicação* e 1 é a *unidade* (ou o *um*) de $\boldsymbol{A}$;
3. A multiplicação $\times$ é distributiva sobre a adição $+$.

Notação. O símbolo de multiplicação $\times$ será suprimido sempre que possível, de modo que denotaremos a_0a_1 para $a_0 \times a_1$, e a notação da multiplicação terá preferência sobre a da adição e a da substração, pois $\times$ é distributiva sobre $+$, de modo que denotaremos $a_0a_1 + a_2$ para $(a_1a_2) + a_3$ e $-a_1a_2$ para $-(a_1a_2)$. Os símbolos operatórios relativos à adição e à multiplicação serão, respectivamente,

$$+ \quad \text{e} \quad \times .$$

O elemento a multiplicado por si mesmo n vezes será denotado a^n. Denotaremos o inverso de um elemento $a \in A$ com respeito a $\times$ por a^{-1} ou $\diagup a$, se ele existir, pois sabemos que é único (5.34).

Na definição de anel aqui adotada, consideramos anéis que têm multiplicação comutativa e unidade. Em contextos mais gerais, esses objetos são chamados de *anéis comutativos unitários*.

Um comentário sobre as identidades 0 e 1. Se $0 = 1$, o anel será *trivial*, no sentido de que 0 será seu único elemento, mas mantemos esse caso na definição pois, mais à frente, no estudo de anéis quocientes, será proveitoso que qualquer anel quociente seja um anel, e o quociente de um anel por ele mesmo é o anel trivial.

$\vdash$ **Proposição 7.1.** *Seja* $\boldsymbol{A}$ *um anel. Então, para todos* $a, a' \in A$*,*

1. $0a = 0$*;*
2. *Se* $0 = 1$*, então* $A = \{0\}$*;*
3. $-(aa') = (-a)a'$*.*

□ *Demonstração.* 1.

$$\begin{aligned} 0a &= 0a + 0 \\ &= 0a + (0a - 0a) \\ &= (0a + 0a) - 0a \\ &= (0 + 0)a - 0a \\ &= 0a - 0a \\ &= 0. \end{aligned}$$

2. Se $0 = 1$, então, para todo $a \in A$,

$$a = 1a = 0a = 0;$$

3.

$$\begin{aligned} -(aa') &= -(aa') + 0 \\ &= -(aa') + 0a' \\ &= -(aa') + (a - a)a' \\ &= -(aa') + (aa' + (-a)a') \\ &= (-(aa') + aa') + (-a)a' \\ &= 0 + (-a)a' \\ &= (-a)a'. \end{aligned}$$

■

:$\vdash$ **Definição 7.2.** Seja $\boldsymbol{A}$ um anel. O conjunto dos elementos invertíveis sob multiplicação de $\boldsymbol{A}$ é denotado por $A^{\diagup}$.

$\vdash$ **Proposição 7.2.** *Seja* $\boldsymbol{A}$ *um anel. A lista* $(A^{\diagup}, \times|_{A^{\diagup} \times A^{\diagup}}, {}^{\diagup}, 1)$ *é um grupo comutativo.*

□ *Demonstração.* A tripla $(A, \times, 1)$ é um monoide comutativo com identidade 1. Portanto segue que a quádrupla $(A^{\diagup}, \times|_{A^{\diagup} \times A^{\diagup}}, {}^{\diagup}, 1)$ é um grupo (em que ${}^{\diagup}a = a^{-1}$ denota o inverso de a sob $\times$). Como $\times$ é comutativa, então $(A^{\diagup}, \times|_{A^{\diagup} \times A^{\diagup}}, {}^{-1}, 1)$ também o é. ■

:⊢ **Definição 7.3.** Seja $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ um anel. Um *subanel* de $\boldsymbol{A}$ é um anel $\boldsymbol{S} = (S, +_S, -_S, 0_S, \times_S, 1_S)$ tal que $S \subseteq A$, $+_S = +|_{S\times S}$, $-_S = -|_S$, $0_S = 0$, $\times_S = \times|_{S\times S}$ e $1_S = 1$. Denota-se $\boldsymbol{S} \leq \boldsymbol{A}$. Um subanel *próprio* de $\boldsymbol{A}$ é um subanel $\boldsymbol{S} \leq \boldsymbol{A}$ em que S é um conjunto próprio de A ($S \subset A$). Denota-se $\boldsymbol{S} < \boldsymbol{A}$.

⊢ **Proposição 7.3.** *Sejam $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ um anel e $S \subseteq A$. Então*

$$\boldsymbol{S} = (S, +|_{S\times S}, -|_S, 0, \times|_{S\times S}, 1)$$

é um anel se, e somente se,

1. $\boldsymbol{S}^+ = (S, +|_{S\times S}, -|_S, 0)$ *é um subgrupo comutativo de* $\boldsymbol{A}^+$

   SG1 (Não-vacuidade) $S \neq \emptyset$;
   SG2 (Fechamento) Para todos $s_1, s_2 \in S$, $s_1 + s_2 \in S$;
   SG3 (Invertibilidade) Para todo $s \in S$, $-s \in S$;

2. $\boldsymbol{S}^\times = (S, \times|_{S\times S}, 1)$ *é um submonoide comutativo de* $\boldsymbol{A}^\times$

   SM1 (Identidade) $1 \in S$;
   SM2 (Fechamento) Para todos $s_1, s_2 \in S$, $s_1 \times s_2 \in S$.

□ *Demonstração.* ($\Rightarrow$) Suponhamos que $\boldsymbol{S}$ é um anel.

1. (Subgrupo) Como $S \subseteq A$ e $\boldsymbol{S}^+$ um grupo comutativo, então é um subgrupo de $\boldsymbol{A}^+$ por definição de subgrupo (o que é equivalente às propriedades listadas) e é comutativo (6.1)).
2. (Submonoide) Como $S \subseteq A$ e $\boldsymbol{S}^\times$ é um monoide comutativo com $1 \in S$, então é um submonoide de $(A, \times)$ por definição de submonoide (o que é equivalente às propriedades listadas) e é comutativo (5.30).

($\Leftarrow$) Suponhamos, agora, que $\boldsymbol{S}^+$ é subgrupo comutativo de $\boldsymbol{A}^+$ e $\boldsymbol{S}^\times$ é submonoide comutativo de $\boldsymbol{A}^\times$.

1. (Grupo comutativo) Como $(S, +|_{S\times S})$ é subgrupo comutativo, então é um grupo comutativo por definição de subgrupo.
2. (Monoide comutativo) Como $(S, \times|_{S\times S})$ é submonoide comutativo, então é um monoide comutativo por definição de monoide.
3. (Distributividade) Sejam $s_1, s_2, s_3 \in S$. Então

$$\begin{aligned} s_1 \times|_{S\times S} (s_2 +|_{S\times S} s_3) &= s_1 \times (s_2 + s_3) \\ &= (s_1 \times s_2) + (s_1 \times s_3) \\ &= (s_1 \times|_{S\times S} s_2) +|_{S\times S} (s_1 \times|_{S\times S} s_3). \end{aligned}$$

   Logo $\times_{S\times S}$ é distributiva sobre $+\times_{S\times S}$.

■

7.1.2 Ideais e anéis quocientes

:⊢ **Definição 7.4.** Seja $\boldsymbol{A}$ um anel. Um *ideal* de $\boldsymbol{A}$ é um conjunto não vazio $I \subseteq A$ tal que

1. Para todos $i_1, i_2 \in I$, $i_1 - i_2 \in I$;
2. Para todos $a \in A$ e $i \in I$, $ai \in I$.

Denotamos que I é ideal de $\boldsymbol{A}$ por $I \trianglelefteq A$. Ainda, $I \triangleleft A$ significa que $I \neq A$ e $I \trianglelefteq A$.

É interessante observar que I é subgrupo de $(A, +)$. A definição de ideal difere da definição de subanel na propriedade 2.

⊢ **Proposição 7.4.** *Sejam $\boldsymbol{A}$ um anel e $I \trianglelefteq A$. Então*

1. $0 \in I$;
2. *Para todos $i_1, i_2 \in I$, $i_1 + i_2 \in I$;*
3. $1 \in I \Rightarrow I = A$;
4. $\{0\}$ *e A são ideais de A.*

□ *Demonstração.*
1. Seja $i \in I$. Então $0 = i - i \in I$.
2. Sejam $i_1, i_2 \in I$. Pelo item anterior, sabemos que $0 \in I$, o que implica $-i_2 = 0 - i_2 \in I$. Logo $i_1 + i_2 = i_1 - (-i_2) \in I$.
3. Se $1 \in I \trianglelefteq A$, então, para todo $a \in A$, temos $a = a \cdot 1 \in A$. Logo $I = A$.
4. Consideremos $\{0\}$. Se $i \in \{0\}$, então $i = 0$. Portanto, para todo $a \in A$ e $i \in \{0\}$, temos $i - i = 0 \in \{0\}$ e $ai = 0 \in \{0\}$. Logo $\{0\} \trianglelefteq A$. Agora, consideremos A. Para todo $a_1, a_2 \in A$, temos $a_1 - a_2 \in A$ e $a_1 a_2 \in A$. Logo $A \trianglelefteq A$.

■

⊢ **Proposição 7.5.** *Sejam $\boldsymbol{A}$ um anel e $(I_j)_{j \in J}$ uma família de ideais de $\boldsymbol{A}$. Então*

$$I := \bigcap_{j \in J} I_j$$

é um ideal de $\boldsymbol{A}$.

□ *Demonstração.* Sejam $i_1, i_2 \in I$ e $a \in A$. Então, para todo $j \in J$, $i_1, i_2 \in I_j$ e, como I_j é ideal de $\boldsymbol{A}$, segue que $i_1 - i_2 \in I_j$ e que $ai_1 \in W_i$. Logo $i_1 - i_2 \in I$ e $ai_1 \in I$, o que mostra que I é ideal de $\boldsymbol{A}$. ■

⊢ **Proposição 7.6.** *Sejam $\boldsymbol{A}$ um anel e $(I_j)_{n \in \mathbb{N}}$ uma sequência crescente de ideais de $\boldsymbol{A}$; ou seja, $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$. Então*

$$I := \bigcup_{n \in \mathbb{N}} I_n$$

é um ideal de $\boldsymbol{A}$.

□ *Demonstração.* Sejam $i_1, i_2 \in I$. Então existem $n, m \in \mathbb{N}$ tais que $i_1 \in I_n$ e $i_2 \in I_m$. Nesse caso, $I_n \subseteq I_m$ ou $I_m \subseteq I_n$. Sem perda de generalidade, suponhamos o primeiro caso. Então segue que $i_1 \in I_m$ e, portanto, $i_1 - i_2 \in I_m$, o que mostra que $i_1 - i_2 \in I$. Agora, seja $a \in A$ e notemos que, como I_n é ideal de $\boldsymbol{A}$, segue que $ai_1 \in I_n$. Logo $ai_1 \in I$, o que mostra que I é ideal de $\boldsymbol{A}$. ■

Lembremos, dados $\boldsymbol{A}$ um anel e $a_1, \ldots, a_n \in A$, definimos o conjunto

$$\mathop{+}_{i=1}^{n} a_i A = a_1 A + \ldots + a_n A = \left\{ \mathop{+}_{i=1}^{n} a_i b_i \;\middle|\; b_i \in A \right\}.$$

⊢ **Proposição 7.7.** *Sejam $\boldsymbol{A}$ um anel e $a_1, \ldots, a_n \in A$. Então*

$$I := \mathop{+}_{k=1}^{n} a_k A \trianglelefteq A.$$

□ *Demonstração.* Sejam $a \in A$ e $i_1, i_2 \in I$ tais que $i_1 = +_{k=1}^{n} a_i b_i$ e $i_2 = +_{k=1}^{n} a_i c_i$. Então

$$i_1 - i_2 = \mathop{+}_{k=1}^{n} a_k b_k - \mathop{+}_{k=1}^{n} a_k c_k = \mathop{+}_{k=1}^{n} a_k (b_k - c_k) \in I,$$

pois $(b_k - c_k) \in A$ para todo $k \in \{1, \ldots, n\}$. Ainda,

$$ak_1 = a \mathop{+}_{k=1}^{n} a_k b_k = \mathop{+}_{k=1}^{n} a_k (ab_k) \in I,$$

pois $ab_k \in A$ para todo $k \in \{1, \ldots, n\}$. Logo $I \trianglelefteq A$. ■

Esse ideal é chamado de ideal de A gerado por $a_1, \ldots, a_n$.

:⊢ **Definição 7.5.** Seja $\boldsymbol{A}$ um anel. Um *ideal principal* de $\boldsymbol{A}$ é um ideal $I \trianglelefteq A$ tal que, para algum $a \in A$, $I = aA$.

⊢ **Proposição 7.8.** *Seja $\boldsymbol{A}$ um anel. Então $\boldsymbol{A}$ é um corpo se, e somente se, é um anel não-trivial cujos únicos ideais são $\{0\}$ e A.*

□ *Demonstração.* Suponha que $\boldsymbol{A}$ é um corpo. Então $(A \setminus \{0\}, \cdot)$ é um grupo e, portanto, $A \neq \emptyset$. Seja $I \trianglelefteq A$ e suponha que $I \neq \{0\}$. Então existe $i \in I \setminus \{0\}$. Como $\boldsymbol{A}$ é corpo, existe $i^{-1} \in A$. Portanto $1 = i^{-1} i \in I$. Logo $I = A$.

Por outro lado, suponha que os únicos ideais de A são $\{0\}$ e A. Como $\boldsymbol{A}$ é não-trivial, seja $a \in A \setminus \{0\}$ e consideremos o ideal $I = aA$. Notemos que $a = a \cdot 1 \in aA$, o que implica $I \neq \{0\}$. Portanto $I = A$. Mas então $1 \in aA$, o que significa que deve existir $b \in A$ tal que $1 = ab$; ou seja, todo $a \in A \setminus \{0\}$ tem inverso em A. Logo $\boldsymbol{A}$ é corpo. ■

$\vdash$ **Proposição 7.9.** *Sejam* $\boldsymbol{A}$ *um anel e* $I \trianglelefteq A$. *A relação binária* $\sim_I$ *em* A, *definida por*

$$a \sim_I b \Leftrightarrow a - b \in I,$$

é uma relação de equivalência.

□ *Demonstração.* Vamos demonstrar as três propriedades de uma relação de equivalência.

1. (Reflexividade) Seja $a \in A$. Então $a - a = 0 \in I$. Logo $a \sim_I a$.
2. (Simetria) Sejam $a_1, a_2 \in A$ tais que $a_1 \sim_I a_2$. Então $(a_1 - a_2) \in I$. Mas $0 \in I$, o que implica $a_2 - a_1 = 0 - (a_1 - a_2) \in I$. Logo $a_2 \sim_I a_1$.
3. (Transitividade) Sejam $a_1, a_2, a_3 \in A$ tais que $a_1 \sim_I a_2$ e $a_2 \sim_I a_3$. Então $(a_1 - a_2), (a_2 - a_3) \in I$, o que implica $a_1 - a_3 = (a_1 - a_2) + (a_2 - a_3) \in I$. Logo $a_1 \sim_I a_3$.

■

:$\vdash$ **Definição 7.6.** Sejam $\boldsymbol{A}$ um anel e $I \trianglelefteq A$. O conjunto $a + I = \{a + i \mid i \in I\}$ é a *classe lateral* de I e a é um *representante* da classe. O conjunto das classes laterais de A é denotado por $A/I := \{a + I \mid a \in A\}$.

$\vdash$ **Proposição 7.10.** *Sejam* $(A, +, \cdot)$ *um anel,* $I \trianglelefteq A$ *e* $\sim_I$ *a relação de equivalência definida na proposição anterior e* $a \in A$. *Então a classe de equivalência* $[a] = \{b \in A \mid b \sim_I a\}$ *é igual à classe lateral* $a + I$ *e, por consequência, o conjunto quociente* $A/\sim_I$ *é igual ao conjunto* A/I.

□ *Demonstração.* Seja $b \in [a]$. Então $b - a \in I$; ou seja, existe $i \in I$ tal que $b - a = i$. Mas isso implica $b = a + i$, que implica, por sua vez, que $b \in a + I$. Por outro lado, seja $b \in a + I$. Então existe $i \in I$ tal que $b = a + i$; ou seja, $b - a = i$, que implica $b - a \in I$ e, assim, $b \in [a]$. Disso, vem que $A/\sim_I = A/I$. ■

Uma consequência disso é que o conjunto A é particionado em classes laterais de I. Outra consequência é que duas classes laterais são iguais se, e somente se, a diferença entre seus representantes está em I.

:$\vdash$ **Definição 7.7.** Sejam $\boldsymbol{A}$ um anel, $I \trianglelefteq A$ e $a_1, a_2 \in A$. Então definimos as operações binárias $\oplus$ e $\odot$ em A/I por

$$(a_1 + I) \oplus (a_2 + I) := (a_1 + a_2) + I \qquad (a_1 + I) \odot (a_2 + I) := (a_1 \cdot a_2) + I$$

Denotaremos $\oplus$ e $\odot$ por $+$ e $\cdot$ quando não existir ambiguidade.

$\vdash$ **Proposição 7.11.** *As operações* $\oplus$ *e* $\odot$ *da definição anterior estão bem definidas.*

□ *Demonstração.* Sejam $a_1, a_2, b_1, b_2 \in A$ tais que $a_1 + I = a_2 + I$ e $b_1 + I = b_2 + I$. Primeiro, vamos mostrar que $\oplus$ está bem definida. Devemos mostrar que $(a_1 + b_1) + I = (a_2 + b_2) + I$. De $a_1 + I = a_2 + I$, sabemos que $a_1 - a_2 \in I$. Da mesma forma, $b_1 - b_2 \in I$. Mas então $(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) \in I$, o que implica $(a_1 + b_1) + I = (a_2 + b_2) + I$.

Agora, vamos mostrar que $\odot$ está bem definida. Devemos mostrar que $(a_1 b_1) + I = (a_2 b_2) + I$. Sejam $c = a_1 - a_2$ e $d = b_1 - b_2$. Notemos que

$$a_1 b_1 = (a_2 + c)(b_2 + d) = a_2 b_2 + a_2 d + c b_2 + cd.$$

Como $c, d \in I$, $(a_2 d + c b_2 + cd) \in I$. Logo $a_1 b_1 - a_2 b_2 \in I$, o que implica $(a_1 b_1) + I = (a_2 b_2) + I$. ■

⊢ **Proposição 7.12.** *Sejam* $\boldsymbol{A} = (A, +, \cdot)$ *um anel e* $I \trianglelefteq A$. *Então* $\boldsymbol{A} \diagup \boldsymbol{I} := (A \diagup I, \oplus, \odot)$ *é um anel, chamado* anel quociente *de* A *por* I.

□ *Demonstração.* Sejam $a_1, a_2, a_3 \in A$. Primeiro, vamos mostrar que $(A \diagup I, \oplus)$ é um grupo comutativo. As propriedades de grupo comutativo decorrem do fato de que $(A, +)$ é grupo comutativo com elemento neutro 0. A operação $\oplus$ é associativa, pois

$$\begin{aligned}((a_1 + I) \oplus (a_2 + I)) \oplus (a_3 + I) &= ((a_1 + a_2) + I) \oplus (a_3 + I) \\ &= ((a_1 + a_2) + a_3) + I \\ &= (a_1 + (a_2 + a_3)) + I \\ &= (a_1 + I) \oplus ((a_2 + a_3) + I) \\ &= (a_1 + I) \oplus ((a_2 + I) \oplus (a_3 + I)),\end{aligned}$$

e comutativa, pois

$$(a_1 + I) \oplus (a_2 + I) = (a_1 + a_2) + I = (a_2 + a_1) + I = (a_2 + I) \oplus (a_1 + I).$$

Ainda, $0 + I$ é elemento neutro, pois

$$(a_1 + I) \oplus (0 + I) = (a_1 + 0) + I = a_1 + I.$$

Por fim, existe $-a_1 \in A$. Assim, $(-a_1) + I$ é inverso de $a_1 + I$, pois

$$(a_1 + I) \oplus ((-a_1) + I) = (a_1 + (-a_1)) + I = 0 + I.$$

Agora, devemos mostrar que $(A \diagup I, \odot)$ é um monoide comutativo. Novamente, as propriedades de monoide comutativo decorrem do fato de que $(A, \cdot)$ é um monoide

comutativo com elemento neutro 1. A operação $\odot$ é associativa, pois

$$\begin{aligned}((a_1+I)\odot(a_2+I))\odot(a_3+I) &= ((a_1\cdot a_2)+I)\odot(a_3+I)\\ &= ((a_1\cdot a_2)\cdot a_3)+I\\ &= (a_1\cdot(a_2\cdot a_3))+I\\ &= (a_1+I)\odot((a_2\cdot a_3)+I)\\ &= (a_1+I)\odot((a_2+I)\odot(a_3+I)),\end{aligned}$$

e comutativa, pois

$$(a_1+I)\odot(a_2+I) = (a_1\cdot a_2)+I = (a_2\cdot a_1)+I = (a_2+I)\odot(a_1+I).$$

Ainda, $1+I$ é elemento neutro, pois

$$(a_1+I)\odot(1+I) = (a_1\cdot 1)+I = a_1+I.$$

Por fim, como $\cdot$ é distributiva sobre $+$, temos que

$$\begin{aligned}(a_1+I)\odot((a_2+I)\oplus(a_3+I)) &= (a_1+I)\odot((a_2+a_3)+I)\\ &= (a_1\cdot(a_2+a_3))+I\\ &= ((a_1\cdot a_2)+(a_1\cdot a_3))+I\\ &= ((a_1\cdot a_2)+I)\oplus((a_1\cdot a_3)+I)\\ &= ((a_1+I)\odot(a_2+I))\oplus((a_1+I)\odot(a_3+I)).\end{aligned}$$

■

7.1.3 Homomorfismo de anel

⊢ **Definição 7.8.** Sejam $\boldsymbol{A} = (A,+,\cdot)$ e $\boldsymbol{B} = (B,\oplus,\odot)$ dois anéis. Um *homomorfismo de anéis* entre $\boldsymbol{A}$ e $\boldsymbol{B}$ é uma função $\phi : A \to B$ que é

1. um homomorfismo de grupos entres $(A,+)$ e $(B,\oplus)$
   1.1. $\forall a_1, a_2 \in A \qquad \phi(a_1+a_2) = \phi(a_1)\oplus\phi(a_2)$;
2. um homomorfismo de monoides entre $(A,\cdot)$ e $(B,\odot)$
   2.1. $\forall a_1, a_2 \in A \qquad \phi(a_1\cdot a_2) = \phi(a_1)\odot\phi(a_2)$;
   2.2. $\phi(1_A) = 1_B$.

O conjunto de todos os homomorfismos de anéis entre $\boldsymbol{A}$ e $\boldsymbol{B}$ é denotado por $\mathcal{H}(\boldsymbol{A},\boldsymbol{B})$.

► **Exemplo 7.1.** *Seja* $(A, +, \cdot)$ *um anel e consideremos o anel dos números inteiros* $(\mathbb{Z}, +, \cdot)$. *Então*

$$\begin{aligned} \phi\colon \mathbb{Z} &\longrightarrow A \\ z &\longmapsto \begin{cases} \displaystyle\mathop{+}_{i=1}^{z} 1_A & z > 0 \\ 0_A & z = 0 \\ -\phi(-z) & z > 0 \end{cases} \end{aligned}$$

é um homomorfismo de anéis.

□ *Demonstração.* Sejam $z_1, z_2 \in \mathbb{Z}$. Para ver que ϕ é um homomorfismo de anéis, provemos primeiro que $\phi(z_1 + z_2) = \phi(z_1) + \phi(z_2)$. Vamos separar a demonstração em vários casos. Primeiro, notemos que, se $z_1 = 0$ ou $z_2 = 0$, a igualdade é trivial; sem perda de generalidade, suponha que $z_2 = 0$. Então

$$\phi(z_1 + z_2) = \phi(z_1) = \phi(z_1) + 0_A = \phi(z_1) + \phi(z_2).$$

Então, suponhamos $z_1 z_2 \neq 0$. Se $z_1 > 0$ e $z_1 > 0$, então $z_1 + z_2 > 0$. Logo

$$\phi(z_1 + z_2) = \mathop{+}_{i=1}^{z_1+z_2} 1_A = \mathop{+}_{i=1}^{z_1} 1_A + \mathop{+}_{i=z_1+1}^{z_1+z_2} 1_A = \mathop{+}_{i=1}^{z_1} 1_A + \mathop{+}_{i=1}^{z_2} 1_A = \phi(z_1) + \phi(z_2).$$

Se $z_1 < 0$ e $z_1 < 0$, então $z_1 + z_2 < 0$. Logo $-z_1$, $-z_2$ e $-(z_1 + z_2)$ são positivos e segue da equação anterior que

$$\begin{aligned} \phi(z_1 + z_2) &= -\phi(-(z_1 + z_2)) \\ &= -\phi((-z_1) + (-z_2)) \\ &= -(\phi(-z_1) + \phi(-z_2)) \\ &= -(-\phi(z_1) - \phi(z_2)) \\ &= \phi(z_1) + \phi(z_2). \end{aligned}$$

No caso que resta, z_1 e z_2 são um positivo e um negativo; sem perda de generalidade, suponhamos que $z_1 > 0$ nesse caso $-z_2 > 0$. Se tivermos $z_1 = -z_2$, então

$$\begin{aligned} \phi(z_1 + z_2) &= \phi(0) \\ &= 0_A \\ &= \mathop{+}_{i=1}^{z_1} 1_A - \mathop{+}_{i=1}^{z_1} 1_A \\ &= \phi(z_1) - \phi(-z_1) \\ &= \phi(z_1) + \phi(z_2). \end{aligned}$$

Se tivermos $z_1 > -z_2$, então

$$\begin{aligned}
\phi(z_1+z_2) &= \mathop{+}_{i=1}^{z_1+z_2} 1_A \\
&= \mathop{+}_{i=1}^{z_1+z_2} 1_A + \mathop{+}_{i=1}^{-z_2} 1_A - \mathop{+}_{i=1}^{-z_2} 1_A \\
&= \mathop{+}_{i=1}^{z_1+z_2} 1_A + \mathop{+}_{i=z_1+z_2+1}^{z_1} 1_A - \mathop{+}_{i=1}^{-z_2} 1_A \\
&= \mathop{+}_{i=1}^{z_1} 1_A - \mathop{+}_{i=1}^{-z_2} 1_A \\
&= \phi(z_1) + \phi(z_2).
\end{aligned}$$

Por fim, se $-z_2 > z_1$, então $-z_1 < 0$ e $-z_2 >$ e segue da equação anterior que

$$\begin{aligned}
\phi(z_1+z_2) &= -\phi(-(z_1+z_2)) \\
&= -\phi((-z_1)+(-z_2)) \\
&= -(\phi(-z_1)+\phi(-z_2)) \\
&= -(-\phi(z_1)-\phi(z_2)) \\
&= \phi(z_1)+\phi(z_2).
\end{aligned}$$

■

:⊢ **Definição 7.9.** Sejam $\boldsymbol{A}$ um anel e $n \in \mathbb{Z}$. O *número inteiro* n em $\boldsymbol{A}$ é o elemento

$$n_A := \begin{cases} +_{i=1}^{n} 1_A, & n > 0 \\ 0, & n = 0 \\ +_{i=1}^{-n}(-1_A), & n < 0. \end{cases}$$

O conjunto dos inteiros de $\boldsymbol{A}$ é denotado $\mathbb{Z}(\boldsymbol{A})$. O *coeficiente binomial* de $n_A, k_A \in \mathbb{Z}(\boldsymbol{A})$ é o número

$$\binom{n_A}{k_A} := \binom{n}{k}_A.$$

⊢ **Proposição 7.13.** *Sejam* $\boldsymbol{A}$ *um anel e* $a, b \in A$. *Então*

$$(a+b)^n = \mathop{+}_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

► **Exemplo 7.2.** *Sejam* $\boldsymbol{A}$ *um anel e* $I \trianglelefteq A$. *Então a* projeção canônica *de* A *em* A/I, *definida por*

$$\begin{aligned}
\pi\colon A &\longrightarrow A/I \\
a &\longmapsto a + I,
\end{aligned}$$

é um homomorfismo de anéis.

□ *Demonstração.* Sejam $a_1, a_2 \in A$. Vemos que π é um homomorfismo de grupos entre $(A, +)$ e $(A/I, +)$, pois

$$\pi(a_1 + a_2) = (a_1 + a_2) + I = (a_1 + I) + (a_2 + I) = \pi(a_1) + \pi(a_2).$$

Também, vemos que π é um homomorfismo de monoides entre $(A, \cdot)$ e $(A/I, \cdot)$, pois

$$\pi(a_1 \cdot a_2) = (a_1 \cdot a_2) + I = (a_1 + I) \cdot (a_2 + I) = \pi(a_1) \cdot \pi(a_2)$$

e $\pi(1) = 1 + I = 1_{A/I}$. ■

⊢ **Corolário 7.14** (Homomorfismos preservam a estrutura algébrica entre anéis)**.** *Sejam* $\boldsymbol{A} = (A, +, \cdot)$ *e* $\boldsymbol{B} = (B, \oplus, \odot)$ *dois anéis e* $\phi : A \to B$ *um homomorfismo de anéis. Então*

1. $\phi(0_A) = 0_B$;
2. $-\phi(a) = \phi(-a)$.

□ *Demonstração.* Como ϕ é um homomorfismo de grupos entres $(A, +)$ e $(B, \oplus)$, sabemos que ϕ preserva a estrutura algébrica de grupo entre os grupos (6.15). ■

⊢ **Corolário 7.15** (Composição de homomorfismos é homomorfismo)**.** *Sejam* $\boldsymbol{A_1}$, $\boldsymbol{A_2}$ *e* $\boldsymbol{A_3}$ *anéis e* $\phi \in \mathcal{H}(\boldsymbol{A_1}, \boldsymbol{A_2})$ *e* $\psi \in \mathcal{H}(\boldsymbol{A_2}, \boldsymbol{A_3})$. *Então* $(\psi \circ \phi) \in \mathcal{H}(\boldsymbol{A_1}, \boldsymbol{A_3})$.

□ *Demonstração.* As duas propriedades de homomorfismo de anéis para $(\psi \circ \phi)$ seguem de propriedades análogas na seção de grupos e monoides.

1. Como ϕ é um homomorfismo de grupos entre $(A_1, +_1)$ e $(A_2, +_2)$ e ψ é homomorfismos de grupos entre $(A_2, +_2)$ e $(A_3, +_3)$, segue da proposição de composição de homomorfismos da seção de grupos (6.16) que $(\psi \circ \phi)$ é homomorfismo de grupos entre $(A_1, +_1)$ e $(A_3, +_3)$.
2. Como ϕ é um homomorfismo de monoides entre $(A_1, \cdot_1)$ e $(A_2, \cdot_2)$ e ψ é homomorfismos de monoides entre $(A_2, \cdot_2)$ e $(A_3, \cdot_3)$, segue da proposição de composição de homomorfismos da seção de monoides (5.31) que $(\psi \circ \phi)$ é homomorfismo de monoides entre $(A_1, \cdot_1)$ e $(A_3, \cdot_3)$.

■

⊢ **Proposição 7.16.** *Sejam* $\boldsymbol{A}$ *e* $\boldsymbol{B}$ *anéis,* $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$ *e* $I \trianglelefteq B$. *Então* $\phi^{-1}(I) \trianglelefteq A$.

□ *Demonstração.* Sejam $i_1, i_2 \in \phi^{-1}(I)$ e $a \in A$. Então, como $\phi(i_1), \phi(i_2) \in I$, temos $\phi(i_1 - i_1) = \phi(i_1) - \phi(i_2) \in I$, o que implica que $i_1 - i_2 \in \phi^{-1}(I)$. Ainda, como $\phi(a) \in A$, temos que $\phi(ai_1) = \phi(a)\phi(i_1) \in I$, o que implica $ai_1 \in \phi^{-1}(I)$. Logo $\phi^{-1}(I) \trianglelefteq A$. ■

$\vdash$ **Proposição 7.17.** *Sejam $\boldsymbol{A}$ e $\boldsymbol{B}$ anéis, $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$ sobrejetivo e $I \trianglelefteq A$. Então $\phi(I) \trianglelefteq B$.*

$\square$ *Demonstração.* Sejam $j_1, j_2 \in \phi(I)$ e $b \in B$. Então existem $i_1, i_2 \in I$ tais que $\phi(i_1) = j_1$ e $\phi(i_2) = j_2$ e, como ϕ é sobrejetiva, existe $a \in A$ tal que $\phi(a) = b$. Então, como $I \trianglelefteq A$, temos que $i_1 - i_2 \in I$ e $ai_1 \in I$, o que implica $j_1 - j_2 = \phi(i_1) - \phi(i_2) = \phi(i_1 - i_2) \in \phi(I)$ e $bj_1 = \phi(a)\phi(i_1) = \phi(ai_1) \in \phi(I)$. Logo $\phi(I) \trianglelefteq B$. ■

$:\vdash$ **Definição 7.10.** Sejam $\boldsymbol{A}$ e $\boldsymbol{B}$ dois anéis e $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$. O *núcleo* de ϕ é o conjunto

$$\mathrm{nuc}(\phi) := \{a \in A : \phi(a) = 0_B\}$$

e a *imagem* de ϕ é o conjunto

$$\mathrm{im}(\phi) := \{b \in B : \exists a \in A, \phi(a) = b\}.$$

$\vdash$ **Proposição 7.18.** *Sejam $\boldsymbol{A} = (A, +, \cdot)$ e $\boldsymbol{B} = (B, \oplus, \odot)$ dois anéis e seja $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$. Então*

1. $\mathrm{nuc}(\phi) \trianglelefteq A$;
2. $\mathrm{im}(\phi)$ *é subanel de* $\boldsymbol{B}$.

$\square$ *Demonstração.* Demonstramos as afirmações separadamente.

1. Primeiro notamos que $\mathrm{nuc}(\phi) \subseteq A$ e que $\mathrm{nuc}(\phi)$ não é vazio, pois, como $\phi(0_A) = 0_B$, então $0_A \in \mathrm{nuc}(\phi)$. Vamos mostrar as duas propriedades de um ideal. Sejam $a \in A$ e $n_1, n_2 \in \mathrm{nuc}(\phi)$. Então $n_1 - n_2 \in \mathrm{nuc}(\phi)$, pois

$$\phi(n_1 - n_2) = \phi(n_1) - \phi(n_2) = 0_B - 0_B = 0_B.$$

   Ainda, $a \cdot n_1 \in \mathrm{nuc}(\phi)$, pois

$$\phi(a \cdot n_1) = \phi(a) \odot \phi(n_1) = \phi(a) \odot 0_B = 0_B.$$

   Portanto $\mathrm{nuc}(\phi)$ é ideal de A.
2. Claramente, $\mathrm{im}(\phi) \subseteq B$ e $\mathrm{im}(\phi)$ não é vazio. Sejam $i_1, i_2 \in \mathrm{im}(\phi)$. Então existem $a_1, a_2 \in A$ tais que $\phi(a_1) = i_1$ e $\phi(a_2) = i_2$. Portanto $i_1 \ominus i_2 \in \mathrm{im}(\phi)$, já que

$$\phi(a_1 - a_2) = \phi(a_1) - \phi(a_2) = i_1 - i_2.$$

   Ainda, $i_1 \odot i_2 \in \mathrm{im}(\phi)$, pois

$$\phi(a_1 \cdot a_2) = \phi(a_1) \odot \phi(a_2) = i_1 \odot i_2.$$

   Por fim, $1_B \in \mathrm{im}(\phi)$, pois $\phi(1_A) = 1_B$. Logo $\mathrm{im}(\phi)$ é subanel de $\boldsymbol{B}$.

■

⊢ **Proposição 7.19.** *Sejam* $\boldsymbol{A} = (A, +, \cdot)$ *e* $\boldsymbol{B} = (B, \oplus, \odot)$ *dois anéis e seja* $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$. *Então* ϕ *é injetiva se, e somente se,* $\mathrm{nuc}(\phi) = \{0_A\}$.

□ *Demonstração.* Como ϕ é um homomorfismo de anéis, ele é um homomorfismo de grupos entre $(A, +)$ e $(B, \oplus)$. Então, pela proposição análoga da seção de grupos (6.20), esta proposição está provada. ■

:⊢ **Definição 7.11.** Sejam $\boldsymbol{A} = (A, +, \cdot)$ e $\boldsymbol{B} = (B, \oplus, \odot)$ dois anéis. Um isomorfismo de anéis é um homomorfismo de anéis $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$ que é bijetivo. O conjunto de todos os homomorfismos de anéis entre $\boldsymbol{A}$ e $\boldsymbol{B}$ é denotado por $\overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A}, \boldsymbol{B})$.

⊢ **Proposição 7.20.** *Sejam* $\boldsymbol{A} = (A, +, \cdot)$ *e* $\boldsymbol{B} = (B, \oplus, \odot)$ *anéis e* $\phi \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A}, \boldsymbol{B})$. *Então* $\phi^{-1} \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{B}, \boldsymbol{A})$.

□ *Demonstração.* Como ϕ é bijetiva, sua inversa ϕ^{-1} também é bijetiva. Devemos provar que ϕ^{-1} é um homomorfismo de anéis. Sejam $b_1, b_1 \in B$. Primeiro, vamos provar que ϕ^{-1} é um homomorfismo de grupos entre $\boldsymbol{B}$ e $\boldsymbol{A}$. Como ϕ é isomorfismo, existem $a_1, a_2 \in A$ tais que $\phi(a_1) = b_1$ e $\phi(a_2) = b_2$. Então

$$\begin{aligned}
\phi^{-1}(b_1 \oplus b_2) &= \phi^{-1}(\phi(a_1) \oplus \phi(a_2)) \\
&= \phi^{-1}(\phi(a_1 + a_2)) \\
&= a_1 + a_2 \\
&= \phi^{-1}(b_1) \oplus \phi^{-1}(b_2)
\end{aligned}$$

e

$$\phi^{-1}(\ominus b_1) = \phi^{-1}(\phi(-a_1)) = -a_1 = \ominus\phi(b_1).$$

Agora, mostramos que ϕ^{-1} é homomorfismo de monoides entre $(B, \odot)$ e $(A, \cdot)$.

$$\begin{aligned}
\phi^{-1}(b_1 \odot b_2) &= \phi^{-1}(\phi(a_1) \odot \phi(a_2)) \\
&= \phi^{-1}(\phi(a_1 \cdot a_2)) \\
&= a_1 \cdot a_2 \\
&= \phi^{-1}(b_1) \odot \phi^{-1}(b_2)
\end{aligned}$$

e, como $\phi(1_A) = 1_B$, temos $\phi^{-1}(1_B) = 1_A$. ■

:⊢ **Definição 7.12.** Sejam $\boldsymbol{A}$ e $\boldsymbol{B}$ dois anéis. Dizemos que $\boldsymbol{A}$ é *isomorfo* a $\boldsymbol{B}$, e denotamos isso por $\boldsymbol{A} \simeq \boldsymbol{B}$, sse existe $\phi \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A}, \boldsymbol{B})$.

⊢ **Proposição 7.21.** *Sejam* $\boldsymbol{A_1}$, $\boldsymbol{A_2}$ *e* $\boldsymbol{A_3}$ *três anéis. Então*

1. *(Reflexividade)* $\boldsymbol{A_1} \simeq \boldsymbol{A_1}$;

2. *(Antissimetria)* $\boldsymbol{A_1} \simeq \boldsymbol{A_2} \Rightarrow \boldsymbol{A_2} \simeq \boldsymbol{A_1}$*;*
3. *(Transitividade)* $\boldsymbol{A_1} \simeq \boldsymbol{A_2}$ *e* $\boldsymbol{A_2} \simeq \boldsymbol{A_3} \Rightarrow \boldsymbol{A_1} \simeq \boldsymbol{A_3}$*.*

□ *Demonstração.* Vamos demonstrar as três propriedades separadamente.

1. Claramente, a função $\phi = \mathrm{I}_A : A \to A$ é um isomorfismo de anéis. Logo $\boldsymbol{A_1} \simeq \boldsymbol{A_1}$
2. Se $\mathbf{A}_1 \simeq \boldsymbol{A_2}$, existe $\phi \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A_1}, \boldsymbol{A_2})$. Pela proposição (7.20), ϕ^{-1} é um isomorfismo de anéis entre $\boldsymbol{A_2}$ e $\boldsymbol{A_1}$. Logo $\boldsymbol{A_2} \simeq \boldsymbol{A_1}$.
3. $\boldsymbol{A_1} \simeq \boldsymbol{A_2}$, existe $\phi \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A_1}, \boldsymbol{A_2})$ e, como $\boldsymbol{A_2} \simeq \boldsymbol{A_3}$, existe $\psi \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A_2}, \boldsymbol{A_3})$. Assim, pela proposição (7.15), sabemos que $(\psi \circ \phi) \in \mathcal{H}(\boldsymbol{A_1}, \boldsymbol{A_3})$. Ainda, como ϕ e ψ são bijeções, sua composição é uma bijeção. Portanto $(\psi \circ \phi) \in \overset{\leftrightarrow}{\mathcal{H}}(\boldsymbol{A_1}, \boldsymbol{A_3})$, o que implica $\boldsymbol{A_1} \simeq \boldsymbol{A_3}$.

■

Não dizemos que $\simeq$ é uma relação de equivalência porque ela não está propriamente definida em um conjunto, já que não existe o conjunto de todos os anéis por não existir o conjunto de todos os conjuntos. No entanto, é claro que o que a proposição afirma é que ela satisfaz as propriedades de uma relação de equivalência.

7.1.4 Teoremas de isomorfismo

⊢ **Teorema 7.22** (Primeiro Teorema de Isomorfismo)**.** *Sejam* $\boldsymbol{A} = (A, +, \cdot)$ *e* $\boldsymbol{B} = (B, \oplus, \odot)$ *dois anéis e* $\phi \in \mathcal{H}(\boldsymbol{A}, \boldsymbol{B})$*. Então*

$$\boldsymbol{A}\big/\mathbf{nuc}(\boldsymbol{\varphi}) \simeq \mathbf{im}\,(\boldsymbol{\varphi}).$$

□ *Demonstração.* Primeiro, vale notar que, como $\mathrm{im}(\phi) \trianglelefteq A$, o $\boldsymbol{A} \diagup \mathbf{nuc}\,(\boldsymbol{\phi})$ é um anel. Agora, consideremos a função

$$\begin{aligned} \psi\colon A\big/\mathrm{nuc}(\phi) &\longrightarrow \mathrm{im}(\phi) \\ a + \mathrm{nuc}(\phi) &\longmapsto \phi(a). \end{aligned}$$

Notemos que a função ψ é bem definida. Para isso, sejam $a_1, a_2 \in A$ tais que $a_1 + \mathrm{nuc}(\phi) = a_2 + \mathrm{nuc}(\phi)$. Então $(a_1 - a_2) \in \mathrm{nuc}(\phi)$, o que implica $\phi(a_1 - a_2) = 0$. Como ϕ é homomorfismo de anéis, segue que $\phi(a_1) = \phi(a_2)$. Então $\psi(a_1 + \mathrm{nuc}(\phi)) = \psi(a_2 + \mathrm{nuc}(\phi))$.

Vamos mostrar que essa função é um isomorfismo de anéis. Primeiro, vamos mostrar que ψ é homomorfismo de anéis. Para isso, vamos denotar $a + \mathrm{nuc}(\phi) \in A \diagup \mathrm{nuc}(\phi)$ por $[a]$ para facilitar a demonstração. Sejam $[a_1], [a_2] \in A \diagup \mathrm{nuc}(\phi)$. Vemos que ψ é homomorfismo de grupos entre $(A \diagup \mathrm{nuc}(\phi), +)$ e $(\mathrm{im}(\phi), \oplus)$, pois

$$\psi([a_1] + [a_2]) = \psi([a_1 + a_2]) = \phi(a_1 + a_2) = \phi(a_1) \oplus \phi(a_2) = \psi([a_1]) \oplus \psi([a_2]).$$

Agora, vemos que ψ é homomorfismo de monoides entre $(A \diagup \operatorname{nuc}(\phi), \cdot)$ e $(\operatorname{im}(\phi), \odot)$, pois

$$\psi([a_1] \cdot [a_2]) = \psi([a_1 \cdot a_2]) = \phi(a_1 \cdot a_2) = \phi(a_1) \odot \phi(a_2) = \psi([a_1]) \odot \psi([a_2])$$

e $\psi([1_A]) = \phi(1_A) = 1_B$.

Por fim, devemos mostrar que ψ é bijetiva. Primeiro, mostramos que ψ é injetiva. Seja $[a] \in \operatorname{nuc}(\psi)$. Então $\psi([a]) = 0_B$, o que implica $\phi(a) = 0_B$. Mas isso implica $a \in \operatorname{nuc}(\phi)$; ou seja, $[a] = [0_A]$. Logo $\operatorname{nuc}(\psi) = \{[0_A]\}$, que quer dizer que ψ é injetiva (7.19). Agora, notamos que ψ é sobrejetiva por construção, pois tem como contradomínio $\operatorname{im}(\phi)$. ■

⊢ **Proposição 7.23** (Teorema Chinês do Resto)**.** *Sejam $m_1, \ldots, m_n \in \mathbb{N} \setminus \{0, 1\}$ dois a dois coprimos entre si. Então*

$$\mathbb{Z}\big/(m_1 m_2 \ldots m_n \mathbb{Z}) \simeq \left(\mathbb{Z}\big/m_1\mathbb{Z}\right) \times \cdots \times \left(\mathbb{Z}\big/m_n\mathbb{Z}\right).$$

:⊢ **Definição 7.13.**

$$B + I := \{b + i : b \in B, i \in I\}$$

⊢ **Teorema 7.24** (Segundo Teorema de Isomorfismo)**.** *Sejam $\boldsymbol{A}$ um anel, B um subanel de $\boldsymbol{A}$ e $I \trianglelefteq A$. Então*

1. *$B + I$ é subanel de $\boldsymbol{A}$;*
2. *$B \cap I \trianglelefteq B$;*
3. $$\boldsymbol{B}\big/\boldsymbol{B \cap I} \simeq \boldsymbol{B + I}\big/\boldsymbol{I}.$$

□ *Demonstração.* 1. Para mostrar que $B + I$ é subanel de $\boldsymbol{A}$, primeiro notamos que $B + I$ não é vazio, pois B não é vazio e I não é vazio. Ainda, notamos que $B + I \subseteq A$, pois $B \subseteq A$ e $I \subseteq A$. Agora, mostramos as propriedades de subanel. Sejam $b_1, b_2 \in B$ e $i_1, i_2 \in I$. Primeiro, mostraremos que $B + I$ é subgrupo de $(A, +)$. Note que $(b_1 + i_1) - (b_2 + i_2) \in B + I$, pois, como B é subgrupo de $(A, +)$, $b_1 - b_2 \in B$ e, como $I \trianglelefteq A$, $i_1 - i_2 \in A$, o que implica $(b_1 + i_1) - (b_2 + i_2) = (b_1 - b_2) + (i_1 - i_2) \in B + I$. Mostremos agora que $B + I$ é submonoide de $(A, \cdot)$. Note que

$$(b_1 + i_1)(b_2 + i_2) = b_1 b_2 + b_1 i_2 + i_1 b_2 + i_1 i_2.$$

Como B é submonoide de $(A, \cdot)$, $b_1 b_2 \in B$ e, como $i \trianglelefteq A$, $b_1 i_2, i_1 b_2, i_1 i_2 \in I$, o que implica $b_1 i_2 + i_1 b_2 + i_1 i_2 \in I$. Logo $(b_1 + i_1)(b_2 + i_2) = (b_1 b_2) + (b_1 i_2 + i_1 b_2 + i_1 i_2) \in B + I$. Ainda, $1 \in B$ e $0 \in I$. Logo $1 = 1 + 0 \in B + I$.

2. Para mostrar que $B \cap I \trianglelefteq B$, notemos primeiro que $B \cap I$ não é vazio. De fato, como B é subanel de $\boldsymbol{A}$, segue que $0 \in B$ e, como $I \trianglelefteq A$, também segue que $0 \in I$, o que implica $0 \in B \cap I$. Claramente, $B \cap I \subseteq A$. Então basta provar as propriedades de ideal. Sejam $b_1, b_2 \in B \cap I$. Como B é subanel de $\boldsymbol{A}$, temos $b_1 - b_2 \in B$ e, como $I \trianglelefteq A$, temos $b_1 - b_2 \in I$. Logo $b_1 - b_2 \in B \cap I$. Seja $b \in B$. Como B é subanel de $\boldsymbol{A}$, temos $bb_1 \in B$ e, como $I \trianglelefteq A$, temos $bb_1 \in I$. Logo $bb_1 \in B \cap I$.
3. O isomorfismo só faz sentido se os dois quocientes fazem sentido. O primeiro faz sentido pelo item anterior. O segundo faz sentido pois, pela definição de ideal, segue direto que $I \trianglelefteq B + I$, pois $I \subseteq B + I \subseteq A$. Então devemos exibir um isomorfismo de anéis entre os dois anéis. Considere a função

$$\begin{aligned} \phi\colon B &\longrightarrow {}^{B+I}\big/_{I} \\ b &\longmapsto b + I. \end{aligned}$$

Essa função é um homomorfismo de anéis. Sejam $b_1, b_2 \in B$. Então $\phi(b_1 + b_2) = (b_1 + b_2) + I = (b_1 + I) + (b_1 + I) = \phi(b_1) + \phi(b_2)$. Ainda, vale $\phi(b_1 b_2) = (b_1 b_2) + I = (b_1 + I)(b_1 + I) = \phi(b_1)\phi(b_2)$. Por fim, $\phi(1) = 1 + I$. Agora, notemos que $\mathrm{nuc}(\phi) = B \cap I$. Seja $b \in B$. Então

$$b \in \mathrm{nuc}(\phi) \Leftrightarrow \phi(b) = 0 \Leftrightarrow b + I = I \Leftrightarrow b \in I.$$

Por fim, notemos que $\mathrm{im}(\phi) = B + I/I$, pois um elemento de $B + I/I$ é da forma $b = i + I$, com $b \in B$ e $i \in I$. Mas então $b + i + I = b + I$. Logo segue do primeiro teorema de isomorfismo (7.22) que

$$\boldsymbol{B}\big/_{\boldsymbol{B} \cap \boldsymbol{I}} = \boldsymbol{B}\big/_{\mathbf{nuc}(\boldsymbol{\phi})} \simeq \mathbf{im}\,(\boldsymbol{\phi}) = \boldsymbol{B} + \boldsymbol{I}\big/_{\boldsymbol{I}}.$$

■

⊢ **Lema 7.25.** *Sejam $\boldsymbol{A}$ um anel e $I \trianglelefteq A$. Então*

1. *Se B é subanel de $\boldsymbol{A}$ tal que $I \subseteq B$, então B/I é subanel de $\boldsymbol{A/I}$. Por outro lado, todo subanel de $\boldsymbol{A/I}$ é da forma B/I para algum B subanel de $\boldsymbol{A}$ tal que $I \subseteq B$.*
2. *Se $J \trianglelefteq A$ tal que $I \subseteq J$, então $J/I \trianglelefteq A/I$. Por outro lado, todo ideal de A/I é da forma J/I para algum $J \trianglelefteq A$, tal que $I \subseteq J$.*

□ *Demonstração.* 1. Seja B um subanel de $\boldsymbol{A}$ tal que $I \subseteq B$. Para mostrar que o conjunto B/I é subanel de $\boldsymbol{A/I}$, sejam $b_1 + I, b_2 + I \in B/I$. Então, como $b_1, b_2 \in B$, vale $b_1 - b_2 \in B$ e $b_1 b_2 \in B$ e segue que $(b_1 + I) - (b_2 + I) = (b_1 - b_2) + I \in B/I$ e $(b_1 + I)(b_2 + I) = (b_1 b_2) + I \in B/I$. Ainda, como $1 \in B$, $1 + I \in B/I$. Logo B/I é subanel de $\boldsymbol{A/I}$.

Seja agora C um subanel de $\boldsymbol{A/I}$. Como C é subconjunto não vazio de $\boldsymbol{A/I}$, é da forma $C = \{b + I : b \in B\}$, com $I \subseteq B \subseteq A$; ou seja, $C = B/I$. Vamos mostrar que B é subanel de $\boldsymbol{A}$. Sejam $b_1, b_2 \in B$. Como B/I é subanel de $\boldsymbol{A/I}$, temos que $(b_1 + I) - (b_2 + I) = (b_1 - b_2) + I \in B/I$ e $(b_1 + I)(b_2 + I) = (b_1 b_2) + I \in B/I$. Então $b_1 - b_2 \in B$ e $b_1 b_2 \in B$. Ainda, como $1 + I \in B/I$, temos que $1 \in B$, e a demonstração está completa.

2. Seja $J \trianglelefteq A$ tal que $I \subseteq J$. Para mostrar que $J/I \trianglelefteq A/I$, sejam $j_1+I, j_2+I \in J/I$ e $a + I \in A/I$. Como $J \trianglelefteq A$, vale $j_1 - j_2 \in J$ e $aj_1 \in J$. Então $(j_1 + I) - (j_2 + I) = (j_1 - j_2) + I \in J/I$ e $(a + I)(j_1 + I) = (aj_1) + I \in J/I$. Seja $C \trianglelefteq A/I$. Como C é subconjunto de A/I, é da forma $C = \{j + I : j \in J\}$, com $I \subseteq J \subseteq A$; ou seja, $C = J/I$. Vamos mostrar que $J \trianglelefteq A$. Sejam $j_1, j_2 \in J$ e $a \in A$. Como $J/I \trianglelefteq A/I$, temos que $(j_1 + I) - (j_2 + I) = (j_1 - j_2) + I \in J/I$ e $(a + I)(j_1 + I) = (aj_1) + I \in J/I$. Então $j_1 - j_2 \in J$ e $aj_1 \in J$, e a demonstração está completa.

■

⊢ **Teorema 7.26** (Terceiro Teorema de Isomorfismo)**.** *Sejam* $\boldsymbol{A}$ *um anel,* $I \trianglelefteq A$ *e* $J \trianglelefteq A$ *tais que* $I \subseteq J$*. Então*

$$\boldsymbol{(A/I)}\big/\boldsymbol{(J/I)} \simeq \boldsymbol{A}\big/\boldsymbol{J}.$$

□ *Demonstração.* Consideremos a função

$$\begin{aligned} \phi\colon A\big/I &\longrightarrow A\big/J \\ a + I &\longmapsto a + J. \end{aligned}$$

Primeiro, notemos que ϕ é bem definida, pois, se $a_1 + I = a_2 + I$, então $a_1 - a_2 \in I \subseteq J$, o que implica $a_1 + J = a_2 + J$. Agora, provemos que ϕ é homomorfismo de anéis. Sejam $a_1 + I, a_2 + I \in A \diagup I$. Então

$$\begin{aligned} \phi((a_1 + I) + (a_2 + I)) &= \phi((a_1 + a_2) + I) \\ &= (a_1 + a_2) + J \\ &= (a_1 + J) + (a_2 + J) \\ &= \phi(a_1) + \phi(a_2). \end{aligned}$$

Também, vale que

$$\begin{aligned} \phi((a_1 + I)(a_2 + I)) &= \phi((a_1 a_2) + I) \\ &= (a_1 a_2) + J \\ &= (a_1 + J)(a_2 + J) \\ &= \phi(a_1)\phi(a_2). \end{aligned}$$

Por fim, notamos que $\phi(1+I) = 1+J$. Assim, provamos que ϕ é homomorfismo de anéis.

Agora, notemos que

$$\operatorname{nuc}(\phi) = \{a+I : \phi(a+I) = 0+J\} = \{a+I : a \in J\} = {}^{J}\!/_{I},$$

o que prova que $J/I \trianglelefteq A/I$ e que, portanto, o quociente $(A/I)/(J/I)$ pode formar um anel quociente. Notemos também que ϕ é sobrejetiva por construção; ou seja, $\operatorname{im}(\phi) = A/J$. Logo, pelo primeiro teorema de isomorfismo (7.22), temos que

$$\boldsymbol{(A/I)/(J/I) = (A/I)/\operatorname{nuc}(\phi) \simeq \operatorname{im}(\phi) = A/J}.$$

■

7.1.5 Produto de anéis

:⊢ **Definição 7.14.** Seja $(\boldsymbol{A_i})_{i\in I} = (A_i, +_i, \times_i)_{i\in I}$ uma família de anéis. O *produto* da família $(\boldsymbol{A_i})_{i\in I}$ é a tripla

$$\prod_{i\in I} \boldsymbol{A_i} := (A, +, \times)$$

em que $A = \prod_{i\in I} A_i$ é o produto de conjuntos,

$$\begin{aligned} +\colon A \times A &\longrightarrow A \\ (a,b) &\longmapsto (a_i +_i b_i)_{i\in I} \end{aligned}$$

e

$$\begin{aligned} \times\colon A \times A &\longrightarrow A \\ (a,b) &\longmapsto (a_i \times_i b_i)_{i\in I}. \end{aligned}$$

Denotaremos as operações $+_i$ todas por $+$ e as operações $\times_i$ todas por $\times$ quando não existir ambiguidade.

⊢ **Proposição 7.27.** *Seja $(\boldsymbol{A_i})_{i\in I} = (A_i, +_i, \times_i)_{i\in I}$ uma família de anéis. Então o produto $\prod_{i\in I} \boldsymbol{A_i}$ é um anel.*

□ *Demonstração.* Como, para todo $i \in I$, o par $(A_i, +_i)$ é um grupo comutativo, segue que o par $(\prod_{i\in I} A_i, +)$ é um grupo comutativo.

Agora, devemos mostrar que $(\prod_{i=1}^{n} A_i, \times)$ é um monoide comutativo. Novamente, as propriedades de monoide comutativo decorrem do fato de que $(A_i, \times_i)$, $i \in I$,

são todos monoides comutativos com elementos neutros 1_i, respectivamente. Sejam $a = (a_i)_{i\in I}, b = (b_i)_{i\in I}, c = (c_i)_{i\in I} \in \prod_{i\in I} A_i$. A operação $\times$ é associativa, pois

$$\begin{aligned}(a \times b) \times c &= (a_i \times_i b_i)_{i\in I} \times c \\ &= ((a_i \times_i b_i) \times_i c_i)_{i\in I} \\ &= (a_i \times_i (b_i \times_i c_i))_{i\in I} \\ &= a \times (b_i \times_i c_i)_{i\in I} \\ &= a \times (b \times c),\end{aligned}$$

e comutativa, pois

$$a \times b = (a_i \times_i b_i)_{i\in I} = (b_i \times_i a_i)_{i\in I} = b \times a.$$

Ainda, $1 := (1_i)_{i\in I}$ é elemento neutro, pois

$$a \times 1 = (a_i \times 1_i)_{i\in I} = (a_i)_{i\in I} = a.$$

Por fim, como $\times_i$ são distributivas sobre $+_i$, temos que

$$\begin{aligned}a \times (b + c) &= a \times (b_i +_i c_i)_{i\in I} \\ &= (a_i \times_i (b_i +_i c_i))_{i\in I} \\ &= ((a_i \times_i b_i) +_i (a_i \times_i c_i))_{i\in I} \\ &= (a_i \times_i b_i)_{i\in I} + (a_i \times_i c_i)_{i\in I} \\ &= (a \times b) + (a \times c).\end{aligned}$$

■

7.1.6 Domínios e corpos

:⊢ **Definição 7.15.** Um *domínio* (ou *domínio integral*, ou *anel integral*) é um anel não trivial $\boldsymbol{A}$ tal que, para todos $a, a' \in A$ tais que $aa' = 0$, vale que $a = 0$ ou $a' = 0$.

:⊢ **Definição 7.16.** Um *corpo* é um anel $\boldsymbol{A}$ tal que todo $a \in A \smallsetminus \{0\}$ tem inverso multiplicativo $a^{-1} \in A \smallsetminus \{0\}$; ou seja, para todo $a \in A \smallsetminus \{0\}$, existe $a' \in A$ tal que $aa' = 1 = a'a$.

O inverso multiplicativo a' de a é único e denotado a^{-1}. A definição de corpo é equivalente a dizer que $(A \smallsetminus \{0\}, \times, \diagup, 1)$ é um grupo.

⊢ **Proposição 7.28.** *Todo corpo* $\boldsymbol{A}$ *é um domínio.*

□ *Demonstração.* Sejam $a_1, a_2 \in A$ tais que $a_1 \cdot a_2 = 0$. Suponhamos que $a_2 \neq 0$. Então existe $a_2^{-1} \in A$ e temos

$$\begin{aligned} a_1 &= a_1 \cdot 1 \\ &= a_1 \cdot (a_2 \cdot a_2^{-1}) \\ &= (a_1 \cdot a_2) \cdot a_2^{-1} \\ &= 0 \cdot a_2^{-1} \\ &= 0. \end{aligned}$$

Logo, se $a_1 \cdot a_2 = 0$, $a_1 = 0$ ou $a_2 = 0$. ■

⊢ **Proposição 7.29** (Lei do corte em domínios)**.** *Sejam* $\boldsymbol{A}$ *um domínio e* $a, a_1, a_2 \in A$*,* $a \neq 0$*. Então*

$$aa_1 = aa_2 \quad \Longrightarrow \quad a_1 = a_2$$

□ *Demonstração.* Se $aa_1 = aa_2$, então $-aa_1 = -aa_2$. Portanto

$$a(a_1 - a_2) = aa_1 - aa_2 = aa_1 - aa_1 = 0.$$

Logo, como $\boldsymbol{A}$ é um domínio e $a \neq 0$, temos que $a_1 - a_2 = 0$, o que implica $a_1 = a_2$. ■

Essa proposição é interessante pois, mesmo sem exigir que $(A \setminus \{0\}, \times, /, 1)$ seja um grupo, como no caso de $\boldsymbol{A}$ ser um corpo, se $\mathbf{A}$ for um domínio, vale a lei do corte da multiplicação para elementos de $A \setminus \{0\}$.

7.1.7 Ideais primos e ideais maximais

:⊢ **Definição 7.17.** Seja $\boldsymbol{A}$ um anel. Um ideal *primo* de $\boldsymbol{A}$ é um ideal próprio $I \lhd A$ tal que

1. $\forall a, b \in A \qquad ab \in I \Rightarrow a \in I \;$ ou $\; b \in I$.

▷ **Exercício 7.1.** *Sejam* $\boldsymbol{A}$ *um anel e* $I \lhd A$ *um ideal primo. Para todos* $a \in A$ *e* $n \in \mathbb{N}$*, se* $a^n \in I$*, então* $a \in I$*.*

▷ **Exercício 7.2.** *Sejam* $\boldsymbol{A}$ *um anel e* $I \lhd A$ *um ideal. Então* I *é um ideal primo de* $\boldsymbol{A}$ *se, e somente se,* $I^{\complement}$ *é um submonoide multiplicativo de* $\boldsymbol{A}^{\times}$*.*

⊢ **Teorema 7.30.** *Seja* $\boldsymbol{A}$ *um anel e* $I \lhd A$*. Então* I *é um ideal primo de* $\boldsymbol{A}$ *se, e somente se,* $\boldsymbol{A}/\boldsymbol{I}$ *é um domínio.*

□ *Demonstração.* Vamos demonstrar as duas implicações ao mesmo tempo. Sejam $a, b \in A$ e $\alpha, \beta \in A/I$ tais que $\alpha = a + I$ e $\beta = b + I$. Note que $\boldsymbol{A/I}$ é um domínio se, e somente se,

$$\forall \alpha, \beta \in {}^{A}\!/_{I} \qquad \alpha\beta = 0_{A/I} \Rightarrow \alpha = 0_{A/I} \ \text{ ou } \ \beta = 0_{A/I}.$$

Mas $\alpha\beta = (a + I)(b + I) = ab + I$ e $0_{A/I} = 0 + I$. Ainda, para qualquer $a' \in A$, $a' + I = 0 + I$ se, e somente se, $a' \in I$. Logo segue que a implicação acima é equivalente a

$$\forall a, b \in A \qquad ab = 0 \Rightarrow a = 0 \ \text{ ou } \ b = 0,$$

que é a definição de ideal primo. ■

:⊢ **Definição 7.18.** Seja $\boldsymbol{A}$ um anel. Um ideal *maximal* de $\boldsymbol{A}$ é um ideal $I \lhd A$ tal que

$$\forall J \in \wp(A) \qquad I \subseteq J \ \text{ e } \ J \trianglelefteq A \Rightarrow J = I \ \text{ ou } \ J = A.$$

⊢ **Teorema 7.31.** *Seja $\boldsymbol{A}$ um anel e $I \lhd A$. Então I é um ideal maximal de $\boldsymbol{A}$ se, e somente se, $\boldsymbol{A/I}$ é um corpo.*

□ *Demonstração.* Em ambas implicações, usaremos o fato que um anel é um corpo se, e somente se, seus únicos ideais são o anel trivial e o anel todo (7.8).

- ⇐ Suponha que I é um ideal maximal de $\boldsymbol{A}$. Vamos mostrar que os únicos ideais de $\boldsymbol{A/I}$ são $\{0 + I\}$ e A/I. Seja $L \trianglelefteq A/I$ e consideremos a projeção canônica

$$\begin{aligned} \pi \colon A &\longrightarrow {}^{A}\!/_{I} \\ a &\longmapsto a + I. \end{aligned}$$

  Sabemos que $\pi^{-1}(L) = \{a \in A : \pi(a) \in L\} \trianglelefteq A$ (7.16). Como $0+I = 0_{A/I} \in L$, isso implica que $\pi^{-1}(0+I) \subseteq \pi^{-1}(L)$. Notando que $\pi^{-1}(0+I) = \mathrm{nuc}(\phi) = I$, temos que $I \subseteq \pi^{-1}(L) \subseteq A$. Como I é ideal maximal, então $\pi^{-1}(L) = I$ ou $\pi^{-1}(L) = A$. Vamos então avaliar os dois casos. Para isso, ressaltamos antes que $L = \pi(\pi^{-1}(L))$. No primeiro caso, $L = \pi(\pi^{-1}(L)) = \pi(I) = 0+I = 0_{A/I}$. No segundo caso, $L = \pi(\pi^{-1}(L)) = \pi(A) = A/I$. Logo A/I é um corpo.
- ⇒ Suponha que $\boldsymbol{A/I}$ é um corpo. Consideremos $J \trianglelefteq A$ tal que $I \subseteq J$ e a projeção canônica $\pi : A \to A/I$. Como π é um homomorfismo de anéis bijetivo, temos que $\pi(J) \trianglelefteq A/I$ (7.17). Como $\boldsymbol{A/I}$ é corpo, $\pi(J) = 0_{A/I}$ ou $\pi(J) = A/I$. No primeiro caso, $J = \mathrm{nuc}(\pi) = I$. No segundo caso, $J = \pi^{-1}(\pi(J)) = \pi^{-1}(A/I) = A$. Logo I é ideal maximal de A. ■

⊢ **Proposição 7.32.** *Seja $\boldsymbol{A}$ um anel e $I \lhd A$ um ideal maximal. Então I é ideal primo.*

□ *Demonstração.* A demonstração é simples. Sabemos que, se I é ideal maximal, então A/I é um corpo (7.31). Mas isso implica que A/I é um domínio (7.28). Concluímos, portanto, que I é um ideal primo (7.30). ■

:⊢ **Definição 7.19.** Seja $\boldsymbol{A}$ um anel. O *espectro* de $\boldsymbol{A}$ é o conjunto de ideal primos de $\boldsymbol{A}$. Denota-se $\mathrm{Esp}(\boldsymbol{A})$.

7.1.8 Radicais

:⊢ **Definição 7.20.** Sejam $\boldsymbol{A}$ um anel e $I \trianglelefteq A$ um ideal. O *radical* de I é o conjunto

$$\sqrt{I} := \{r \in A \mid \exists_{n\in\mathbb{N}}\, r^n \in I\}\,.$$

⊢ **Proposição 7.33.** *Sejam $\boldsymbol{A}$ um anel e $I \trianglelefteq A$.*

1. $I \subseteq \sqrt{I}$;
2. $\sqrt{I} \trianglelefteq A$.

□ *Demonstração.*

1. Seja $i \in I$. Como $i = i^1$ e $1 \in \mathbb{N}$, segue que $i \in \sqrt{I}$, logo $i \subseteq \sqrt{I}$.
2. Demonstramos por partes.
   - 2.1. (Não-vacuidade) Como $0 \in I$, segue que $0 \in \sqrt{I}$.
   - 2.2. (Fechamento da soma) Sejam $r, r' \in \sqrt{I}$ e $n, n' \in \mathbb{N}$ tais que $r^n, r'^{n'} \in I$. Temos
   $$(r+r')^{n+n'+1} = \mathop{+}_{k=0}^{n+n'+1} \binom{n+n'+1}{k} r^k r'^{n+n'+1-k}.$$
   Como para todo k vale $k \geq n$ ou $n+n'+1-k \geq n'$, segue que $r^k r'^{n+n'+1-k} \in I$, portanto $(r+r')^{n+n'+1} \in I$, o que implica que $r+r' \in \sqrt{I}$.
   - 2.3. (Invertibilidade) Sejam $r \in \sqrt{I}$ e $n \in \mathbb{N}$ tal que $r^n \in I$. Então $(-r)^n = (-1)^n r^n \in I$, portanto $-r \in \sqrt{I}$.
   - 2.4. (Absorção da multiplicação) Sejam $a \in A$, $r \in \sqrt{I}$ e $n \in \mathbb{N}$ tal que $r^n \in I$. Então $(ar)^n = a^n r^n \in I$, portanto $ar \in \sqrt{I}$. ■

:⊢ **Definição 7.21.** Seja $\boldsymbol{A}$ um anel. O *nilradical* de $\boldsymbol{A}$ é o radical $\sqrt{0} := \sqrt{\{0\}}$ e um elemento *nilpotente* de $\boldsymbol{A}$ é um elemento de $\sqrt{0}$, ou seja, um elemento $r \in A$ tal que, para algum $n \in \mathbb{N}$, $r^n = 0$.

⊢ **Proposição 7.34.** *Seja $\boldsymbol{A}$ um anel.*

$$\sqrt{0} = \bigcap_{P\in\mathrm{Esp}(\boldsymbol{A})} P.$$

□ *Demonstração.* ($\subseteq$) Sejam $r \in \sqrt{0}$ e $n \in \mathbb{N}$ tal que $r^n = 0$. Para todo ideal primo $P \in \mathrm{Esp}(\boldsymbol{A})$, segue que $r^n \in P$; como P é primo, segue que $r \in P$, o que implica que $r \in \bigcap_{P \in \mathrm{Esp}(\boldsymbol{A})} P$.

($\supseteq$) Seja $r \notin \sqrt{0}$. Consideremos o conjunto Σ de ideais $I \lhd A$ tais que, para todo $n \in \mathbb{N}$, $r^n \notin I$. Notemos que Σ não é vazio, pois $\{0\} \in \Sigma$, já que $r \notin \{0\}$. Pelo lema de Zorn, Σ tem elemento maximal $P \in \Sigma$. Mostremos que P é primo. Sejam $a, a' \in P^{\complement}$. Os ideais $P + aA$ e $P + a'A$ contém P estritamente, portanto não pertencem a Σ, o que implica, para alguns $n, n' \in \mathbb{N}$, $r^n \in P + aA$ e $r^{n'} \in P + a'A$. Segue que $r^{n+n'} \in P + aa'A$, logo que $P + aa'A \notin \Sigma$, o que implica que $aa' \in P^{\complement}$, logo que P é primo. Como em particular $r = r^1 \notin P$, pois $P \in \Sigma$, segue que $r \notin \bigcap_{P \in \mathrm{Esp}(\boldsymbol{A})} P$.

■

▷ **Exercício 7.3.** *Sejam $\boldsymbol{A}$ um anel, $I \trianglelefteq A$ um ideal e $\mathrm{p}\colon A \longrightarrow A/I$ a projeção quociente.*

$$\sqrt{I} = \mathrm{p}^{-1}\left(\sqrt{0}_{A \,/\, I}\right).$$

⊢ **Proposição 7.35.** *Seja $\boldsymbol{A}$ um anel.*

$$\sqrt{I} = \bigcap \{P \in \mathrm{Esp}(\boldsymbol{A}) \mid I \subseteq P\}.$$

□ *Demonstração.* Segue direto dos resultados anteriores, pois

$$\begin{aligned}\sqrt{I} &= \mathrm{p}^{-1}\left(\sqrt{0}_{A \,/\, I}\right)\\ &= \mathrm{p}^{-1}\left(\bigcap_{P \in \mathrm{Esp}(\boldsymbol{A} \,/\, \boldsymbol{I})} P\right)\\ &= \bigcap \{P \in \mathrm{Esp}(\boldsymbol{A}) \mid I \subseteq P\}.\end{aligned}$$

■

7.1.9 Matrizes

:⊢ **Definição 7.22.** Seja $\boldsymbol{A}$ um anel e $l, c \in \mathbb{N}$. Uma *matriz* de dimensão $l \times c$ sobre $\boldsymbol{A}$ é uma função $M : [l] \times [c] \to A$. Representa-se isso por

$$M = \begin{bmatrix} m_{1,1} & \cdots & m_{1,c} \\ \vdots & \ddots & \vdots \\ m_{l,1} & \cdots & m_{l,c} \end{bmatrix},$$

em que $m_{i,j} := M(i,j) \in A$. O conjunto $[l]$ é o conjunto dos *índices das linhas* e $[c]$ é o conjunto dos *índices das colunas* da matriz M. A imagem de M é o conjunto das *entradas* da matriz M e o elemento $m_{i,j}$ é a entrada da linha i e coluna j.

O conjunto de todas as matrizes de dimensão $l \times c$ sobre $\boldsymbol{A}$ é denotado por $\mathbb{M}_{l \times c}(\boldsymbol{A})$.

:⊢ **Definição 7.23.** Seja $\boldsymbol{A}$ um anel e $d \in \mathbb{N}$. Uma *matriz quadrada* de dimensão d sobre $\boldsymbol{A}$ é uma matriz $M \in \mathbb{M}_{d\times d}(\boldsymbol{A})$. O conjunto de todas as matrizes quadradas de dimensão d sobre $\boldsymbol{A}$ é denotado por $\mathbb{M}_d(\boldsymbol{A})$.

:⊢ **Definição 7.24.** Sejam $\boldsymbol{A}$ um anel e $M \in \mathbb{M}_{l\times c}(\boldsymbol{A})$. A *matriz transposta* de M é a matriz $M^\intercal \in \mathbb{M}_{c\times l}(\boldsymbol{A})$ definida por

$$(M^\intercal)(i,j) := m_{j,i}.$$

7.1.9.1 Soma de matrizes

:⊢ **Definição 7.25.** Sejam $\boldsymbol{A}$ um anel e $M, N \in \mathbb{M}_{l\times c}(\boldsymbol{A})$. A *matriz soma* das matrizes M e N é a matriz $(M+N) \in \mathbb{M}_{l\times c}(\boldsymbol{A})$ definida por

$$(M+N)(i,j) := m_{i,j} + n_{i,j}.$$

:⊢ **Definição 7.26.** Sejam $\boldsymbol{A}$ um anel e 0 o elemento neutro da soma de $\boldsymbol{A}$.

1. A *matriz nula* de dimensão $l \times c$ sobre $\boldsymbol{A}$ é a matriz $\mathbb{O}_{l\times c} \in \mathbb{M}_{l\times c}(\boldsymbol{A})$, definida por
$$\mathbb{O}_{l\times c}(i,j) := 0.$$
2. Se $M \in \mathbb{M}_{l\times c}$, a *matriz negativa* de M é a matriz $-M \in \mathbb{M}_{l\times c}(\boldsymbol{A})$, definida por
$$(-M)(i,j) := -m_{i,j}.$$

⊢ **Proposição 7.36.** *Seja $\boldsymbol{A}$ um anel e $+$ a operação binária em $\mathbb{M}_{l\times c}(\boldsymbol{A})$ definida por*

$$\begin{aligned} +\colon \mathbb{M}_{l\times c}(\boldsymbol{A}) \times \mathbb{M}_{l\times c}(\boldsymbol{A}) &\longrightarrow \mathbb{M}_{l\times c}(\boldsymbol{A}) \\ (M,N) &\longmapsto M+N. \end{aligned}$$

Então $(\mathbb{M}_{l\times c}(\boldsymbol{A}), +)$ é um grupo com elemento neutro $\mathbb{O}_{l\times c}$. Se $\boldsymbol{A}$ é comutativo, então $(\mathbb{M}_{l\times c}(\boldsymbol{A}), +)$ é comutativo.

□ *Demonstração.* Sejam $M, N, P \in \mathbb{M}_{l\times c}(\boldsymbol{A})$. Primeiro, notemos que $+$ é associativa, pois, como a soma no anel é associativa, segue que

$$(m_{i,j} + n_{i,j}) + p_{i,j} = m_{i,j} + (n_{i,j} + p_{i,j})$$

e, portanto, $(M+N)+P = M+(N+P)$. Então, notemos que $\mathbb{O}$ é elemento neutro de $+$. Como 0 é elemento neutro da soma da anel, segue que

$$m_{i,j} + 0 = 0 + m_{i,j} = m_{i,j}$$

e, portanto, $M + \mathbb{O} = \mathbb{O} + M = M$. Ainda, notemos que, como $-m_{i,j}$ é o inverso aditivo de $m_{i,j}$ no anel, segue que

$$m_{i,j} + (-m_{i,j}) = (-m_{i,j}) + m_{i,j} = 0$$

e, portanto, $M + (-M) = (-M) + M = \mathbb{O}$. Assim, concluímos que $(\mathbb{M}_{l\times c}(\boldsymbol{A}), +)$ é um anel. Por fim, notemos que, se $\boldsymbol{A}$ é comutativo, então $+$ é comutativa, pois, como a soma no anel é comutativa, segue que

$$m_{i,j} + n_{i,j} = n_{i,j} + m_{i,j}$$

e, portanto, $M + N = N + M$. Assim, concluímos que, se $\boldsymbol{A}$ é comutativo, então $(\mathbb{M}_{l\times c}(\boldsymbol{A}), +)$ é comutativo. ■

7.1.9.2 Produto de matrizes e produto por escalar

:⊢ **Definição 7.27.** Sejam $\boldsymbol{A}$ um anel e $M \in \mathbb{M}_{l\times d}(\boldsymbol{A})$ e $N \in \mathbb{M}_{d\times c}(\boldsymbol{A})$. A *matriz produto* das matrizes M e N é a matriz $(MN) \in \mathbb{M}_{l\times c}(\boldsymbol{A})$ definida por

$$(MN)(i,j) := \mathop{\mathbin{+}}_{k=1}^{d} m_{i,k} n_{k,j}.$$

⊢ **Proposição 7.37.** *Sejam $\boldsymbol{A}$ um anel e $M \in \mathbb{M}_{l\times d_1}(\boldsymbol{A})$, $N \in \mathbb{M}_{d_1\times d_2}(\boldsymbol{A})$ e $P \in \mathbb{M}_{d_2\times c}(\boldsymbol{A})$. Então*

$$(MN)P = M(NP).$$

□ *Demonstração.* Um elemento de MN é dado por

$$(mn)_{i,j} = \mathop{\mathbin{+}}_{k_1=1}^{d_1} m_{i,k_1} n_{k_1,j}.$$

Logo, um elemento de $(MN)P$ é dado por

$$((mn)p)_{i,j} = \mathop{\mathbin{+}}_{k_2=1}^{d_2} (mn)_{i,k_2} p_{k_2,j} = \mathop{\mathbin{+}}_{k_2=1}^{d_2} \left(\mathop{\mathbin{+}}_{k_1=1}^{d_1} m_{i,k_1} n_{k_1,k_2} \right) p_{k_2,j}.$$

Analogamente, um elemento de $M(NP)$ é dado por

$$(m(np))_{i,j} = \mathop{\mathbin{+}}_{k_1=1}^{d_1} m_{i,k_1} (np)_{k_1,j} = \mathop{\mathbin{+}}_{k_1=1}^{d_1} m_{i,k_1} \left(\mathop{\mathbin{+}}_{k_2=1}^{d_2} n_{k_1,k_2} p_{k_2,j} \right).$$

Mas então, como $\boldsymbol{A}$ é um anel, segue que

$$\begin{aligned}
((mn)p)_{i,j} &= \mathop{\plus}_{k_2=1}^{d_2} \left(\mathop{\plus}_{k_1=1}^{d_1} m_{i,k_1} n_{k_1,k_2} \right) p_{k_2,j} \\
&= \mathop{\plus}_{k_2=1}^{d_2} \left(\mathop{\plus}_{k_1=1}^{d_1} m_{i,k_1} n_{k_1,k_2} p_{k_2,j} \right) \\
&= \mathop{\plus}_{k_1=1}^{d_1} \left(\mathop{\plus}_{k_2=1}^{d_2} m_{i,k_1} n_{k_1,k_2} p_{k_2,j} \right) \\
&= \mathop{\plus}_{k_1=1}^{d_1} m_{i,k_1} \left(\mathop{\plus}_{k_2=1}^{d_2} n_{k_1,k_2} p_{k_2,j} \right) \\
&= (m(np))_{i,j}.
\end{aligned}$$

■

:⊢ **Definição 7.28.** Sejam $\boldsymbol{A}$ um anel e 0 e 1 os elementos neutros da soma e da multiplicação de $\boldsymbol{A}$ respectivamente. A *matriz identidade* de dimensão d sobre $\boldsymbol{A}$ é a matriz $\mathrm{I}_d \in \mathbb{M}_d(\boldsymbol{A})$ definida por

$$\mathrm{I}_d(i,j) := \delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}$$

Essa função δ é conhecida como delta de Kronecker.

⊢ **Proposição 7.38.** *Seja $\boldsymbol{A}$ um anel e $\cdot$ a operação binária em $\mathbb{M}_d(\boldsymbol{A})$ definida por*

$$\begin{aligned}
\cdot \colon \mathbb{M}_d(\boldsymbol{A}) \times \mathbb{M}_d(\boldsymbol{A}) &\longrightarrow \mathbb{M}_d(\boldsymbol{A} \\
(M,N) &\longmapsto MN.
\end{aligned}$$

Então $(\mathbb{M}_d(\boldsymbol{A}), \cdot)$ é um monoide com elemento neutro I_d.

□ *Demonstração.* Sejam $M, N, P \in \mathbb{M}_d(\boldsymbol{A})$. Pela proposição anterior, sabemos que vale $(MN)P = M(NP)$ e que, portanto, $\cdot$ é associativa. Agora, notemos que um elemento de $M1_d$ é da forma

$$\mathop{\plus}_{k=0}^{d-1} m_{i,k}\delta_{k,j}.$$

Mas, para $k \in [d]$, se $k \neq j$, então $\delta_{k,j} = 0$ e, se $k = j$, então $\delta_{k,j} = 1$ e, portanto, segue que

$$\mathop{\plus}_{k=1}^{d} m_{i,k}\delta_{k,j} = m_{i,j}.$$

Assim, concluímos que $M1_d = M$. Analogamente, mostra-se que $\mathrm{I}_d M = M$, e concluímos que I_d é elemento neutro de $\cdot$. Isso mostra que $(\mathbb{M}_d(\boldsymbol{A}), \cdot)$ é um monoide. ■

:⊢ **Definição 7.29.** Seja $\boldsymbol{A}$ um anel. Uma *matriz invertível* é uma matriz $M \in \mathbb{M}_d(\boldsymbol{A})$ que é invertível com respeito ao produto do monoide $(\mathbb{M}_d(\boldsymbol{A}), \cdot)$. A matriz inversa de M é denotada M^{-1}.

:⊢ **Definição 7.30.** Seja $\boldsymbol{A}$ um anel, $a \in A$ e $M \in \mathbb{M}_{l\times c}(\boldsymbol{A})$. O *produto por escalar* de a e M é a matriz $aM \in \mathbb{M}_{l\times c}$ definida por

$$(aM)(i,j) := am_{i,j}.$$

7.1.9.3 Matrizes quadradas

:⊢ **Definição 7.31.** Seja $\boldsymbol{A}$ um anel.

1. Uma *matriz triangular superior* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que satisfaz
$$\forall i,j \in [d] \qquad i > j \Rightarrow m_{i,j} = 0.$$
2. Uma *matriz triangular inferior* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ sobre $\boldsymbol{A}$ que satisfaz
$$\forall i,j \in [d] \qquad i < j \Rightarrow m_{i,j} = 0.$$
3. Uma *matriz triangular superior* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que satisfaz
$$\forall i,j \in [d] \qquad i > j \Rightarrow m_{i,j} = 0.$$
4. Uma *matriz triangular* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que é triangular superior ou triangular inferior.
5. Uma *matriz diagonal* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que é triangular superior e triangular inferior; ou seja, que satisfaz
$$\forall i,j \in [d] \qquad i \neq j \Rightarrow m_{i,j} = 0.$$
6. Uma *matriz simétrica* uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que é igual a sua transposta
$$M = M^{\intercal}.$$
7. Uma *matriz antissimétrica* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ que é igual à negativa da sua transposta
$$M = -M^{\intercal}.$$
8. Uma *matriz ortogonal* é uma matriz quadrada $M \in \mathbb{M}_d(\boldsymbol{A})$ cuja transposta é igual à sua inversa
$$M^{\intercal} = M^{-1}.$$

7.1.9.4 Traço

:⊢ **Definição 7.32.** Sejam $\boldsymbol{A}$ um anel e $M \in \mathbb{M}_d(\boldsymbol{A})$. O *traço* de M é o elemento $\operatorname{tr} M \in A$ definido por

$$\operatorname{tr} M := \mathop{\mathbin{+}}_{i=1}^{d} m_{i,i}.$$

⊢ **Proposição 7.39.** *Sejam $\boldsymbol{A}$ um anel, $a \in A$ e $M, N \in \mathbb{M}_d(\boldsymbol{A})$. Então*

1. $\operatorname{tr} M^{\intercal} = \operatorname{tr} M$;
2. $\operatorname{tr} MN = \operatorname{tr} NM$;
3. $\operatorname{tr} M + N = \operatorname{tr} M + \operatorname{tr} N$;
4. $\operatorname{tr} aM = a \operatorname{tr} M$.

□ *Demonstração.* 1.

$$\operatorname{tr} M = \mathop{\mathbin{+}}_{i=1}^{d} m_{i,i} = \operatorname{tr} M^{\intercal}.$$

2.

$$\begin{aligned}
\operatorname{tr} MN &= \mathop{\mathbin{+}}_{i=1}^{d} \left(\mathop{\mathbin{+}}_{k=1}^{d} m_{i,k} n_{k,i} \right) \\
&= \mathop{\mathbin{+}}_{i=1}^{d} \left(\mathop{\mathbin{+}}_{k=1}^{d} n_{k,i} m_{i,k} \right) \\
&= \mathop{\mathbin{+}}_{k=1}^{d} \left(\mathop{\mathbin{+}}_{i=1}^{d} n_{k,i} m_{i,k} \right) \\
&= \operatorname{tr} NM.
\end{aligned}$$

3.

$$\begin{aligned}
\operatorname{tr} M + N &= \mathop{\mathbin{+}}_{i=1}^{d} (m_{i,i} + n_{i,i}) \\
&= \mathop{\mathbin{+}}_{i=1}^{d} m_{i,i} + \mathop{\mathbin{+}}_{i=1}^{d} n_{i,i} \\
&= \operatorname{tr} M + \operatorname{tr} N.
\end{aligned}$$

4.

$$\operatorname{tr} aM = \mathop{\mathbin{+}}_{i=1}^{d} a m_{i,i} = a \mathop{\mathbin{+}}_{i=1}^{d} m_{i,i} = a \operatorname{tr} M.$$

■

7.2 Divisão em anéis

7.2.1 Anel de frações

Lembremos que $\boldsymbol{A}^\times = (A, \times, 1)$ é o monoide multiplicativo de um anel $\boldsymbol{A}$ e que um submonoide multiplicativo de $\boldsymbol{A}$ é um submonoide $\boldsymbol{D} \leq \boldsymbol{A}^\times$; ou seja, $D \subseteq A$ e

SM1 (Identidade) $1 \in D$;
SM2 (Fechamento) Para todos $d, d' \in D$, $dd' \in D$.

:⊢ **Definição 7.33.** Sejam $\boldsymbol{A}$ um anel e $\boldsymbol{D}$ um submonoide multiplicativo de $\boldsymbol{A}$. A *equivalência fracionária* em $A \times D$ é a relação $\rightsquigarrow$ em $A \times D$ definida por

$$(a, d) \rightsquigarrow (a', d') \Leftrightarrow \exists_{u \in D}(ad' - a'd)u = 0.$$

⊢ **Proposição 7.40.** *Sejam $\boldsymbol{A}$ um anel e $\boldsymbol{D}$ um submonoide multiplicativo de $\boldsymbol{A}$. A equivalência fracionária é uma equivalência.*

□ *Demonstração.*

1. (Reflexividade) Seja $(a, d) \in A \times D$. Como $1 \in D$, segue que $(ad - ad)1 = 0$, logo $(a, d) \rightsquigarrow (a, d)$;
2. (Simetria) Sejam $(a, d), (a', d') \in A \times D$ tais que $(a, d) \rightsquigarrow (a', d')$. Então existe $u \in D$ tal que $(ad' - a'd)u = 0$, portanto

$$(a'd - ad')u = -(ad' - a'd)u = 0,$$

   o que mostra que $(a', d') \rightsquigarrow (a, d)$;
3. (Transitividade) Sejam $(a, d), (a', d'), (a'', d'') \in A \times D$ tais que $(a, d) \rightsquigarrow (a', d')$ e $(a', d') \rightsquigarrow (a'', d'')$. Então existem u, u' tais que $(ad' - a'd)u = 0$ e $(a'd'' - a''d')u' = 0$, portanto segue da comutatividade que

$$\begin{aligned}(ad'' - a''d)(d'uu') &= ad'ud''u' - a''d'u'du \\ &= ad'ud''u' - a'dud''u' + a'd''u'du - a''d'u'du \\ &= (ad' - a'd)ud''u' + (a'd'' - a''d')u'du \\ &= 0d''u' + 0du \\ &= 0.\end{aligned}$$

   Como $d'uu' \in D$, segue que $(a, d) \rightsquigarrow (a'', d'')$.

■

:⊢ **Definição 7.34.** Sejam $\boldsymbol{A}$ um anel e $\boldsymbol{D}$ um submonoide multiplicativo de $\boldsymbol{A}$. O *anel de frações* de A sobre D é a lista

$$\boldsymbol{A} \div \boldsymbol{D} := (A \div D, +, -, 0, \times, 1),$$

em que

1. $A \div D$ é o *conjunto de frações*, definido por
$$A \div D := A \times D \big/ \dot{\sim}$$
e um elemento de $A \div D$ é denotado
$$\frac{a}{d} := [\![(a,d)]\!] = \{(a',d') \in A \times D \mid (a,d) \dot{\sim} (a',d')\},$$
e denominado *fração* de a sobre d, em que a é o *numerador* e d o *denominador* da fração.
2. $+$ é a *adição* de $A \div D$, definida por
$$\begin{aligned} +\colon (A \div D) \times (A \div D) &\longrightarrow A \div D \\ \left(\frac{a}{d}, \frac{a'}{d'}\right) &\longmapsto \frac{ad' + a'd}{dd'}. \end{aligned}$$
3. $-$ é a *inversão aditiva* de $A \div D$, definida por
$$\begin{aligned} -\colon A \div D &\longrightarrow A \div D \\ \frac{a}{d} &\longmapsto \frac{-a}{d}. \end{aligned}$$
4. 0 é a *nulidade* de $A \div D$, definida por
$$0 := \frac{0}{1}.$$
5. $\times$ é a *multiplicação* de $A \div D$, definida por
$$\begin{aligned} \times\colon (A \div D) \times (A \div D) &\longrightarrow A \div D \\ \left(\frac{a}{d}, \frac{a'}{d'}\right) &\longmapsto \frac{aa'}{dd'}. \end{aligned}$$
6. 1 é a *unidade* de $A \div D$, definida por
$$1 := \frac{1}{1}.$$

$\triangleright$ **Exercício 7.4.** *Sejam* $\boldsymbol{A}$ *um anel e* $\boldsymbol{D}$ *um submonoide multiplicativo de* $\boldsymbol{A}$*. As operações* $+, -, \times$ *em* $A \div D$ *definidas acima são bem definidas.*

$\vdash$ **Proposição 7.41.** *Sejam* $\boldsymbol{A}$ *um anel e* $\boldsymbol{D}$ *um submonoide multiplicativo de* $\boldsymbol{A}$*. O anel de frações*
$$\boldsymbol{A} \mathbin{\dot{\div}} \boldsymbol{D} = (A \div D, +, -, 0, \times, 1)$$
é um anel.

□ *Demonstração.* Demonstramos por partes.

1. $((A \div D, +, -, 0)$ é grupo comutativo)

   1.1. (Associatividade) Sejam $\frac{a}{d}, \frac{a'}{d'}, \frac{a''}{d''} \in A \div D$. Então

$$\begin{aligned}\left(\frac{a}{d}+\frac{a'}{d'}\right)+\frac{a''}{d''} &= \frac{ad'+a'd}{dd'}+\frac{a''}{d''}\\ &= \frac{(ad'+a'd)d''+a''(dd')}{(dd')d''}\\ &= \frac{a(d'd'')+(a'd''+a''d')d}{d(d'd'')}\\ &= \frac{a}{d}+\frac{a'd''+a''d'}{d'd''}\\ &= \frac{a}{d}+\left(\frac{a'}{d'}+\frac{a''}{d''}\right).\end{aligned}$$

   1.2. (Identidade) Seja $\frac{a}{d} \in A \div D$. Então

$$0+\frac{a}{d}=\frac{0}{1}+\frac{a}{d}=\frac{0d+a1}{1d}=\frac{a}{d}.$$

   1.3. (Invertibilidade) Seja $\frac{a}{d} \in A \div D$. Então

$$\frac{a}{d}+(-\frac{a}{d})=\frac{a}{d}+\frac{-a}{d}=\frac{ad-ad}{dd}=\frac{0}{dd}=0.$$

   1.4. (Comutatividade) Sejam $\frac{a}{d}, \frac{a'}{d'} \in A \div D$. Então

$$\frac{a}{d}+\frac{a'}{d'}=\frac{ad'+a'd}{dd'}=\frac{a'd+ad'}{d'd}=\frac{a'}{d'}+\frac{a}{d}.$$

2. $((A \div D, \times, 1)$ é monoide comutativo)

   2.1. (Associatividade) Sejam $\frac{a}{d}, \frac{a'}{d'}, \frac{a''}{d''} \in A \div D$. Então

$$\left(\frac{a}{d}\frac{a'}{d'}\right)\frac{a''}{d''}=\frac{aa'}{dd'}\frac{a''}{d''}=\frac{(aa')a''}{(dd')d''}=\frac{a(a'a'')}{d(d'd'')}=\frac{a}{d}\frac{a'a''}{d'd''}=\frac{a}{d}\left(\frac{a'}{d'}\frac{a''}{d''}\right).$$

   2.2. (Identidade) Seja $\frac{a}{d} \in A \div D$. Então

$$1\frac{a}{d}=\frac{1}{1}\frac{a}{d}=\frac{1a}{1d}=\frac{a}{d}.$$

   2.3. (Comutatividade) Sejam $\frac{a}{d}, \frac{a'}{d'} \in A \div D$. Então

$$\frac{a}{d}\frac{a'}{d'}=\frac{aa'}{dd'}=\frac{a'a}{d'd}=\frac{a'}{d'}\frac{a}{d}.$$

3. (Distributividade de $\times$ sobre $+$) Sejam $\frac{a}{d}, \frac{a'}{d'}, \frac{a''}{d''} \in A \div D$. Então

$$\begin{aligned}
\frac{a}{d}\left(\frac{a'}{d'} + \frac{a''}{d''}\right) &= \frac{a}{d}\frac{a'd'' + a''d'}{d'd''} \\
&= \frac{a(a'd'' + a''d')}{d(d'd'')} \\
&= \frac{a(a'd'') + a(a''d')}{d(d'd'')} \\
&= \frac{(aa')(dd'') + (aa'')(dd')}{(dd')(dd'')} \\
&= \frac{aa'}{dd'} + \frac{aa''}{dd''} \\
&= \frac{a}{d}\frac{a'}{d'} + \frac{a}{d}\frac{a''}{d''}.
\end{aligned}$$

■

⊢ **Proposição 7.42.** *Seja* $\boldsymbol{A \div D}$ *um anel de frações.*

1. $\boldsymbol{A \div D}$ *é trivial se, e somente se,* $0 \in D$*;*
2. *Para todo* $d \in D$*,*
$$\frac{d}{d} = 1;$$
3. *Para todo* $d \in D$*,*
$$\frac{1}{d} = \left(\frac{d}{1}\right)^{-1}.$$

□ *Demonstração.*

1. A trivialidade do anel de frações é equivalente a $\frac{0}{1} = \frac{1}{1}$, que ocorre se, e somente se, existe $u \in D$ tal que $(01 - 11)u = 0$. Como $(01 - 11)u = -u$, isso é equivalente a $0 \in D$.
2. Basta notar que $(d1 - 1d)1 = 0$.
3. Basta notar que
$$\frac{1}{d}\frac{d}{1} = \frac{d}{d} = 1.$$

■

:⊢ **Definição 7.35.** Seja $\boldsymbol{A \div D}$ um anel de frações. O *homomorfismo quociente* é a função

$$\begin{aligned}
\frac{\mathrm{I}}{1} \colon A &\longrightarrow A \div D \\
a &\longmapsto \frac{a}{1}.
\end{aligned}$$

⊢ **Proposição 7.43.** *Seja* $\boldsymbol{A \div D}$ *um anel de frações. O homomorfismo quociente é um homomorfismo de anel.*

□ *Demonstração.* Demonstramos por partes.

1. (Homomorfismo de grupo) Sejam $a, a' \in A$. Então

$$\frac{\mathrm{I}}{1}(a + a') = \frac{a + a'}{1} = \frac{a1 + a'1}{1 \times 1} = \frac{a}{1} + \frac{a'}{1} = \frac{\mathrm{I}}{1}(a) + \frac{\mathrm{I}}{1}(a').$$

2. (Homomorfismo de monoide)

   2.1. Sejam $a, a' \in A$. Então

$$\frac{\mathrm{I}}{1}(aa') = \frac{aa'}{1} = \frac{a}{1}\frac{a'}{1} = \frac{\mathrm{I}}{1}(a)\frac{\mathrm{I}}{1}(a').$$

   2.2. Por definição da unidade do anel de frações,

$$\frac{\mathrm{I}}{1}(1) = \frac{1}{1} = 1.$$

■

⊢ **Proposição 7.44** (Propriedade universal)**.** *Sejam* $\boldsymbol{A \div D}$ *um anel de frações,* $\boldsymbol{A'}$ *um anel e* $h\colon A \longrightarrow A'$ *um homomorfismo de anel tal que, para todo* $d \in D$*,* $h(d) \in A'$ *é invertível. Existe único homomorfismo de anel* $h'\colon A \div D \longrightarrow A'$ *tal que* $h = h' \circ \frac{\mathrm{I}}{1}$ *(o diagrama comuta).*

$$\begin{tikzcd}
\boldsymbol{A} \arrow[r, "h"] \arrow[d, "\frac{\mathrm{I}}{1}"'] & \boldsymbol{A'} \\
\boldsymbol{A \div D} \arrow[ur, dashed, "h'"'] &
\end{tikzcd}$$

□ *Demonstração.* Definimos

$$\begin{aligned} h'\colon A \div D &\longrightarrow A' \\ \frac{a}{d} &\longmapsto h(a)h(d)^{-1}. \end{aligned}$$

Mostremos que essa função está bem definida. Sejam $a, a' \in A$ e $d, d' \in D$ tais que $\frac{a}{d} = \frac{a'}{d'}$. Isso significa que existe $u \in D$ tal que $(ad' - a'd)u = 0$. Isso significa que

$$(h(a)h(d') - h(a')h(d))h(u) = h(ad' - a'd)h(u) = h((ad' - a'd)u) = 0$$

Como $h(u)$ é invertível, segue que $h(a)h(d') - h(a')h(d)$, ou seja, que $h(a)h(d') = h(a')h(d)$. Como $h(d)$ e $h(d')$ são invertíveis, segue que

$$h'\left(\frac{a}{d}\right) = h(a)h(d)^{-1} = h(a')h(d')^{-1} = h'\left(\frac{a'}{d'}\right).$$

Agora, mostremos que h' é homomorfismo de anel.

1. (Homomorfismo de grupo) Sejam $\frac{a}{d}, \frac{a'}{d'} \in A \div D$. Então

$$\begin{aligned} h'\left(\frac{a}{d} + \frac{a'}{d'}\right) &= h'\left(\frac{ad' + a'd}{dd'}\right) \\ &= h(ad' + a'd)h(dd')^{-1} \\ &= (h(a)h(d') + h(a')h(d))h(d)^{-1}h(d')^{-1} \\ &= h(a)h(d)^{-1} + h(a')h(d')^{-1} \\ &= h'\left(\frac{a}{d}\right) + h\left(\frac{a'}{d'}\right). \end{aligned}$$

2. (Homomorfismo de monoide)

   2.1. Sejam $\frac{a}{d}, \frac{a'}{d'} \in A \div D$. Então

$$\begin{aligned} h'\left(\frac{a}{d}\frac{a'}{d'}\right) &= h'\left(\frac{aa'}{dd'}\right) \\ &= h(aa')h(dd')^{-1} \\ &= h(a)h(a')h(d)^{-1}h(d')^{-1} \\ &= h(a)h(d)^{-1}h(a')h(d')^{-1} \\ &= h'\left(\frac{a}{d}\right)h\left(\frac{a'}{d'}\right). \end{aligned}$$

   2.2. Segue de

$$h'\left(\frac{1}{1}\right) = h(1)h(1)^{-1} = 1.$$

Por fim, para mostrar a unicidade, suponhamos que $h''\colon A \div D \longrightarrow A'$ é um

homomorfismo de anel tal que $h = h'' \circ \frac{\mathrm{I}}{1}$. Para todo $\frac{a}{d} \in A \div D$,

$$\begin{aligned} h''\left(\frac{a}{d}\right) &= h''\left(\frac{a}{1}\frac{1}{d}\right) \\ &= h''\left(\frac{a}{1}\right) h''\left(\left(\frac{d}{1}\right)^{-1}\right) \\ &= h''\left(\frac{a}{1}\right) h''\left(\frac{d}{1}\right)^{-1} \\ &= \left(h'' \circ \frac{\mathrm{I}}{1}\right)(a) \left(h'' \circ \frac{\mathrm{I}}{1}\right)(d)^{-1} \\ &= h(a)h(d)^{-1} \\ &= h'\left(\frac{a}{d}\right), \end{aligned}$$

o que mostra que $h'' = h'$. ■

▷ **Exercício 7.5.** *Sejam $\boldsymbol{A}$ e $\boldsymbol{A}'$ anéis, $\boldsymbol{D}$ um submonoide multiplicativo de $\boldsymbol{A}$ e $q\colon A \longrightarrow A'$ um homomorfismo de anel tal que*

1. *Para todo $d \in D$, $q(d) \in A'$ é invertível;*
2. *Para todo $a \in A$ tal que $q(a) = 0$, existe $u \in D$ tal que $au = 0$;*
3. *Para todo $a' \in A'$, existem $a \in A$ e $d \in D$ tais que $a' = q(a)q(d)^{-1}$.*

Existe único isomorfismo de anel $h\colon A \div D \longrightarrow A'$ tal que $q = h \circ \frac{\mathrm{I}}{1}$.

7.2.2 Corpo de frações e números racionais $\mathbb{Q}$

▷ **Exercício 7.6.** *Seja $\boldsymbol{A}$ um anel. Então $A \smallsetminus \{0\}$ é um submonoide multiplicativo de $\boldsymbol{A}$ se, e somente se, $\boldsymbol{A}$ é um domínio.*

:⊢ **Definição 7.36.** Seja $\boldsymbol{A}$ um domínio. O *corpo de frações* de $\boldsymbol{A}$ é o anel de frações

$$\boldsymbol{A}^{\div} := \boldsymbol{A} \div (\boldsymbol{A} \smallsetminus \{\boldsymbol{0}\}).$$

⊢ **Proposição 7.45.** *Seja $\boldsymbol{A}$ um domínio. O corpo de frações $\boldsymbol{A}^{\div}$ de $\boldsymbol{A}$ é um corpo.*

□ *Demonstração.* Como $\boldsymbol{A}^{\div} = \boldsymbol{A} \div (\boldsymbol{A} \smallsetminus \{\boldsymbol{0}\})$ é um anel, basta mostrar que todo elemento não zero tem inversa multiplicativa. Seja $\frac{a}{d} \in A^{\div} \smallsetminus \{0\}$. Como $\frac{a}{d} \neq 0$, então $a \neq 0$, portanto $a \in A \smallsetminus \{0\}$. Segue que

$$\frac{a}{d}\frac{d}{a} = \frac{1}{1} = 1,$$

portanto $\frac{a}{d}$ é invertível. ■

7.2.2.1 Os números racionais $\mathbb{Q}$

A construção de corpo de frações de um domínio aplicada ao anel dos números inteiros $\mathbb{Z}$ é o conjunto dos números racionais.

:⊢ **Definição 7.37.** O *corpo dos números racionais* é o corpo

$$\mathbb{Q} := \mathbb{Z}^{\div} = \mathbb{Z} \div (\mathbb{Z} \smallsetminus \{0\}).$$

7.2.2.2 Funções de arredondamento

Chão e teto

:⊢ **Definição 7.38.** A função *chão* é a função

$$\begin{aligned} \lfloor \cdot \rfloor \colon \mathbb{R} &\longrightarrow \mathbb{Z} \\ x &\longmapsto \max \{ i \in \mathbb{Z} \mid i \leq x \} . \end{aligned}$$

A função *teto* é a função

$$\begin{aligned} \lceil \cdot \rceil \colon \mathbb{R} &\longrightarrow \mathbb{Z} \\ x &\longmapsto \min \{ i \in \mathbb{Z} \mid i \geq x \} . \end{aligned}$$

▷ **Exercício 7.7.** *Para todo $x \in \mathbb{R}$,*

1. *(Conjugação)*

$$\begin{aligned} -\lfloor x \rfloor &= \lceil -x \rceil \\ -\lceil x \rceil &= \lfloor -x \rfloor ; \end{aligned}$$

2. *(Idempotência)*

$$\begin{aligned} \lfloor \lfloor x \rfloor \rfloor &= \lfloor x \rfloor \\ \lceil \lceil x \rceil \rceil &= \lceil x \rceil ; \end{aligned}$$

▷ **Exercício 7.8.** *Para todo $x, x' \in \mathbb{R}$ e todo $i \in \mathbb{Z}$,*

1.

$$\begin{aligned} \lfloor x + i \rfloor &= \lfloor x \rfloor + i \\ \lceil x + i \rceil &= \lceil x \rceil + i; \end{aligned}$$

2.

$$\begin{aligned} \lfloor x \rfloor + \lfloor x' \rfloor \leq \lfloor x + x' \rfloor &\leq \lfloor x \rfloor + \lfloor x' \rfloor + 1 \\ -1 + \lceil x \rceil + \lceil x' \rceil \leq \lceil x + x' \rceil &\leq \lceil x \rceil + \lceil x' \rceil ; \end{aligned}$$

▷ **Exercício 7.9.** *Para todos $x, x' \in \mathbb{R}$, $x' > 0$ e todos $i, i' \in \mathbb{Z}$, $i > 0$,*

1.
$$\left\lfloor \frac{\lfloor x \rfloor + i'}{i} \right\rfloor = \left\lfloor \frac{x + i'}{i} \right\rfloor$$
$$\left\lceil \frac{\lceil x \rceil + i'}{i} \right\rceil = \left\lceil \frac{x + i'}{i} \right\rceil;$$

2.
$$\left\lceil \frac{i'}{i} \right\rceil - 1 = \left\lfloor \frac{i' - 1}{i} \right\rfloor$$
$$\left\lfloor \frac{i'}{i} \right\rfloor + 1 = \left\lceil \frac{i' + 1}{i} \right\rceil;$$

3.
$$\left\lfloor \frac{\lfloor x \mathbin{/} x' \rfloor}{i} \right\rfloor = \left\lfloor \frac{x}{x'i} \right\rfloor$$
$$\left\lceil \frac{\left\lceil \frac{x}{x'} \right\rceil}{i} \right\rceil = \left\lceil \frac{x}{x'i} \right\rceil.$$

Resto e parte fracionária

⊢ **Definição 7.39.** Sejam $x, d \in \mathbb{R}$, $d \neq 0$. O *resto* de x com respeito a d é
$$[\![x]\!]_d := x - d \left\lfloor \frac{x}{d} \right\rfloor.$$
A *parte fracionária* de x é o resto de x com respeito a 1:
$$[\![x]\!] := [\![x]\!]_1.$$

▷ **Exercício 7.10.** *Sejam $x, d \in \mathbb{R}$, $d \neq 0$.*

1. *Se $d > 0$,*
$$0 \leq [\![x]\!]_d < d;$$
2. *Se $d < 0$,*
$$d < [\![x]\!]_d \leq 0;$$

▷ **Exercício 7.11.** *Seja $x \in \mathbb{R}$,*
$$\lceil [\![x]\!] \rceil = \lceil x \rceil - \lfloor x \rfloor = \begin{cases} 0, & x \in \mathbb{Z} \\ 1, & x \notin \mathbb{Z}. \end{cases}$$

▷ **Exercício 7.12.** *Sejam $x \in \mathbb{R}$, $i \in \mathbb{Z}$, $i > 0$.*
$$\left\lceil \left[\!\!\left[\frac{x}{i} \right]\!\!\right] \right\rceil = \left\lceil \frac{x}{i} \right\rceil - \left\lfloor \frac{x}{i} \right\rfloor = \begin{cases} 0, & x \in i\mathbb{Z} \\ 1, & x \notin i\mathbb{Z}. \end{cases}$$

Arredondamento

:⊢ **Definição 7.40.** A função *arredondamento meio para baixo* é a função

$$\begin{aligned} \lceil\cdot\rfloor : \mathbb{R} &\longrightarrow \mathbb{Z} \\ x &\longmapsto \left\lceil x - \frac{1}{2} \right\rceil. \end{aligned}$$

A função *arredondamento meio para cima* é a função

$$\begin{aligned} \lfloor\cdot\rceil : \mathbb{R} &\longrightarrow \mathbb{Z} \\ x &\longmapsto \left\lfloor x + \frac{1}{2} \right\rfloor. \end{aligned}$$

7.2.3 Números reais $\mathbb{R}$

Definiremos o conjunto dos números reais, suas operações de corpo e sua relação de ordem total pela construção usando cortes de números racionais.

:⊢ **Definição 7.41.** Um *corte* de $\mathbb{Q}$ é um conjunto $c \subseteq \mathbb{Q}$ tal que

1. (Finitude) $c \neq \emptyset$ e $c \neq \mathbb{Q}$;
2. (Fechamento inferior) Para todos $q, q' \in \mathbb{Q}$ tais que $q < q'$, se $q' \in c$ então $q \in c$;
3. (Não-maximalidade) Para todo $q \in c$, algum $q' \in c$ satisfaz $q < q'$.

O *conjunto dos números reais* é o conjunto dos cortes de $\mathbb{Q}$, denotado $\mathbb{R}$.

:⊢ **Definição 7.42.** A *ordem dos números reais* é a ordem $\subseteq$ de inclusão de subconjuntos de $\mathbb{Q}$, denotada $\leq$.

⊢ **Proposição 7.46.** *O par $(\mathbb{R}, \leq)$ é um conjunto totalmente ordenado, e completo no sentido que todo conjunto não-vazio limitado superiormente tem supremo.*

□ *Demonstração.* O par $(\mathbb{R}, \leq)$ é um conjunto parcialmente ordenado porque é um subconjunto parcialmente ordenado de $(\wp(\mathbb{Q}), \subseteq)$. Basta mostrarmos que a ordem é total. Sejam $c, c' \in \mathbb{R}$ e suponhamos que $c' \not\leq c$, ou seja, $c' \not\subseteq c$. Mostremos que $c \leq c'$, ou seja, $c \subseteq c'$. Seja $q \in c$. Como $c' \not\subseteq c$, algum $q' \in c'$ satisfaz $q' \notin c$. Se valesse $q > q'$, então, como $q \in c$, seguiria do fechamento inferior que $q' \in c$, uma contradição, portanto $q \leq q'$; se valesse $q = q'$, então, como $q \in c$, seguiria que $q' \in c$, uma contradição, portanto $q < q'$. Segue assim do fechamento inferior que, como $q' \in c'$, $q \in c'$, portanto $c \subseteq c'$, ou seja, $c \leq c'$.

Agora, seja $C \subseteq \mathbb{R}$ um conjunto não-vazio limitado superiormente. Mostremos que o supremo de C é dado por

$$\sup C := \bigcup C = \bigcup_{c \in C} c \subseteq \mathbb{Q}.$$

Primeiro mostremos que $\sup C \in \mathbb{R}$.

- (Finitude) Como C é não-vazio, algum $c \in C$. Como $c \in \mathbb{R}$, $c \neq \emptyset$, portanto algum $q \in c$. Assim segue que $q \in c \subseteq \bigcup C = \sup C$, portanto $\sup C \neq \emptyset$. Como C é limitado superiormente, algum $l \in \mathbb{R}$ satisfaz que, para todo $c \in C$, $c \leq l$, ou seja, $c \subseteq l$. Como $l \in \mathbb{R}$, $l \neq \mathbb{Q}$, portanto algum $q \in \mathbb{Q} \smallsetminus l$. Como, para todo $c \in C$, $c \subseteq l$, segue que $q \notin c$, logo $q \notin \bigcup C$, o qu implica que $\sup C \neq \mathbb{Q}$.
- (Fechamento inferior) Sejam $q, q' \in \mathbb{Q}$ tais que $q < q'$ e $q' \in \sup C = \bigcup C$. Então, para algum $c \in \sup C$, $q' \in c$. Pelo fechamento inferior, segue que $q \in c$, portanto $q \in \bigcup C = \sup C$.
- (Não-maximalidade) Seja $q \in \sup C = \bigcup C$. Então, para algum $c \in C$, $q \in c$. Pela não-maximalidade, segue que algum $q' \in c$ satisfaz $q < q'$. Mas então $q' \in c \subseteq \bigcup C = \sup C$.

Agora mostremos que $\sup C$ é supremo de C. Primeiro mostremos que $\sup C$ é limitante superior de C. Como, para todo $c \in C$, $c \subseteq \bigcup C$, ou seja, $c \leq \sup C$. Agora, mostremos que $\sup C$ é o menor limitante superior. Seja $l \in \mathbb{R}$ um limitante superior de C, ou seja, para todo $c \in C$, $c \leq l$. Mas isso é equivalente a $c \subseteq l$, portanto $\sup C = \bigcup C \subseteq l$, ou seja, $\sup C \leq l$. ■

:⊢ **Definição 7.43.** O *grupo aditivo dos números reais* é a lista

$$\mathbb{R}^+ := (\mathbb{R}, +, -, 0),$$

em que

1. $+$ é a *adição* de $\mathbb{R}$, definida por

$$\begin{aligned} +\colon \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (c, c') &\longmapsto c + c' = \{q + q' \mid q \in c \wedge q' \in c'\}\,; \end{aligned}$$

2. 0 é o *zero* de $\mathbb{R}$, definido por

$$0_{\mathbb{R}} := \{q \in \mathbb{Q} \mid q < 0\}\,;$$

3. $-$ é a *inversa aditiva* de $\mathbb{R}$, definida por

$$\begin{aligned} -\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ c &\longmapsto -c := \left\{-q \;\middle|\; q \in c^{\complement} \wedge \exists_{r \in c^{\complement}} r < q\right\}. \end{aligned}$$

A menos que gere ambiguidade, denotaremos $0_{\mathbb{R}}$ por 0.

▷ **Exercício 7.13.** *As operações $+$ e $-$ estão bem definidas, $0 \in \mathbb{R}$, o grupo aditivo dos números reais $\mathbb{R}^+$ é um grupo comutativo e $(\mathbb{R}^+, \leq)$ é um grupo ordenado.*

:⊢ **Definição 7.44.** O *monoide multiplicativo dos números reais* é a lista

$$\mathbb{R}^{\times} := (\mathbb{R}, \times, 1)$$

e o *grupo multiplicativo dos números reais* é a lista

$$\mathbb{R}^{\diagup} := (\mathbb{R} \smallsetminus \{0\}, \times, \diagup, 1),$$

em que

1. $\times \colon \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ é a *multiplicação* de $\mathbb{R}$, definida para todos $c, c' \in \mathbb{R}$ por

$$c \times c' := \begin{cases} \left\{ qq' \;\middle|\; q \in c_{\geq 0} \wedge q' \in c'_{\geq 0} \right\} \cup 0_{\mathbb{R}}, & c \geq 0_{\mathbb{R}} \wedge c' \geq 0_{\mathbb{R}}, \\ c \times (-c'), & c \geq 0_{\mathbb{R}} \wedge c' < 0_{\mathbb{R}}, \\ (-c) \times c', & c < 0_{\mathbb{R}} \wedge c' \geq 0_{\mathbb{R}}, \\ (-c) \times (-c'), & c < 0_{\mathbb{R}} \wedge c' < 0_{\mathbb{R}}; \end{cases}$$

2. 1 é o *um* de $\mathbb{R}$, definido por

$$1_{\mathbb{R}} := \{q \in \mathbb{Q} \mid q < 1\};$$

3. $\diagup$ é a *inversa multiplicativa* de $\mathbb{R}$, definida por

$$\begin{aligned} \diagup \colon \mathbb{R} \smallsetminus \{0\} &\longrightarrow \mathbb{R} \smallsetminus \{0\} \\ c &\longmapsto c^{-1} := \begin{cases} \left\{ u \diagup q \;\middle|\; u \in 1_{\mathbb{R}} \wedge q \in c^{\complement} \right\} & c > 0, \\ -((-c)^{-1}) & c < 0. \end{cases} \end{aligned}$$

A menos que gere ambiguidade, denotaremos $1_{\mathbb{R}}$ por 1.

▷ **Exercício 7.14.** *As operações $\times$ e $\diagup$ estão bem definidas, $1 \in \mathbb{R}$, o monoide multiplicativo dos números reais $\mathbb{R}^{\times}$ é um monoide comutativo e o grupo multiplicativo dos números reais $\mathbb{R}^{\diagup}$ é um grupo comutativo.*

:⊢ **Definição 7.45.** O *corpo dos números reais* é a lista

$$\mathbb{R} = (\mathbb{R}, +, -, 0, \times, \diagup, 0).$$

Por abuso de notação, denotamos tanto o corpo como o conjunto dos números reais pelo símbolo $\mathbb{R}$.

▷ **Exercício 7.15.** *O corpo dos números reais $\mathbb{R}$ é um corpo e $(\mathbb{R}, \leq)$ é um corpo ordenado.*

▷ **Exercício 7.16.** *Para todo* $q \in \mathbb{Q}$,

$$q_{\mathbb{R}} := \{q' \in \mathbb{Q} \mid q' < q\} \in \mathbb{R}$$

e a função

$$\begin{aligned} \iota \colon \mathbb{Q} &\longrightarrow \mathbb{R} \\ q &\longmapsto q_{\mathbb{R}} = \{q' \in \mathbb{Q} \mid q' < q\} \end{aligned}$$

é um homomorfismo injetivo de corpo ordenado de $(\mathbb{Q}, \leq)$ *para* $(\mathbb{R}, \leq)$.

Por essa identificação dos racionais $\mathbb{Q}$ dentro dos números reais $\mathbb{R}$, vamos sempre nos referir aos racionais como elementos de $\mathbb{R}$. A menos que gere ambiguidade, denotaremos $q_{\mathbb{R}}$ por q. A partir de agora, não vamos mais utilizar a construção de $\mathbb{R}$ como um conjunto de subconjuntos de $\mathbb{R}$, somente o fato de que ele é um corpo ordenado e completo.

7.2.3.1 Raiz de números reais

⊢ **Proposição 7.47.** *Sejam* $n \in \mathbb{Z}_>$ *e* $c \in \mathbb{R}_\geq$. *O conjunto*

$$\{x \in \mathbb{R}_\geq \mid x^n \leq c\}$$

é não vazio e limitado superiormente por $1 \vee c$.

□ *Demonstração.* O conjunto $\{x \in \mathbb{R}_\geq \mid x^n \leq c\}$ é não vazio pois, como $c \geq 0$ e $0^n = 0$, então $c \geq 0^n$, portanto $0 \in \{x \in \mathbb{R}_\geq \mid x^n \leq c\}$.

O conjunto $\{x \in \mathbb{R}_\geq \mid x^n \leq c\}$ é limitado superiormente por $1 \vee c$, seja $x \in \mathbb{R}_\geq$ tal que $x^n \leq c$. Se $x < 1$, então $x < 1 \vee c$. Se $x \geq 1$, como $\mathbb{R}$ é corpo ordenado segue que $x \leq x^n$, logo $x \leq c \leq 1 \vee c$. Isso mostra que $1 \vee c$ é um limitante superior de $\{x \in \mathbb{R}_\geq \mid x^n \leq c\}$. ■

Isso mostra que o conjunto $\{x \in \mathbb{R}_\geq \mid x^n \leq c\}$ tem supremo. Esse supremo define a função raiz n-ésima.

:⊢ **Definição 7.46.** Seja $n \in \mathbb{Z}_>$. A *n-ésima raiz* em $\mathbb{R}_\geq$ é a função

$$\begin{aligned} \sqrt[n]{} \colon \mathbb{R}_\geq &\longrightarrow \mathbb{R}_\geq \\ c &\longmapsto \sqrt[n]{c} := \sup\{x \in \mathbb{R}_\geq \mid x^n \leq c\} \end{aligned}$$

⊢ **Proposição 7.48.** *Sejam* $n \in \mathbb{Z}_>$ *e* $c \in \mathbb{R}_\geq$.

1. *Se* $c \geq 1$, *então* $1 \leq \sqrt[n]{c} \leq c$;
2. *Se* $c \leq 1$, *então* $c \leq \sqrt[n]{c} \leq 1$;

3. *$\sqrt[n]{c} = c$ se, e somente se, $c = 1$ ou $c = 0$;*
4. *$(\sqrt[n]{c})^n = c$.*

□ *Demonstração.* 1. Nesse caso, $1 \vee c = c$, portanto c é limitante superior de $\{x \in \mathbb{R}_{\geq} \mid x^n \leq c\}$; como $\sqrt[n]{c}$ é supremo dese conjunto, é o menor limitante superior, e segue que $\sqrt[n]{c} \leq c$.
Como $1 \leq c$ e $1^n = 1$, então $1^n \leq c$, e segue que $1 \in \{x \in \mathbb{R}_{\geq} \mid x^n \leq c\}$; como $\sqrt[n]{c}$ é supremo dese conjunto, é limitante superior, e segue que $1 \leq \sqrt[n]{c}$.
2. Nesse caso, $1 \vee c = 1$, portanto 1 é limitante superior de $\{x \in \mathbb{R}_{\geq} \mid x^n \leq c\}$; como $\sqrt[n]{c}$ é supremo dese conjunto, é o menor limitante superior, e segue que $\sqrt[n]{c} \leq 1$.
Como $c \leq 1$, então $c^n \leq c$, e segue que $c \in \{x \in \mathbb{R}_{\geq} \mid x^n \leq c\}$; como $\sqrt[n]{c}$ é supremo dese conjunto, é limitante superior, e segue que $c \leq \sqrt[n]{c}$.
3. Segue dos itens anteriores que $\sqrt[n]{1} \leq 1$ e $\sqrt[n]{1} \geq 1$, portanto $\sqrt[n]{1} = 1$. Além disso, $\{x \in \mathbb{R}_{\geq} \mid x^n \leq 0\} = \{0\}$, portanto $\sqrt[n]{0} = 0$. A recíproca é exercício.
4. Exercício.

■

7.2.4 Divisão e associação

7.2.4.1 Divisão

:⊢ **Definição 7.47.** Sejam $\boldsymbol{A}$ um anel e $a, a' \in A$. O elemento a *divide* o elemento a' em $\boldsymbol{A}$ se, e somente se, existe $q \in A$ tal que $aq = a'$. O elemento a é um *divisor* de a' em $\boldsymbol{A}$, o elemento a' é um *múltiplo* de a em $\boldsymbol{A}$ e denota-se $a \preceq_{\boldsymbol{A}} a'$. A relação $\preceq_{\boldsymbol{A}}$ é a relação de *divisão* em $\boldsymbol{A}$. Sempre que possível, o subíndice de $\preceq_{\boldsymbol{A}}$ será omitido.

⊢ **Proposição 7.49.** *Seja $\boldsymbol{A}$ um anel. A relação de divisão $\preceq$ em A é uma pré-ordem.*

□ *Demonstração.* (Reflexividade) Seja $a \in A$. Então $a \preceq a$, pois $a \cdot 1 = a$. (Transitividade) Sejam $a, a', a'' \in A$ tais que $a \preceq a'$ e $a' \preceq a''$. Então existem $q, q' \in A$ tais que $aq = a'$ e $a'q' = a''$. Logo $aqq' = a''$. Como $qq' \in A$, segue que $a \preceq a''$. ■

⊢ **Proposição 7.50.** *Seja $\boldsymbol{A}$ um anel.*

1. *Para todos $a \in A$ e $u \in A^{\prime}$,*

$$u \preceq a \preceq 0;$$

2. *Para todos $a \in A$ e $u \in A^{\prime}$,*

$$a \preceq u \Leftrightarrow a \in A^{\prime};$$

3. *Para todo $a \in A$,*
$$0 \preceq a \Leftrightarrow a = 0.$$
4. *Para todos $a, a', a'' \in A$ tais que $a' \preceq a''$,*
$$aa' \preceq aa''.$$
5. *Para todos $u \in A^{\prime}$ e $a, a' \in A$ tais que $a \preceq a'$,*
$$ua \preceq a \quad e \quad a \preceq ua'.$$

□ *Demonstração.* Sejam $a \in A$ e $u \in A^{\prime}$.

1. Como $u \in A^{\prime}$, existe $u^{-1} \in A^{\prime}$. Então $u \cdot (u^{-1}a) = a$ e segue que $u \preceq a$; como $a \cdot 0 = 0$, segue que $a \preceq 0$;
2. Se $a \preceq u$, existe $q \in A$ tal que $aq = u$. Como $u \in A^{\prime}$, segue que $aqu^{-1} = uu^{-1} = 1$. Logo $a \in A^{\prime}$. Reciprocamente, se $a \in A^{\prime}$, existe a^{-1} tal que $aa^{-1} = 1$. Então $aa^{-1}u = u$ Logo $a \preceq u$;
3. Se $0 \preceq a$, existe $q \in A$ tal que $0q = a$. Mas então $a = 0$. A recíproca segue da reflexividade de $\preceq$.
4. Se $a' \preceq a''$, existe $q \in D$ tal que $a'q = a''$, logo $aa'q = aa''$, portanto $aa' \preceq aa''$.
5. Usaremos a comutatividade. Se $a \preceq a'$, existe $q \in A$ tal que $aq = a'$, portanto
$$a' = uu^{-1}aq = uau^{-1}q,$$
logo $ua \preceq a'$, e
$$ua' = uaq = auq,$$
logo $a \preceq ua'$. ■

⊢ **Proposição 7.51.** *Sejam $\boldsymbol{A}$ um anel e $a, a' \in A$. Então $aA \subseteq a'A$ se, e somente se, $a' \preceq a$.*

□ *Demonstração.* Se $aA \subseteq a'A$, então $a \in a'A$. Mas isso significa que existe $q \in A$ tal que $a = a'q$, o que mostra que $a' \preceq a$. A implicação contrária segue a mesma demonstração com as implicação invertidas. ■

:⊢ **Definição 7.48.** Sejam $\boldsymbol{A}$ um anel e $Q \subseteq A$. Um *divisor comum* de Q em $\boldsymbol{A}$ é um elemento $d \in A$ que satisfaz, para todo $q \in Q$,
$$d \preceq q.$$

O conjunto dos divisores comuns de Q em $\boldsymbol{A}$ é $\mathrm{Div}_{\boldsymbol{A}}(Q)$. O subíndice $\boldsymbol{A}$ será omitido sempre que possível.

Dualmente, um *múltiplo comum* de Q em $\boldsymbol{A}$ é um elemento $m \in A$ que satisfaz, para todo $q \in Q$,

$$q \preceq m.$$

O conjunto dos múltiplos comuns de Q em $\boldsymbol{A}$ é $\mathrm{Mul}_{\boldsymbol{A}}(Q)$. O subíndice $\boldsymbol{A}$ será omitido sempre que possível.

⊢ **Proposição 7.52.** *Sejam $\boldsymbol{A}$ um anel e $Q \subseteq A$. Então*

1. $\mathrm{Div}(A) = \mathrm{Div}(A^{\prime}) = A^{\prime}$ *e* $\mathrm{Div}(\emptyset) = \mathrm{Div}(\{0\}) = A$*;*
2. $A^{\prime} \subseteq \mathrm{Div}(Q) \subseteq A$*;*
3. $\{0\} \cap \mathrm{Div}(Q) \neq \emptyset \Leftrightarrow Q \subseteq \{0\} \Leftrightarrow \mathrm{Div}(Q) = A$.

Dualmente,

1. $\mathrm{Mul}(A) = \mathrm{Mul}(\{0\}) = \{0\}$ *e* $\mathrm{Mul}(\emptyset) = \mathrm{Mul}(A^{\prime}) = A$*;*
2. $\{0\} \subseteq \mathrm{Mul}(Q) \subseteq A$*;*
3. $A^{\prime} \cap \mathrm{Mul}(Q) \neq \emptyset \Leftrightarrow Q \subseteq A^{\prime} \Leftrightarrow \mathrm{Mul}(Q) = A$.

□ *Demonstração.*

1. Seja $u \in A^{\prime}$. Então $u \preceq a$ para todo $a \in A$. Logo $u \in \mathrm{Div}(A)$. Por outro lado, seja $d \in \mathrm{Div}(A)$. Então $d \preceq a$ para todo $a \in A$. Em particular, $d \preceq 1$. Como $1 \in A^{\prime}$, então $d \in A^{\prime}$.
   Seja $u \in A^{\prime}$. Então $u \preceq v$ para todo $v \in A^{\prime}$. Logo $u \in \mathrm{Div}(A^{\prime})$. Por outro lado, seja $d \in \mathrm{Div}(A^{\prime})$. Então $d \preceq u$ para todo $u \in A^{\prime}$. Logo $d \in A^{\prime}$.
   Seja $a \in A$. Se $a \notin \mathrm{Div}(\emptyset)$, existe $q \in \emptyset$ tal que $a \not\preceq q$, o que é absurdo. Logo $a \in \mathrm{Div}(\emptyset)$.
   Seja $a \in A$. Então $a \preceq 0$, o que implica $a \in \mathrm{Div}(\{0\})$. A inclusão contrária é óbvia pela definição do conjunto dos divisores comuns.
2. Se $Q = \emptyset$, então $\mathrm{Div}(Q) = A$. Se $Q \neq \emptyset$, sejam $q \in Q$ e $u \in A^{\prime}$. Então $u \preceq q$. Logo $u \in \mathrm{Div}(Q)$. A inclusão $\mathrm{Div}(Q) \subseteq A$ é óbvia pela definição do conjunto de divisores comuns;
3. Suponhamos que $\{0\} \cap \mathrm{Div}(Q) \neq \emptyset$. Então $0 \in \mathrm{Div}(Q)$. Se $Q = \emptyset$, então $Q \subseteq \{0\}$. Caso contrário, seja $q \in Q$. Como $0 \preceq q$, segue que $q = 0$. Em ambos os casos, $Q \subseteq \{0\}$.
   Suponhamos que $Q \subseteq \{0\}$. Então $Q = \emptyset$ ou $Q = \{0\}$. Pelo item 1, segue que $\mathrm{Div}(Q) = A$.
   Suponhamos que $\mathrm{Div}(Q) = A$. Então $\{0\} \cap \mathrm{Div}(Q) = \{0\} \neq \emptyset$.

Dualmente,

1. Note que $a \preceq 0$ para todo $a \in A$. Logo $0 \in \mathrm{Mul}(A)$. Por outro lado, seja $m \in \mathrm{Mul}(A)$. Então $a \preceq m$ para todo $a \in A$. Em particular, $0 \preceq m$, o que implica $m = 0$.
   Note que $0 \preceq 0$, o que implica $0 \in \mathrm{Mul}(\{0\})$. Por outro lado, seja $m \in \mathrm{Mul}(\{0\})$. Então $0 \preceq m$, o que implica $m = 0$.

Seja $a \in A$. Se $a \notin \mathrm{Mul}(\emptyset)$, existe $q \in \emptyset$ tal que $q \not\preceq a$, o que é absurdo. Logo $a \in \mathrm{Mul}(\emptyset)$.
Sejam $a \in A$ e $u \in A^{\prime}$. Então $u \preceq a$, o que implica $a \in \mathrm{Mul}_{bmA}(A^{\prime})$. A inclusão contrária é óbvia pela definição do conjunto dos múltiplos comuns.

2. Se $Q = \emptyset$, então $\mathrm{Mul}(Q) = A$. Se $Q \neq \emptyset$, seja $q \in Q$. Então $q \preceq 0$, o que implica $\{0\} \in \mathrm{Mul}(Q)$. A inclusão $\mathrm{Mul}(Q) \subseteq A$ é óbvia pela definição do conjunto dos múltiplos comuns;
3. Suponhamos que $A^{\prime} \cap \mathrm{Mul}(Q) \neq \emptyset$. Seja $m \in A^{\prime} \cap \mathrm{Mul}(Q)$. Se $Q = \emptyset$, então $Q \subseteq A^{\prime}$. Caso contrário, seja $q \in Q$. Como $q \preceq m$, pois $m \in \mathrm{Mul}(Q)$, então $q \in A^{\prime}$, pois $m \in A^{\prime}$. Logo $Q \subseteq A^{\prime}$.
Suponhamos que $Q \subseteq A^{\prime}$. Se $Q = \emptyset$, do item 1 segue que $\mathrm{Mul}(Q) = A$. Caso contrário, sejam $q \in Q$ e $a \in A$. Então $q \in A^{\prime}$ e segue que $q \preceq a$. Logo $a \in \mathrm{Mul}(Q)$. Por outro lado, a inclusão $\mathrm{Mul}(Q) \subseteq A$ é óbvia pela definição do conjunto de múltiplos comuns;
Suponhamos que $\mathrm{Mul}(Q) = A$. Então $A^{\prime} \cap \mathrm{Mul}(Q) = A^{\prime}$. Como $1 \in A^{\prime}$, segue que $A^{\prime} \cap \mathrm{Mul}(Q) \neq \emptyset$. ■

:⊢ **Definição 7.49.** Sejam $\boldsymbol{A}$ um anel e $Q \subseteq A$. Um *máximo divisor comum* de Q em $\boldsymbol{A}$ é um elemento $d \in A$ que satisfaz

1. $d \in \mathrm{Div}(Q)$;
2. Para todo $d' \in \mathrm{Div}(Q)$, $d' \preceq d$.

O conjunto dos máximos divisores comuns de Q em $\boldsymbol{A}$ é $\curlywedge_{\boldsymbol{A}}(Q)$. Se Q é um conjunto finito $Q = \{a_1, \ldots, a_n\}$, denota-se $\curlywedge_{\boldsymbol{A}}(Q) = \curlywedge_{\boldsymbol{A}}(a_1, \ldots, a_n)$ ou $\curlywedge_{\boldsymbol{A}}(Q) = (a_1, \ldots, a_n)$. O subíndice $\boldsymbol{A}$ será omitido sempre que possível.

Dualmente, um *mínimo múltiplo comum* de Q em $\boldsymbol{A}$ é um elemento $m \in A$ que satisfaz

1. $m \in \mathrm{Mul}_{\boldsymbol{A}}(Q)$;
2. Para todo $m' \in \mathrm{Mul}(Q)$, $m \preceq m'$.

O conjunto dos mínimos múltiplos comuns de Q em $\boldsymbol{A}$ é $\curlyvee_{\boldsymbol{A}}(Q)$. Se Q é um conjunto finito $Q = \{a_1, \ldots, a_n\}$, denota-se $\curlyvee_{\boldsymbol{A}} = \curlyvee_{\boldsymbol{A}}(a_1, \ldots, a_n)$ ou $\curlyvee_{\boldsymbol{A}}(Q) = [a_1, \ldots, a_n]$. O subíndice $\boldsymbol{A}$ será omitido sempre que possível.

⊢ **Proposição 7.53.** *Sejam $\boldsymbol{A}$ um anel e $Q \subseteq A$. Então*

1. $Q \subseteq \{0\} \Leftrightarrow \curlywedge(Q) = \{0\}$;
2. $A^{\prime} \cap Q \neq \emptyset \Rightarrow \curlywedge(Q) = A^{\prime}$.

Dualmente,

1. $Q \subseteq A^{\prime} \Leftrightarrow \curlyvee(Q) = A^{\prime}$;

2. $\{0\} \cap Q \neq \emptyset \Rightarrow \curlyvee(Q) = \{0\}$.

□ *Demonstração.* 1. Suponha que $Q \subseteq \{0\}$. Então $A = \mathrm{Div}(Q)$. Em particular, $0 \in \mathrm{Div}(Q)$. Ainda, para todo $d \in \mathrm{Div}(Q)$, vale que $d \preceq 0$, pois $d \in A$. Logo $0 \in \curlywedge(Q)$. Ainda, se $d \in \curlywedge(Q)$, então $0 \preceq d$, pois $0 \in \mathrm{Div}(Q)$ e $d \in \curlywedge(Q)$. Portanto $d = 0$. Reciprocamente, se $\curlywedge(Q) = \{0\}$, então $0 \in \mathrm{Div}(Q)$; ou seja, $\{0\} \cap \mathrm{Div}(Q) \neq \emptyset$, o que implica que $Q \subseteq \{0\}$.

2. Como $A^{\prime} \cap Q \neq \emptyset$, seja $a \in A^{\prime} \cap Q$. Se $u \in A^{\prime}$, então $u \in \mathrm{Div}(Q)$. Ainda, se $d \in \mathrm{Div}(Q)$, então, em particular, $d \preceq a$. Mas como $a \in A^{\prime}$, segue que $d \in A^{\prime}$. Logo $d \preceq u$, o que mostra que $u \in \curlywedge(Q)$. Por outro lado, se $d \in \curlywedge(Q)$, então $d \preceq a$. Como $a \in A^{\prime}$, então $d \in A^{\prime}$, e concluímos que $A^{\prime} = \curlywedge(Q)$.

Dualmente,

1. Suponha que $Q \subseteq A^{\prime}$. Então $A = \mathrm{Mul}(Q)$. Seja $u \in A^{\prime}$. Então $u \in \mathrm{Mul}(Q)$. Ainda, para todo $m \in \mathrm{Mul}(Q)$, vale que $u \preceq m$, pois $m \in A$. Logo $u \in \curlyvee(Q)$. Ainda, se $m \in \curlyvee(Q)$, então $m \preceq u$, pois $u \in \mathrm{Mul}(Q)$ e $m \in \curlyvee(Q)$. Portanto $m \in A^{\prime}$. Reciprocamente, se $\curlyvee(Q) = A^{\prime}$, então $1 \in \mathrm{Mul}(Q)$; ou seja, $A^{\prime} \cap \mathrm{Mul}(Q) \neq \emptyset$, o que implica $Q \subseteq A^{\prime}$.
2. Como $\{0\} \cap Q \neq \emptyset$, então $0 \in Q$. Note que $0 \in \mathrm{Mul}(Q)$. Ainda, se $m \in \mathrm{Mul}(Q)$, então $0 \preceq m$. Então segue que $0 \in \curlyvee(Q)$. Por outro lado, se $m \in \curlyvee(Q)$, então $0 \preceq m$, o que implica $m = 0$, e concluímos que $\curlyvee(Q) = \{0\}$. ■

7.2.4.2 Relação de associação

:⊢ **Definição 7.50.** Sejam $\boldsymbol{A}$ um anel e $a, a' \in A$. O elemento a é *associado* ao elemento a' em $\boldsymbol{A}$ se, e somente se, $a \preceq_{\boldsymbol{A}} a'$ e $a' \preceq_{\boldsymbol{A}} a$. Denota-se $a \sim_{\boldsymbol{A}} a'$. A relação $\sim_{\boldsymbol{A}}$ é a relação de *associação* em $\boldsymbol{A}$. Sempre que possível, o subíndice de $\sim_{\boldsymbol{A}}$ será omitido.

A relação $\sim$ de associação em anéis é uma equivalência, pois é a equivalência induzida pela pré-ordem de divisão $\preceq$ em anéis. Quando o anel A é um domínio, essa relação é equivalente a existir $u \in A^{\prime}$ tal que $au = a'$.

⊢ **Proposição 7.54.** *Sejam $\boldsymbol{A}$ um domínio e $a, a' \in A$. Então $a \sim a'$ se, e somente se, existe $u \in A^{\prime}$ tal que $au = a'$.*

□ *Demonstração.* Se $a \sim a'$, então $a \preceq a'$ e $a' \preceq a$. Isso é equivalente a existirem $q, q' \in A$ tais que $aq = a'$ e $a'q' = a$, o que é equivalente a $a = aqq'$ e $a' = a'q'q$. Como $\boldsymbol{A}$ é um domínio, a e a' não são divisores de 0, e concluímos que $qq' = q'q = 1$, portanto $q, q' \in A^{\prime}$.

Reciprocamente, se existe $u \in A^{\prime}$ tal que $au = a'$, então $a = a'u^{-1}$, portanto $a \preceq a'$ e $a' \preceq a$, logo $a \sim a'$. ■

⊢ **Proposição 7.55.** *Seja* $\boldsymbol{A}$ *um anel.*

1. *Para todos* $a_0, \ldots, a_{n-1}, a'_0, \ldots, a'_{n-1} \in A$*, se para todo* $i \in [n]$ *vale* $a_i \sim a'_i$*, então*
$$\bigtimes_{i\in[n]} a_i \sim \bigtimes_{i\in[n]} a'_i.$$

□ *Demonstração.* 1. Sejam $a_0, \ldots, a_{n-1}, a'_0, \ldots, a'_{n-1} \in A$ e $i \in [n]$. Se $a_i \sim a'_i$, então existe $u_i \in A^{\times}$ tal que $a_i u_i = a'_i$. Logo $a_0 \cdots a_{n-1} \cdot u_0 \cdots u_{n-1} = a'_0 \cdots a'_{n-1}$. Como $u_0 \cdots u_{n-1} \in A^{\times}$, segue que
$$\bigtimes_{i\in[n]} a_i \sim \bigtimes_{i\in[n]} a'_i.$$
■

⊢ **Proposição 7.56.** *Sejam* $\boldsymbol{D}$ *um domínio e* $d_0, \ldots, d_{n-1} \in D$ *não todos nulos. Se* d, d' *são máximos divisores comuns de* $d_0, \ldots, d_{n-1}$*, então* $d \sim d'$*.*

□ *Demonstração.* Como d é máximo divisor comum de $d_0, \ldots, d_{n-1}$ e d' é divisor comum de $d_0, \ldots, d_{n-1}$, então $d' \preceq d$. Analogamente, $d \preceq d'$. Portanto $d \sim d'$. ■

7.2.5 Domínios euclidianos

:⊢ **Definição 7.51.** Seja $\boldsymbol{D}$ um domínio. Uma *função euclidiana* em $\boldsymbol{D}$ é uma função $\phi : D \setminus \{0\} \to \mathbb{N}$ que satisfaz

1. Para todos $a \in D$ e $b \in D \setminus \{0\}$, existem $q, r \in D$ tais que
   1.1. $a = qb + r$;
   1.2. $r = 0$ ou $\phi(r) < \phi(b)$;
2. Para todos $a, b \in D \setminus \{0\}$, $\phi(a) \leq \phi(ab)$.

Nesse caso, q é chamado *quociente* e r é chamado de *resto* da divisão de a por b.

É possível mostrar que a segunda propriedade é desnecessária no seguinte sentido.

⊢ **Proposição 7.57.** *Sejam* $\boldsymbol{D}$ *um domínio e* $\phi : D \setminus \{0\} \to \mathbb{N}$ *uma função que satisfaz*

1. *Para todos* $a \in D$ *e* $b \in D \setminus \{0\}$*, existem* $q, r \in D$ *tais que*
   1.1. $a = qb + r$*;*
   1.2. $r = 0$ *ou* $\phi(r) < \phi(b)$*;*

Então existe uma função euclidiana em $\boldsymbol{D}$*.*

$\vdash$ **Proposição 7.58.** *Sejam $\boldsymbol{D}$ um domínio, ϕ uma função euclidiana em $\boldsymbol{D}$ e $a, b \in D$.*

$:\vdash$ **Definição 7.52.** Um *domínio euclidiano* é um domínio em que existe uma função euclidiana.

$\triangleright$ **Exercício 7.17.** *Os anéis $\mathbb{Z}$ e $\mathbb{Z}[i]$ são domínios euclidianos.*

$\triangleright$ **Exercício 7.18.** *Seja $\boldsymbol{C}$ um corpo. Então $\boldsymbol{C}[\boldsymbol{x}]$ é um domínio euclidiano.*

$:\vdash$ **Definição 7.53.** Um *domínio principal* é um domínio em que todo ideal é principal.

$\vdash$ **Proposição 7.59.** *Seja $\boldsymbol{D}$ um domínio euclidiano. Então $\boldsymbol{D}$ é um domínio principal.*

$\square$ *Demonstração.* Seja ϕ uma função euclidiana em $\boldsymbol{D}$ e $I \trianglelefteq D$. Se $I = \{0\}$, então $I = 0I$. Se $I \neq 0$, seja $a \in I \smallsetminus \{0\}$ tal que $\phi(a) = \min\{\phi(i) \mid a \in I \smallsetminus \{0\}\}$. Tome $b \in I$. Então existem $q, r \in D$ tais que $b = aq + r$. Então $r = aq - b \in I$, pois $a, b \in I$. Se $r \neq 0$, então $\phi(r) < \phi(a)$. Mas isso é absurdo, pois $\phi(a) = \min\{\phi(i) \mid a \in I \smallsetminus \{0\}\}$. Logo $r = 0$, o que implica $b = aq$; ou seja, $I \subseteq aD$. Como a inclusão inversa sempre vale, então $I = aD$. ■

$\vdash$ **Proposição 7.60.** *Sejam $\boldsymbol{D}$ um domínio euclidiano, ϕ uma função euclidiana em $\boldsymbol{D}$ e $d_1, d_2 \in D \smallsetminus \{0\}$. Então $d_1 \in D^*$ se, e somente se, $\phi(d_2) = \phi(d_1 d_2)$.*

$\square$ *Demonstração.* Se $d_1 \in D^*$, então existe $d_1^{-1} \in D$. Assim, temos que $\phi(d_1 d_2) \leq \phi(d_1 d_2 d_1^{-1}) = \phi(d_2)$. Por outro lado, sempre vale $\phi(d_2) \leq \phi(d_1 d_2)$. Logo $\phi(d_1 d_2) = \phi(d_2)$. Por outro lado, suponha $\phi(d_1 d_1) = \phi(d_2)$. Existem $q, r \in D$ tais que $d_2 = d_1 d_2 q + r$. Se $r \neq 0$, então, como $r = d_2 - d_1 d_2 q = d_2(1 - d_1 q)$ e $\boldsymbol{D}$ é domínio, segue que $1 - qa \neq 0$. Logo $\phi(r) = \phi(d_2(1 - d_1 q)) \geq \phi(d_2) = \phi(d_1 d_2)$, contradição. Logo $r = 0$ e temos $d_2 = d_1 d_2 q$. Como $d_2 \neq 0$ e $\boldsymbol{D}$ é domínio, então $d_1 q = 1$, o que implica $d_1 \in D^*$. ■

$\vdash$ **Proposição 7.61.** *Seja $\boldsymbol{D}$ um domínio euclidiano e ϕ uma função euclidiana em $\boldsymbol{D}$. Então*

$$D^* = \{d \in D \mid \phi(d) = \phi(1)\}.$$

$\square$ *Demonstração.* Tome $d_2 = 1$ na proposição anterior. ■

7.2.6 Domínios principais

7.2.7 Domínios de fatoração única

7.2.7.1 Irredutíveis e primos

:⊢ **Definição 7.54.** Seja $\boldsymbol{D}$ um domínio. Um elemento *irredutível* em $\boldsymbol{D}$ é um elemento $i \in D \smallsetminus (D^* \cup \{0\})$ que satisfaz, para todos $d, d' \in D$ tais que $i = dd'$,

$$d \in D^* \quad \text{ou} \quad d' \in D^*.$$

⊢ **Proposição 7.62.** *Sejam $\boldsymbol{D}$ um domínio e $i \in D$ irredutível. Para todo $i' \in D$ tal que $i \sim i'$, i' é irredutível.*

□ *Demonstração.* Se $i \sim i'$, existe $u \in D^*$ tal que $i' = ui$. Primeiro, devemos mostrar que $ui \notin (D^* \cup \{0\})$. Como $u \neq 0$ e $i \neq 0$ e $\boldsymbol{D}$ é domínio, então $ui \neq 0$. Supondo que $ui \in D^*$, então existe $u' \in D^*$ tal que $u'(ui) = 1$. Mas então $u'u = i^{-1}$, uma contradição porque $i \notin D^*$. Logo $ui \notin D^*$.

Agora, sejam $d, d' \in D$ tais que $ui = dd'$. Então $i = u^{-1}dd'$. Se $d \notin D^*$ e $d' \notin D^*$, então $u^{-1}d \notin D^*$ e $u^{-1}d' \neq 0$. Logo segue que $i = (u^{-1}d)d'$, uma contradição porque i é irredutível. ■

:⊢ **Definição 7.55.** Seja $\boldsymbol{D}$ um domínio. Um elemento *primo* em $\boldsymbol{D}$ é um elemento $p \in D \smallsetminus (D^* \cup \{0\})$ que satisfaz, para todos $d, d' \in D$ tais que $p \preceq dd'$,

$$p \preceq d \quad \text{ou} \quad p \preceq d'.$$

⊢ **Proposição 7.63.** *Sejam $\boldsymbol{D}$ um domínio e p primo.*

1. *Para todo $p' \in D$ tal que $p \sim p'$, p' é primo;*
2. *Para todos $d_0, \ldots, d_{n-1}$ tais que $p \preceq d_0 \cdots d_{n-1}$, existe $i \in [n]$ tal que $p \preceq d_i$;*
3. *O elemento p é irredutível.*

□ *Demonstração.* 1. Se $p \sim p'$, existe $u \in D^*$ tal que $p' = up$. Primeiro, devemos mostrar que $up \notin (D^* \cup \{0\})$. Como $u \neq 0$, $p \neq 0$ e $\boldsymbol{D}$ é domínio, então $up \neq 0$. Supondo que $up \in D^*$, então existe $u' \in D^*$ tal que $u'(up) = 1$. Mas isso implica $u'u = p^{-1}$, uma contradição porque $p \notin D^*$. Logo $p' = up \notin D^*$.

Agora, usaremos a comutatividade. Sejam $d, d' \in D$ tais que $up \preceq dd'$. Então

$$p = u^{-1}up \preceq u^{-1}dd'.$$

Como p é primo, $p \preceq u^{-1}d$ ou $p \preceq d'$. No primeiro caso, $up \preceq uu^{-1}d = d$; no segundo caso, segue pela comutatividade que $p' = up \preceq d'$.

2. Vamos mostrar por indução em n. O caso base é trivial. Para demonstrar o passo indutivo, suponhamos que a propriedade vale para um natural n. Então, se $p \preceq d_0 \cdots d_n$, então como p é primo, $p \preceq d_0 \cdots d_{n-1}$ ou $p \preceq d_n$. Se $p \preceq d_0 \cdots d_{n-1}$, pela hipótese de indução, existe $i \in [n]$ tal que $p \preceq d_i$; caso contrário, $p \preceq d_n$. Logo existe $i \in [n+1]$ tal que $p \preceq d_i$.
3. Se existem $d, d' \in D$ tais que $p = dd'$, então $p \preceq dd'$. Como p é primo, então $p \preceq d$ ou $p \preceq d'$. Se $p \preceq d$, então existe $q \in D$ tal que $pq = d$. Assim, segue que $p = dd' = pqd'$ e, como $\boldsymbol{D}$ é domínio, $1 = qd'$. Logo $d' \in D^*$. Analogamente, se $p \preceq d'$, segue que $d \in D^*$ (usamos a comutatividade). Portanto p é irredutível.

■

⊢ **Proposição 7.64.** *Seja $\boldsymbol{D}$ um domínio euclidiano que não é um corpo e ϕ uma função euclidiana em $\boldsymbol{D}$. Então $d_0 \in D$ tal que*

$$\phi(d_0) = \min \{\phi(d) \mid d \in D \smallsetminus (D^* \cup \{0\})\}$$

é um elemento irredutível em $\boldsymbol{D}$.

□ *Demonstração.* Primeiro, definamos $m := \min \{\phi(d) \mid d \in D \smallsetminus (D^* \cup \{0\})\}$ e notemos que existe tal mínimo porque o conjunto dos números naturais é bem ordenado. Notemos que $d_0 \notin D^* \cup \{0\}$ por definição. Suponha que $d_0 = d_1 d_2$, com $d_1, d_2 \in D$. Como $\boldsymbol{D}$ é um domínio, então $d_1 \neq 0$ e $d_2 \neq 0$. Suponha que $d_1, d_2 \notin D^*$. Então $\phi(d_1) \geq m = \phi(d_1 d_2) \geq \phi(d_1)$ e $\phi(d_2) \geq m = \phi(d_1 d_2) \geq \phi(d_2)$. Logo $\phi(d_1) = \phi(d_1 d_2)$ e $\phi(d_2) = \phi(d_1 d_2)$, o que implica $d_1, d_2 \in D^*$, que é absurdo. ■

7.2.7.2 Fatoração

⊢ **Definição 7.56.** Sejam $\boldsymbol{D}$ um domínio, $a \in D \smallsetminus \{0\}$ e $n \in \mathbb{N}$. Uma *fatoração* de a em n fatores é uma sequência $(p_i)_{i \in [n]}$ de irredutíveis tal que

$$a \sim \bigtimes_{i \in [n]} p_i.$$

Duas fatorações $(p_i)_{i \in [n]}$ e $(p'_i)_{i \in [n']}$ de a são *associadas* se, e somente se, $n = n'$ e existe uma permutação $\sigma \in \mathfrak{S}_n$ tal que, para todo $i \in [n]$, $p_i \sim p'_{\sigma(i)}$.

Claramente, a relação de associação de fatorações é uma relação de equivalência. Note que só exite uma sequência $(p_i)_{i \in [0]}$ — a sequência vazia $\emptyset$ — pois $D^0 = \emptyset$, e nesse caso

$$\bigtimes_{i \in [0]} p_i = 1,$$

pois $[0]$. Esse é um caso degenerado que admitimos para facilitar as demonstrações. Portanto um elemento $u \in D^*$ tem uma única fatoração $(p_i)_{i\in[0]}$ em 0 fatores, já que

$$u \sim 1 = \bigtimes_{i\in[0]} p_i.$$

Reciprocamente, se $a \in D \smallsetminus \{0\}$ tem fatoração em 0 fatores, então

$$a \sim \bigtimes_{i\in[0]} p_i = 1,$$

logo $a \in D^*$. Para $p \in D$ irredutível, $(p)_{i\in[1]}$ é uma fatoração de p e, se $(p_i)_{i\in[n]}$ é uma fatoração de p, então

$$p \sim \bigtimes_{i\in[n]} p_i.$$

Como $p \notin D^*$, $n \neq 0$. Se $n > 1$, como p é irredutível segue que $\bigtimes_{i\in[n-1]} p_i \in D^*$ ou $p_{n-1} \in D^*$, o que é uma contradição, pois os p_i são irredutíveis; logo $n = 1$ e $p \sim p_0$, portanto as fatorações $(p)_{i\in[1]}$ e $(p_0)_{i\in[1]}$ são associadas.

Isso mostra que elementos unitários têm sempre fatorações associadas, mas isso nem sempre é verdade para outros elementos do domínio. Também é verdade que todo elemento irredutível tem fatoração e que todas suas fatorações são associadas. Ressaltamos na proposição seguinte as propriedades aqui comentadas.

$\vdash$ **Proposição 7.65.** *Seja* $\boldsymbol{D}$ *um domínio. Então*

1. *A relação de associação de fatorações é uma relação de equivalência;*
2. *Todo elemento unitário ou irredutível tem fatoração e todas suas fatorações são associadas.*

$:\vdash$ **Definição 7.57.** Um *domínio de fatoração única* é um domínio $\boldsymbol{D}$ que satisfaz

1. (Existência de Fatoração) Todo elemento $d \in D \smallsetminus \{0\}$ tem fatoração;
2. (Unicidade de Fatoração) Todas fatorações de $d \in D \smallsetminus \{0\}$ são associadas.

Domínios de fatoração única também são chamados de domínios fatoriais. Antes de demonstrar a próxima proposição, enunciamos um lema simples que usaremos na demonstração.

$\vdash$ **Lema 7.66.** *Para todo* $k \in [n]$*, a função*

$$\begin{aligned} f_k \colon [n-1] &\longrightarrow [n] \smallsetminus \{k\} \\ i &\longmapsto \begin{cases} i, & i < k \\ i+1, & i \geq k. \end{cases} \end{aligned}$$

é bijetiva e sua inversa é

$$\begin{aligned} {f_k}^{-1}\colon [n] \smallsetminus \{k\} &\longrightarrow [n-1] \\ i &\longmapsto \begin{cases} i, & i < k \\ i-1, & i > k \end{cases} \end{aligned}$$

Para toda permutação $\sigma'\colon [n-1] \to [n-1]$*, a função*

$$\begin{aligned} \sigma\colon [n] &\longrightarrow [n] \\ i &\longmapsto \begin{cases} f_k \circ \sigma'(i), & i < n-1 \\ k, & i = n-1, \end{cases} \end{aligned}$$

é uma permutação.

A proposição a seguir relaciona duas noções da teoria de fatoração em anéis. A primeira é a unicidade de fatoração e a segunda é a relação entre elementos irredutíveis e primos.

⊢ **Proposição 7.67.** *Seja* $\boldsymbol{D}$ *um domínio tal que todo elemento não nulo tem fatoração. Então as fatorações de* $d \in D \smallsetminus \{0\}$ *são todas associadas se, e somente se, todo elemento irredutível de* $\boldsymbol{D}$ *é primo.*

□ *Demonstração.* Suponhamos, primeiro, que todas fatorações de $d \in D \smallsetminus \{0\}$ são associadas. Sejam p irredutível e $d, d' \in D$ tais que $p \preceq dd'$. Então existe $q \in D$ tal que $pq = dd'$. Se $d = 0$ ou $d' = 0$, então $p \preceq d$ ou $p \preceq d'$. Caso $d, d \in D \smallsetminus \{0\}$, então existem fatorações $(p_i)_{i\in[n]}$ de d e $(p'_i)_{i\in[n']}$ de d'. Ainda, como $p \neq 0$, segue que $q \neq 0$, pois D é domínio, e portanto existe fatoração $(q_i)_{i\in[m]}$ de q. Isso implica que

$$p \bigtimes_{i\in[m]} q_i \sim pq \sim \left(\bigtimes_{i\in[n]} p_i\right)\left(\bigtimes_{i\in[n']} p'_i\right).$$

Temos assim duas fatorações para pq, o que implica que existe $i \in [n]$ tal que $p \sim p_i$ ou $i \in [n']$ tal que $p \sim p'_i$. Então $p \preceq p_i$ ou $p \preceq p'_i$. No primeiro caso, como $p_i \preceq d$, segue que $p \preceq d$. No segundo caso, como $p'_i \preceq d'$, segue que $p \preceq d'$. Portanto p é primo.

Reciprocamente, suponhamos todo irredutível de $\boldsymbol{D}$ é primo. Sejam $d \in D \smallsetminus \{0\}$, e $(p_i)_{i\in[n]}$ e $(p'_i)_{i\in[n']}$ fatorações de d, de modo que

$$\bigtimes_{i\in[n]} p_i \sim d \sim \bigtimes_{i\in[n']} p'_i.$$

Todos os irredutíveis $p_0, \ldots, p_{n-1}$ e $p'_0, \ldots, p'_{n'-1}$ são primos. Mostraremos que essas fatorações são associadas por indução em $m := n \wedge n'$. Para o caso base, se $m = 0$,

então $n = 0$ ou $n' = 0$. Em ambos os casos, $d \sim 1$, ou seja, $d \in D^*$, portanto todas suas fatorações são associadas. Agora, seja $m > 0$ e suponhamos que todas fatorações com k e k' fatores de um elemento $d \in D \smallsetminus \{0\}$ tais que $k \wedge k' < m$ sejam associadas. Como p_{n-1} é primo e $p_{n-1} \preceq d \sim p'_0 \cdots p'_{n'-1}$, existe $k \in [n']$ tal que $p_{n-1} \preceq p'_k$. Portanto existe $q \in D$ tal que $p_{n-1}q = p'_k$. Como p_{n-1} e p'_k são irredutíveis, então $q \in D^*$. Portanto segue que

$$\begin{aligned}\left(\bigtimes_{i\in[n-1]} p_i\right) p_{n-1} &\sim \left(\bigtimes_{i\in[k]} p'_i\right) p_{n-1} \left(\bigtimes_{i\in[n'-k-1]} p'_{i+k+1}\right)\\ &= \left(\bigtimes_{i\in[k]} p'_i\right) \left(\bigtimes_{i\in[n'-k-1]} p'_{i+k+1}\right) p_{n-1}.\end{aligned}$$

Como $\boldsymbol{D}$ é domínio e $p_{n-1} \neq 0$, então

$$\bigtimes_{i\in[n-1]} p_i \sim \left(\bigtimes_{i\in[k]} p'_i\right) \left(\bigtimes_{i\in[n'-k-1]} p'_{i+k+1}\right).$$

Considerando a bijeção (7.66)

$$\begin{aligned} f_k \colon [n-1] &\longrightarrow [n] \smallsetminus \{k\}\\ i &\longmapsto \begin{cases} i, & i < k\\ i+1, & i \geq k\end{cases}\end{aligned}$$

e definindo, para todo $i \in [n'] \smallsetminus \{k\}$, $q_i := p'_{f_k(i)}$, temos uma sequência de irredutíveis $(q_i)_{i\in[n'-1]}$ tais que

$$\bigtimes_{i\in[n-1]} p_i \sim \bigtimes_{i\in[n'-1]} q_i.$$

Como $m > 0$, então $n > 0$ e $n' > 0$, portanto as expressões são duas fatorações, uma com $n-1$ e outra com $n'-1$ fatores. Como $(n-1) \wedge (n'-1) = m-1 < m$, vale a hipótese de indução, e então, $n-1 = n'-1$ e existe permutação σ' de $[n-1]$ tal que, para todo $i \in [n-1]$, $p_i \sim q_{\sigma'(i)}$. Então $n = n'$ e, considerando a permutação (7.66)

$$\begin{aligned} \sigma \colon [n] &\longrightarrow [n]\\ i &\longmapsto \begin{cases} f_k \circ \sigma'(i), & i < n-1\\ k, & i = n-1,\end{cases}\end{aligned}$$

segue que, para todo $i \in [n]$, $p_i \sim q_{\sigma(i)}$, pois, se $i < n-1$, então

$$p_i \sim q_{\sigma'(i)} = p'_{f_k(\sigma'(i))} = p'_{\sigma(i)},$$

e, se $i = n-1$, então $p_{n-1} \sim p'_k = p'_{\sigma(n-1)}$. Isso mostra que as fatorações $(p_i)_{i\in[n]}$ e $(p'_i)_{i\in[n']}$ são associadas. ■

:⊢ **Definição 7.58.** Sejam $\boldsymbol{D}$ um domínio, $a \in D \setminus \{0\}$ e $n \in \mathbb{N}$. Uma *fatoração reduzida* de a em n fatores é uma sequência $(p_i, e_i)_{i\in[n]}$ em que $(p_i)_{i\in[n]}$ é uma sequência de irredutíveis e $(e_i)_{i\in[n]}$ é uma sequência de naturais estritamente positivos tais que

$$a \sim \bigtimes_{i\in[n]} p_i^{e_i}$$

e, para todos $i, i' \in [n]$, $i \neq i'$ implica que $p_i \not\sim p_{i'}$.

⊢ **Proposição 7.68.** *Seja $\boldsymbol{D}$ um domínio de fatoração única. Então*

1. *(Existência de Fatoração Reduzida) Todo $d \in D \setminus \{0\}$ tem fatoração reduzida;*
2. *(Unicidade de Fatoração Reduzida) Para todo $d \in D \setminus \{0\}$, se*

$$(p_i, e_i)_{i\in[n]} \quad e \quad (p'_i, e'_i)_{i\in[n']}$$

são fatorações reduzidas de d, então $n = n'$ e existe permutação $\sigma \in \mathfrak{S}_n$ tal que, para todo $i \in [n]$, $p_i \sim p'_{\sigma(i)}$ e $e_i = e'_{\sigma(i)}$.

□ *Demonstração.* Seja $d \in D \setminus \{0\}$. Como $\boldsymbol{D}$ é domínio de fatoração única, existe fatoração $(p'_i)_{i\in[n']}$ de d em irredutíveis tais que

$$d \sim \bigtimes_{i\in[n']} p'_i.$$

Vamos considerar os conjuntos $I_i := \left\{ j \in [n'] \;\middle|\; p'_j \sim p'_i \right\}$ dos índices de elementos da fatoração de d que são associados. Esses conjuntos são uma partição de $[n']$: ou seja, $P := \{I_i \mid i \in [n']\}$ é uma partição de $[n']$. Para mostrar isso, seja $n := \#(P)$ e sejam $P_0, \ldots, P_{n-1}$ os elementos de P (note que os conjuntos P_i são os conjuntos I_i, mas reindexados). Primeiro, notemos que $\emptyset \notin P$ por definição. Segundo, notemos que

$$\bigcup_{j\in[n]} P_j = [n']$$

pois, para todo $j \in [n']$, existe $k \in [n]$ tal que $I_j = P_k$ e, como $j \in I_j$, isso implica que $j \in \bigcup_{i\in[n]} P_i$. Terceiro, sejam $j, k \in [n]$; se $P_j \cap P_k \neq \emptyset$, então seja $i \in P_j \cap P_k$; logo, para todos $i_j \in P_j, i_k \in P_k$, segue que $i_j \sim i \sim i_k$, o que implica $P_j = P_k$. Portanto podemos escrever

$$d \sim \bigtimes_{i\in[n']} p'_i = \bigtimes_{k\in[n]} \bigtimes_{i\in P_k} p'_i.$$

Sendo assim, seja $k \in [n]$. Definamos $p_k := p'_{\min P_k}$ e $e_k := \#(P_k)$. Segue claramente que p_k é irredutível e que $e_k \in \mathbb{N}^*$. Como, para todo $i \in P_k$, vale $p'_i \sim p_k$,

$$\bigtimes_{i\in P_k} p'_i \sim \bigtimes_{i\in P_k} p_k = p_k^{e_k}.$$

Assim, segue que

$$d \sim \bigtimes_{k\in[n]} \bigtimes_{i\in P_k} p_i' = \bigtimes_{k\in[n]} p_k^{e_k}.$$

A unicidade segue da unicidade de fatorações. ■

A proposição seguinte vale para um número finito de elementos, mas mostramos somente para dois elementos.

⊢ **Proposição 7.69.** *Seja* $\boldsymbol{D}$ *um domínio de fatoração única. Então, para todos* $a, b \in D \smallsetminus \{0\}$*, existe* $d \in \curlywedge(a, b)$*.*

□ *Demonstração.* Se $\{a, b\} \cap D^* \neq \emptyset$, então $\curlywedge(a, b) = D^*$. Suponhamos, então, que $\{a, b\} \notin D \smallsetminus (D^* \cup \{0\})$. Como $\boldsymbol{D}$ é domínio de fatoração única, existem fatorações $(p_i)_{i\in[n]}$ de a e $(q_i)_{i\in[n']}$ de b. Seja k o número de fatores p_i e q_j associados, e suponhamos, sem perda de generalidade, que, para todo $i \leq k$, $p_i \sim q_i$ e, para todo $i \geq k+1$ e $j \geq k+1$, $p_i \not\sim q_j$.

Se $k = 0$, mostraremos que $\curlywedge(a, b) = D^*$. Para isso, mostraremos que $1 \in \curlywedge(a, b)$. Notamos que $1 \in D^* \subseteq \mathrm{Div}(a, b)$. Agora, seja $d \in \mathrm{Div}(a, b)$. Como $0 \notin \{a, b\}$, então $0 \notin \mathrm{Div}(a, b)$. Se $d \in D^*$, então $d \preceq 1$, logo $1 \in \curlywedge(a, b)$. Se $d \in D \smallsetminus (D^* \cup \{0\})$, então existe p irredutível tal que $p \preceq d$, portanto $p \preceq a$ e $p \preceq b$. Como $\boldsymbol{D}$ é domínio de fatoração única, p é primo. Então segue que existem i e j tais que $p \preceq p_i$ e $p \preceq q_j$. Como esses elementos são irredutíveis, $p \sim p_i$ e $p \sim q_j$, portanto $p_i \sim q_j$, o que é uma contradição. Logo $1 \in \curlywedge(a, b)$. Como todos os máximos divisores comuns são associados, então $D^* = \curlywedge(a, b)$.

Se $k \geq 1$, mostraremos que $p_1 \cdots p_k \in \curlywedge(a, b)$. Primeiro, notamos que $p_1 \cdots p_k \preceq a$ e $q_1 \cdots q_k \preceq b$. Como $p_i \sim q_i$ para todo $i \leq k$, então $p_1 \cdots p_k \sim q_1 \cdots q_k$. Logo $p_1 \cdots p_k \preceq q_1 \cdots q_k$ e, da transitividade de $\preceq$, segue que $p_1 \cdots p_k \preceq b$. Logo $p_1 \cdots p_k \in \mathrm{Div}(a, b)$. Então, seja $d \in \mathrm{Div}(a, b)$. Se $d \in D^*$, então $d \preceq p_1 \cdots p_k$. Logo $p_1 \cdots p_k \in \curlywedge(a, b)$. Se $d \in D \smallsetminus (D^* \cup \{0\})$, então existem $r_1, \ldots, r_l$ irredutíveis tais que $d = r_1 \cdots r_l$. Por $\boldsymbol{D}$ ser domínio, $r_1, \ldots, r_l$ são primos. Como $d \preceq a$, então $r_1 \preceq p_1 \cdots p_k$. Como r_1 é primo, existe $i_1 \in \{1, \ldots, k\}$ tal que $r_1 \preceq p_{i_1}$. Seja $s_1 \in D$ tal que $r_1 s_1 = p_{i_1}$. Como r_1 e p_{i_1} são irredutíveis, então $r_1 \sim p_{i_1}$, o que implica $s_1 \in D^*$. Então $r_2 \cdots r_l \preceq p_1 \cdots p_{i_1-1} \cdot s_1 \cdot p_{i_1+1} \cdots p_k$, pois $\boldsymbol{D}$ é domínio. Repetindo o processo indutivamente, conclui-se que existem $i_1, \ldots, i_m$ tais que $r_j \sim p_{i_j}$ para todo $j \in \{1, \ldots, l\}$. Da mesma forma, conclui-se que que existem $i_1', \ldots, i_m'$ tais que $r_j \sim q_{i_j'}$ para todo $j \in \{1, \ldots, l\}$. Portanto, da reflexividade e transitividade de $\sim$, segue que $p_{i_j} \sim q_{i_j'}$ para todo i, j. Então $d \sim p_{i_1} \cdots p_{i_m} \sim q_{i_1'} \cdots q_{i_m'}$. Como $p_{i_1} \cdots p_{i_m} \preceq p_1 \cdots p_k$ e $q_{i_1'} \cdots q_{i_m'} \preceq q_1 \cdots q_k$, então $d \preceq p_1 \cdots p_k$. ■

7.3 Anéis polinomiais

7.3.1 Anel de polinômios

:⊢ **Definição 7.59.** Seja $\boldsymbol{A}$ um anel. O *anel de polinômios* em $\boldsymbol{A}$ é a lista

$$\boldsymbol{A}[\boldsymbol{x}] := (A[x], +, -, 0, \times, 1),$$

em que

1. $A[x]$ é o *conjunto de polinômios* em $\boldsymbol{A}$, definido por

$$A[x] := \left\{ p \in A^{\mathbb{N}} \;\middle|\; \#\left(\operatorname{supp}(p)\right) < \#\left(\mathbb{N}\right) \right\}$$

   e os elementos de $A[x]$ são *polinômios* em $\boldsymbol{A}$
2. $+$ é a *adição* de $A[x]$, definida por

$$\begin{aligned} +\colon A[x] \times A[x] &\longrightarrow A[x] \\ (p, q) &\longmapsto (p_n + q_n)_{n \in \mathbb{N}} \end{aligned}$$

3. $-$ é a *inversão aditiva* de $A[x]$, definida por

$$\begin{aligned} -\colon A[x] &\longrightarrow A[x] \\ p &\longmapsto (-p_n)_{n \in \mathbb{N}} \end{aligned}$$

4. 0 é a *nulidade* de $A[x]$, definida por

$$\begin{aligned} 0\colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto 0 \end{aligned}$$

5. $\times$ é a *multiplicação* de $A[x]$, definida por

$$\begin{aligned} \times\colon A[x] \times A[x] &\longrightarrow A[x] \\ (p, q) &\longmapsto \left(\mathop{\Large +}_{i=0}^{n} p_i q_{n-i} \right)_{n \in \mathbb{N}} \end{aligned}$$

6. 1 é a *unidade* de $A[x]$, definida por

$$\begin{aligned} 1\colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto \begin{cases} 1, & n = 0 \\ 0, & n \neq 0. \end{cases} \end{aligned}$$

Definimos os polinômios em $\boldsymbol{A}$ como as sequências em A com suporte finito. O suporte da sequência $p\colon \mathbb{N} \to A$, nesse caso, é o suporte com respeito ao grupo $(A, +, -, 0)$, ou seja, sequências que têm uma quantidade finita de entradas não nulas. Essas sequências são também chamadas de sequências eventualmente nulas. Se consideramos a topologia discreta em $\mathbb{N}$, que é sua topologia usual, toda função $s\colon \mathbb{N} \longrightarrow A$ é contínua, independente da topologia de A, o que significa que $A^{\mathbb{N}} = \mathscr{C}(\mathbb{N}, A)$. Além disso, o suporte topológico (fechado) de funções de $\mathbb{N}$ coincide com o suporte algébrico, e os compactos de $\mathbb{N}$ são os conjuntos finitos, portanto ter suporte finito é o mesmo que ter suporte compacto. Assim, os polinômios são as funções (contínuas) de $\mathbb{N}$ para A com suporte compacto.

Por enquanto, x é um símbolo qualquer, e pode ser substituído por qualquer outro símbolo de preferência, mas mais para frente ele será definido como um elemento de $A[x]$ de modo que os polinômios $p \in A[x]$ tenham a forma usual

$$p = \mathop{\vdots\!+\!\vdots}_{i=0}^{n} p_i x^i.$$

Antes disso, mostraremos que $A[x]$ é um anel, o que justifica seu nome.

$\vdash$ **Proposição 7.70.** *Seja $\boldsymbol{A}$ um anel. O anel de polinômios $\boldsymbol{A}[\boldsymbol{x}]$ em $\boldsymbol{A}$ é um anel.*

$\square$ *Demonstração.* Sabemos que $(A[x], +, -, 0)$ é um grupo comutativo, pois o grupo $(A, +, -, 0)$ é comutativo. Mostremos que $(A[x], \times, 1)$ é um monoide comutativo. Como $(A, \times, 1)$ é um monoide comutativo, segue que

1. (Associatividade) Para todos $p, q, r \in A[x]$,

$$\begin{aligned}
(pq)r &= \left(\mathop{+}_{i=0}^{n} (pq)_i r_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{i=0}^{n} \left(\mathop{+}_{j=0}^{i} p_j q_{i-j}\right) r_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{i=0}^{n} \mathop{+}_{j=0}^{i} p_j q_{i-j} r_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{0\leq j\leq i\leq n} p_j q_{i-j} r_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{i=0}^{n} \mathop{+}_{j=0}^{n-i} p_i q_j r_{n-i-j}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{i=0}^{n} p_i \left(\mathop{+}_{j=0}^{n-i} q_j r_{n-i-j}\right)\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{+}_{i=0}^{n} p_i (qr)_{n-i}\right)_{n\in\mathbb{N}} \\
&= p(qr);
\end{aligned}$$

2. (Unidade) Para todo $p \in A[x]$,

$$1p = \left(\left(1p_n + \mathop{+}_{i=1}^{n} p_i 0\right)\right)_{n\in\mathbb{N}} = (p_n)_{n\in\mathbb{N}} = p.$$

3. (Comutatividade) Para todos $p, q \in A[x]$,

$$pq = \left(\mathop{+}_{i=0}^{n} p_i q_{n-i}\right)_{n\in\mathbb{N}} = \left(\mathop{+}_{i=0}^{n} q_i p_{n-i}\right)_{n\in\mathbb{N}} = qp.$$

Por fim, mostremos que vale a distributividade. Para todos $p, q, r \in A[x]$,

$$\begin{aligned}
p(q+r) &= \left(\mathop{\bigplus}_{i=0}^{n} p_i(q+r)_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{\bigplus}_{i=0}^{n} p_i(q_{n-i}+r_{n-i})\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{\bigplus}_{i=0}^{n} (p_i q_{n-i}+p_i r_{n-i})\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{\bigplus}_{i=0}^{n} p_i q_{n-i} + \mathop{\bigplus}_{i=0}^{n} p_i r_{n-i}\right)_{n\in\mathbb{N}} \\
&= \left(\mathop{\bigplus}_{i=0}^{n} p_i q_{n-i}\right)_{n\in\mathbb{N}} + \left(\mathop{\bigplus}_{i=0}^{n} p_i r_{n-i}\right)_{n\in\mathbb{N}} \\
&= pq + pr.
\end{aligned}$$

■

:⊢ **Definição 7.60.** Seja $\boldsymbol{A}$ um anel. O *grau* (ou *grau máximo*) de $p \in A[x] \smallsetminus \{0\}$ é o número natural

$$|p|^\circ := \max \operatorname{supp}(p) = \max \{n \in \mathbb{N} \mid s_n \neq 0\}.$$

O grau de p é o maior índice de uma entrada não nula de p, que é sempre inteiro porque o suporte de p é finito. O grau do polinômio 0 não é definido. Se a definição acima fosse mudada do máximo max para o supremo sup, ele seria 0 pois $\sup \emptyset = 0$ em $\mathbb{N}$.

▷ **Exercício 7.19.** *Seja* $\boldsymbol{A}$ *um anel.*

1. *Para todos* $p, p' \in A[x] \smallsetminus \{0\}$ *tais que* $p + p' \neq 0$, $|p+p'|^\circ \geq |p|^\circ \vee |p'|^\circ$*;*
2. *Para todo* $p \in A[x] \smallsetminus \{0\}$, $|-p|^\circ = |p|^\circ$*;*
3. *Para todos* $p, p' \in A[x] \smallsetminus \{0\}$ *tais que* $pp' \neq 0$, $|pp'|^\circ \leq |p|^\circ + |p'|^\circ$*;*
4. $|1|^\circ = 0$.

A próxima definição justifica a notação $A[x]$, no sentido de que dá significado ao símbolo x.

:⊢ **Definição 7.61.** Seja $\boldsymbol{A}$ um anel. A *variável* de $A[x]$ é a função

$$\begin{aligned}
x \colon \mathbb{N} &\longrightarrow A \\
n &\longmapsto \begin{cases} 1, & n = 1 \\ 0, & n \neq 1. \end{cases}
\end{aligned}$$

Notemos que, para todo $k \in \mathbb{N}$, x^k é a função

$$\begin{aligned} x^k \colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto \begin{cases} 1, & n = k \\ 0, & n \neq k \end{cases} \end{aligned}$$

e assim podemos escrever todo polinômio $p \in A[x]$ como uma soma finita

$$p = \mathop{\bigplus}_{i=0}^{|p|^\circ} p_i x^i.$$

⊢ **Proposição 7.71.** *Sejam $\boldsymbol{A}$ um domínio e $p, q \in A[x] \smallsetminus \{0\}$. Então*

$$|pq|^\circ = |p|^\circ + |q|^\circ .$$

□ *Demonstração.* Seja $k := |p|^\circ + |q|^\circ$. O produto pq terá coeficientes $\bigplus_{i=0}^{n} p_i q_{n-i}$, com $n \in [k+1]$. Notemos que, para $k = |p|^\circ + |q|^\circ$,

$$\mathop{\bigplus}_{i=0}^{k} p_i q_{k-i} = p_{|p|^\circ} q_{|q|^\circ}.$$

Isso ocorre porque todos os termos desse somatório se anulam, menos quando $i = |p|^\circ$. Se $i > |p|^\circ$, temos $p_i = 0$; se $i < |p|^\circ$, então $k - i > |p|^\circ + |q|^\circ - |p|^\circ = |q|^\circ$, e temos $q_{k-i} = 0$. Em ambos os casos, $p_i q_{k-i} = 0$. Por definição, $p_{|p|^\circ} \neq 0$ e $q_{|q|^\circ} \neq 0$ e, como $\boldsymbol{A}$ é um domínio, isso implica que $p_{|p|^\circ} q_{|q|^\circ} \neq 0$. Logo $pq \neq 0$ e $|pq|^\circ = |p|^\circ + |q|^\circ$. ■

⊢ **Proposição 7.72.** *Seja $\boldsymbol{A}$ um anel. Então $\boldsymbol{A}$ é um domínio se, e somente se, $\boldsymbol{A}[\boldsymbol{x}]$ é um domínio.*

□ *Demonstração.* Suponha que $\boldsymbol{A}$ é um domínio e sejam $p_0, p_1 \in A[x] \smallsetminus \{0\}$. Então $|p_0 p_1|^\circ = |p_0|^\circ + |p_1|^\circ$, o que mostra que $p_0 p_1 \neq 0$. Logo $\boldsymbol{A}[\boldsymbol{x}]$ é um domínio. Reciprocamente, suponha que $\boldsymbol{A}[\boldsymbol{x}]$ é um domínio e sejam $a_0, a_1 \in A \smallsetminus \{0\}$. Então $a_0, a_1 \in A[x] \smallsetminus \{0\}$ e, portanto, $a_0 a_1 \neq 0$. Logo $\boldsymbol{A}$ é um domínio. ■

Podemos ver que $\boldsymbol{A}[\boldsymbol{x}]$ nunca é um corpo, pois x não tem inverso.

7.3.1.1 Anel de polinômios multivariados

:⊢ **Definição 7.62.** Sejam $\boldsymbol{A}$ um anel e $n \in \mathbb{N}$. O *anel de polinômios* em n variáveis em $\boldsymbol{A}$ é o conjunto

$$A[x_0, \ldots, x_{n-1}] := \left\{ p \in A^{\mathbb{N}^n} \;\middle|\; \#\left(\operatorname{supp}(p)\right) < \#\left(\mathbb{N}\right) \right\}.$$

Os elementos de $A[x]$ são os *polinômios em n variáveis* em $\boldsymbol{A}$.

7.3.2 Raízes de polinômios

⊢ **Proposição 7.73.** *Seja* $\boldsymbol{A}$ *um anel e* $f, g \in A[x]$. *Se* g *é mônico, então existem* $q, r \in A[x]$ *tais que* $f = qg + r$ *e* $r = 0$ *ou* $|r|^\circ < |g|^\circ$.

□ *Demonstração.* Sejam $n := |f|^\circ$, $m := |g|^\circ$,

$$f = \mathop{+}_{i=0}^{n} a_i x^i \quad \text{e} \quad g = \mathop{+}_{i=0}^{m-1} b_i x^i + x^m$$

Se $m \leq n$, definimos $q(x) := a_n x^{n-m}$ e $r := f - qg$ e temos

$$\begin{aligned} r(x) &= f(x) - q(x)g(x) \\ &= \mathop{+}_{i=0}^{n} a_i x^i - a_n x^{n-m} \left(\mathop{+}_{i=0}^{m-1} b_i x^i + x^m \right) \\ &= \mathop{+}_{i=0}^{n} a_i x^i - \left(\mathop{+}_{i=0}^{m-1} a_n b_i x^{n-m+i} + a_n x^n \right) \\ &= \mathop{+}_{i=0}^{n-1} a_i x^i - \mathop{+}_{i=n-m}^{n-1} a_n b_{i-n+m} x^i \\ &= \mathop{+}_{i=0}^{n-m-1} a_i x^i + \mathop{+}_{i=n-m}^{n-1} (a_i - a_n b_{i-n+m}) x^i. \end{aligned}$$

Daí segue que, se $r \neq 0$, $|r|^\circ \leq n - 1 < m \leq |g|^\circ$. ■

:⊢ **Definição 7.63.** Sejam $\boldsymbol{A}$ um anel, $p = +_{i=0}^{n} p_i x^i \in A[x]$ e $a \in A$. O *valor* de p em a é o elemento

$$p(a) := \mathop{+}_{i=0}^{n} p_i a^i \in A.$$

:⊢ **Definição 7.64.** Sejam $\boldsymbol{A}$ um anel e $p \in A[x]^* = A[x] \smallsetminus A$. Uma *raiz* de p é um elemento $r \in A$ tal que $p(r) = 0$.

▷ **Exercício 7.20.** *Sejam* $\boldsymbol{A}$ *um anel,* $p \in A[x]^*$ *e* $r \in A$. *Então* r *é raiz de* p *se, e somente se, existe* $q \in A[x]$ *tal que* $p(x) = (x - r)q(x)$.

7.3.3 Anel de séries

:⊢ **Definição 7.65.** Seja $\boldsymbol{A}$ um anel. O *anel de séries* (ou *séries de potências*) em $\boldsymbol{A}$ é a lista

$$\boldsymbol{A}[\![\boldsymbol{x}]\!] := (A[\![x]\!], +, -, 0, \times, 1),$$

em que

1. $A[\![x]\!] := A^{\mathbb{N}}$ é o *conjunto de séries* e um elemento de $A[\![x]\!]$ é denominado *série* (ou *série de potências*) em $\boldsymbol{A}$
2. $+$ é a *adição* de $A[\![x]\!]$, definida por
$$\begin{aligned} +\colon A[\![x]\!] \times A[\![x]\!] &\longrightarrow A[\![x]\!] \\ (s, s') &\longmapsto (s_n + s'_n)_{n\in\mathbb{N}} \end{aligned}$$
3. $-$ é a *inversão aditiva* de $A[\![x]\!]$, definida por
$$\begin{aligned} -\colon A[\![x]\!] &\longrightarrow A[\![x]\!] \\ s &\longmapsto (-s_n)_{n\in\mathbb{N}} \end{aligned}$$
4. 0 é a *nulidade* de $A[\![x]\!]$, definida por
$$\begin{aligned} 0\colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto 0 \end{aligned}$$
5. $\times$ é a *multiplicação* de $A[\![x]\!]$, definida por
$$\begin{aligned} \times\colon A[\![x]\!] \times A[\![x]\!] &\longrightarrow A[\![x]\!] \\ (s, s') &\longmapsto \left(\mathop{\bigplus}_{i=0}^{n} s_i s'_{n-i}\right)_{n\in\mathbb{N}} \end{aligned}$$
6. 1 é a *unidade* de $A[\![x]\!]$, definida por
$$\begin{aligned} 1\colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto \begin{cases} 1, & n = 0 \\ 0, & n \neq 0. \end{cases} \end{aligned}$$

Em particular, todo polinômio $p \in A[x]$ é uma série, ou seja, $A[x] \subseteq A[\![x]\!]$.

▷ **Exercício 7.21.** *Seja $\boldsymbol{A}$ um anel. O anel de séries $\boldsymbol{A}[\![\boldsymbol{x}]\!]$ de $\boldsymbol{A}$ é um anel.*

:⊢ **Definição 7.66.** Seja $\boldsymbol{A}$ um anel. A *variável* de $A[\![x]\!]$ é a função

$$\begin{aligned} x\colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto \begin{cases} 1, & n = 1 \\ 0, & n \neq 1. \end{cases} \end{aligned}$$

Assim como no anel de polinômios $A[x]$, notemos que, para todo $k \in \mathbb{N}$, x^k é a função

$$\begin{aligned} x^k \colon \mathbb{N} &\longrightarrow A \\ n &\longmapsto \begin{cases} 1, & n = k \\ 0, & n \neq k \end{cases} \end{aligned}$$

e assim podemos escrever todo polinômio $p \in A[x]$ como uma soma finita

$$p = \mathop{\boldsymbol{+}}_{n=0}^{|p|^\circ} p_n x^n.$$

Os elementos de $A[\![x]\!]$ que não são todos polinômios, no entanto; as sequências $s \in A[\![x]\!]$ tais que, para todo $n \in \mathbb{N}$, algum $n' \in \mathbb{N}$, $n' > n$, satisfaz $s_{n'} \neq 0$, não podem de fato ser escritos como

$$s = \mathop{\boldsymbol{+}}_{n \in \mathbb{N}} s_n x^n,$$

pois essa soma é infinita e a princípio soma infinita não está definida em um anel. No entanto, podemos adotar a expressão anterior como uma notação para a série s. Um outro método de resolver esse problema será apresentado a seguir, introduzindo uma topologia no anel de séries que possibilitará tratar de convergência de somas finitas.

7.3.3.1 Topologia do anel de séries

Além disso, podemos tentar munir $A[\![x]\!]$ com uma topologia, de modo que a sequencia

$$\left(\mathop{\boldsymbol{+}}_{n=0}^{m} s_n x^n \right)_{m \in \mathbb{N}}$$

seja convergente para s. Isso permitira escrever a soma infinita acima sem ser uma notação. Definimos uma topologia em $A[\![x]\!]$ como sendo a topologia produto de $A^{\mathbb{N}}$, com A munido da topologia discreta.

:⊢ **Definição 7.67.** Seja $\boldsymbol{A}$ um anel. A *topologia do anel de séries* $A[\![x]\!]$ é a topologia produto de $A^{\mathbb{N}}$, em que a topologia de A é considerada a topologia discreta.

Se o anel $\boldsymbol{A}$ tiver uma topologia diferente da discreta, podemos definir a topologia de $A[\![x]\!]$ como a topologia produto dessa topologia de $\boldsymbol{A}$. Isso permite que mais séries convirjam. No entanto, aqui consideraremos somente esse caso, que

é mais simples. Essa segunda topologia é mais grossa e é a topologia mais adequada para considerar, por exemplo, o anel de séries de um anel de séries, $A[\![x]\!][\![x']\!]$; nesse anel, a topologia que definimos considera $A[\![x]\!]$ com tendo a topologia discreta e não a topologia que definimos, e por isso o anel $A[\![x]\!][\![x']\!]$ não terá a mesma topologia de $A[\![x, x']\!]$, que pode ser definia analogamente à topologia acima definida.

Essa topologia pode ser dada por uma métrica e para isso definimos antes a ordem de uma série.

:⊢ **Definição 7.68.** Sejam $\boldsymbol{A}$ um anel e $s \in A[\![x]\!] \smallsetminus \{0\}$. A *ordem* (ou *grau mínimo*) da série s é o número natural

$$|s|_\circ := \min \operatorname{supp}(s) = \min \{n \in \mathbb{N} \mid s_n \neq 0\}.$$

A ordem da série s é o menor índice de um coeficiente não nulo de s (o grau de um polinômio é o maior índice), que existe a menos que $s = 0$. No caso em que $s = 0$, podemos definir sua ordem como sendo ∞, mas não definiremos a ordem de 0.

▷ **Exercício 7.22.** *Seja $\boldsymbol{A}$ um anel.*

1. *Para todos $s, s' \in A[x] \smallsetminus \{0\}$ tais que $s + s' \neq 0$, $|s + s'|_\circ \geq |s|_\circ \wedge |s'|_\circ$;*
2. *Para todo $s \in A[x] \smallsetminus \{0\}$, $|-s|_\circ = |s|_\circ$;*
3. *Para todos $s, s' \in A[x] \smallsetminus \{0\}$ tais que $ss' \neq 0$, $|ss'|_\circ \geq |s|_\circ + |s'|_\circ$;*
4. *$|1|_\circ = 0$.*

⊢ **Proposição 7.74.** *Seja $\boldsymbol{A}$ um anel. A topologia de $A[\![x]\!]$ é dada pela distância*

$$\begin{aligned} |\cdot, \cdot| \colon A[\![x]\!] \times A[\![x]\!] &\longrightarrow \mathbb{R} \\ (s, s') &\longmapsto |s, s'| := \begin{cases} 0 & s = s' \\ 2^{-|s-s'|_\circ} & s \neq s'. \end{cases} \end{aligned}$$

□ *Demonstração.* Segue direto de 17.3, para $\boldsymbol{G} = \boldsymbol{A}^+$, $m = 1$, $I = \mathbb{N}$ e $\alpha = 2^{-1}$. ■

▷ **Exercício 7.23.** *Seja $\boldsymbol{A}$ um anel e $(\sigma_n)_{n \in \mathbb{N}}$ uma sequência de séries $\sigma_n \in \boldsymbol{A}[\![\boldsymbol{x}]\!]$. A sequência de somas parciais*

$$\sigma'_{n' \in \mathbb{N}} = \left(\mathop{\Bigg+}_{n=0}^{n'} \sigma_n \right)_{n' \in \mathbb{N}}$$

converge se, e somente se, para todo $i \in \mathbb{N}$, a sequência de coeficientes de x_i de $\sigma'_{n'}$ é eventualmente constante.

No caso de uma série $s \in \boldsymbol{A}[\![\boldsymbol{x}]\!]$, considerarmos $\sigma_n := s_n x^n$, e claramente a sequência de somas parciais de $(\sigma_n)_{n\in\mathbb{N}}$ converge, portanto obtemos

$$s = \mathop{+}_{n\in\mathbb{N}} s_n x^n.$$

Isso mostra que com essa topologia, a igualdade anterior deixa de ser somente uma notação e se torna de fato um resultado de convergência. Além disso, podemos expressar o produto de duas séries mais simplesmente usando esse resultado de convergência. Para todas séries $s, s' \in A[\![x]\!]$,

$$ss' = \left(\mathop{+}_{n\in\mathbb{N}} s_n x^n\right)\left(\mathop{+}_{n'\in\mathbb{N}} s'_{n'} x^{n'}\right) = \mathop{+}_{n,n'\in\mathbb{N}} s_n s'_{n'} x^{n+n'}.$$

7.4 Corpos

7.4.1 Corpo e subcorpo

Já demos a definição de um corpo, mas agora pretendemos ressaltar essa definição em termos da estrutura algébrica do corpo.

:⊢ **Definição 7.69.** Um *corpo* é uma lista $\boldsymbol{C} = (C, +, -, 0, \times, \diagup, 1)$ tal que

1. $\boldsymbol{C}^+ := (C, +, -, 0)$ é um grupo comutativo, em que $+$ é a *adição*, $-$ é a *inversão aditiva* e 0 é a *nulidade* (ou o *zero*) de $\boldsymbol{C}$;
2. $\boldsymbol{C}^\times := (C, \times, 1)$ é um monoide comutativo, em que $\times$ é a *multiplicação* e 1 é a *unidade* (ou o *um*) de $\boldsymbol{C}$;
3. $\boldsymbol{C}^\diagup := (C^\diagup, \times, \diagup, 1)$ é um grupo comutativo, em que $C^\diagup = C \setminus \{0\}$ é conjunto de elementos invertíveis de C e $\diagup$ é a *inversão multiplicativa* de $\boldsymbol{C}$;
4. A multiplicação $\times$ é distributiva sobre a adição $+$.

7.4.2 Extensões de corpos

:⊢ **Definição 7.70.** Seja $\boldsymbol{C}$ um corpo. Uma *extensão* de $\boldsymbol{C}$ é um corpo $\boldsymbol{E}$ tal que $\boldsymbol{C}$ é um subcorpo de $\boldsymbol{E}$.

⊢ **Proposição 7.75.** *Sejam $\boldsymbol{C}$ um corpo e $\boldsymbol{E} = (E, +, \cdot)$ uma extensão de $\boldsymbol{C}$. Então $(E, \oplus, \odot)$ é um espaço vetorial sobre $\boldsymbol{C}$, em que $\oplus := +$ e $\odot := \cdot|_{C \times E}$.*

□ *Demonstração.* Como $\boldsymbol{E}$ é um anel, então $(E, \oplus) = (E, +)$ é um grupo comutativo com elemento neutro 0 do corpo. Agora, como $\odot$ é a restrição de $\cdot$ a $C \times E$, então, para todo $e \in E$, $1 \odot e = 1e = e$ e, para todos $c_1, c_2 \in C$, $(c_1 \cdot c_2) \odot e = c_1 c_2 e = c_1 \odot (c_2 \odot e)$. Por fim, como $\cdot$ é distributiva sobre $+$, segue que, para todos $c \in C$ e $e_1, e_2 \in E$, $c \odot (e_1 \oplus e_2) = c(e_1 + e_2) = ce_1 + ce_2 = c \odot e_1 \oplus c \odot e_2$ e, para todos $c_1, c_2 \in C$ e $e \in E$, $(c_1 + c_2) \odot e = (c_1 + c_2)e = c_1 e + c_2 e = c_1 odote \oplus c_2 \odot e$. Por tanto, concluímos que $(E, \oplus, \odot)$ é um espaço vetorial sobre $\boldsymbol{C}$. ■

Notação. Nesse caso, quando não houver ambiguidade, denotaremos $\oplus$ como $+$ e $\odot$ como $\cdot$, bem como todas outras notações relacionadas às operações do espaço vetorial.

:⊢ **Definição 7.71.** Sejam $\boldsymbol{C}$ um corpo e $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$. Então a *dimensão* da extensão $\boldsymbol{E}$ com respeito a $\boldsymbol{C}$ é a dimensão do espaço vetorial $\boldsymbol{E}$ sobre $\boldsymbol{C}$. A extensão $\boldsymbol{E}$ é *finita* de sua dimensão é finita e *infinita* se sua dimensão é infinita.

:⊢ **Definição 7.72.** Sejam $\boldsymbol{C}$ um corpo e $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$. Um *elemento algébrico* sobre $\boldsymbol{C}$ é um elemento $\alpha \in E$ que é raiz de um polinômio em $C[x]^*$.

$\vdash$ **Proposição 7.76.** *Sejam $\boldsymbol{C}$ um corpo e $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$. Então*

1. *Todo elemento de C é algébrico sobre $\boldsymbol{C}$;*
2. *Se $\alpha \in E \smallsetminus \{0\}$ é um elemento algébrico sobre $\boldsymbol{C}$, então existem $c_0, \ldots, c_n \in C$ tais que $c_0 \neq 0$ e*

$$\mathop{+}_{i=0}^{n} c_i \alpha^i = 0.$$

$\square$ *Demonstração.* 1. Seja $\alpha \in C$. Então $\alpha \in E$, pois $C \subseteq E$. Tomando $p(x) = x - \alpha$, temos que $p(\alpha) = \alpha - \alpha = 0$.

2. Sejam $\alpha \in E$ um elemento algébrico sobre $\boldsymbol{C}$ e $c'_0, \ldots, c'_m \in C$ tais que

$$\mathop{+}_{i=0}^{m} c'_i \alpha^i = 0.$$

Como $c'_0, \ldots, c'_m$ não são todos nulos, seja $k := \min\{i \in [m+1] \mid c'_i \neq 0\}$. Então, para todo $i < k$, $c'_i = 0$, e segue que

$$0 = \mathop{+}_{i=0}^{m} c'_i \alpha^i = \mathop{+}_{i=k}^{m} c'_i \alpha^i = \alpha^k \mathop{+}_{i=k}^{m} c'_i \alpha^{i-k}$$

e, como E é corpo e $\alpha \neq 0$, podemos dividir por α^k em ambos lados e temos

$$\mathop{+}_{i=k}^{m} c'_i \alpha^{i-k} = 0.$$

Fazendo, $n := m - k$ e, para todo $i \in [n+1]$, $c_i := c'_{k+i}$, temos que $c_0, \ldots, c_n \in E$ — com $c_0 = c'_k \neq 0$ e, portanto, não todos nulos — tais que

$$\mathop{+}_{i=0}^{n} c_i \alpha^i = \mathop{+}_{i=0}^{n} c'_{k+i} \alpha^i = \mathop{+}_{i=k}^{m} c'_i \alpha^{i-k} = 0.$$

$\blacksquare$

$:\vdash$ **Definição 7.73.** Sejam $\boldsymbol{C}$ um corpo, $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$ e $\alpha \in E$ um elemento algébrico sobre $\boldsymbol{C}$. Um *polinômio minimal* de α sobre $\boldsymbol{C}$ é um polinômio $p \in C[x]$ que satisfaz

1. $p(\alpha) = 0$;
2. p é mônico;
3. $|p|^\circ = \min\{|f|^\circ \mid f \in C[x]^*, f(\alpha) = 0 \text{ e } f \text{ é mônico}\}$.

$\vdash$ **Proposição 7.77.** *Sejam $\boldsymbol{C}$ um corpo, $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$ e $\alpha \in E$ um elemento algébrico sobre $\boldsymbol{C}$. Então existe um único polinômio minimal de α sobre $\boldsymbol{C}$.*

□ *Demonstração.* Pela definição de elemento algébrico, existe $p \in C[x]^*$ tal que $p'(\alpha) = 0$. Seja $n := |p'|^\circ$ e $p'(x) = +_{i=0}^{n} a'_i x^i$. Como $\boldsymbol{C}$ é corpo e $a'_n \neq 0$, existe $(a'_n)^{-1} \in C$. Definindo

$$p := (a'_n)^{-1} p = \mathop{+}_{i=0}^{n-1} (a'_n)^{-1} a_i x^i + x^n,$$

segue que $p \in C[x]^*$, $|p|^\circ = n$, $p(\alpha) = (a'_n)^{-1} p'(\alpha) = 0$ e p é mônico. Portanto $\{|f|^\circ \mid f \in C[x]^*, f(\alpha) = 0 \text{ e } f \text{ é mônico}\}$ não é vazio, e, como é subconjunto de $\mathbb{N}$, admite mínimo, o que mostra que existe polinômio minimal de α sobre $\boldsymbol{C}$.

Para mostrar a unicidade, suponhamos que p_1 e p_2 são polinômios minimais de α sobre $\boldsymbol{C}$. Pela primeira propriedade de polinômio minimal, $(p_1 - p_2)(\alpha) = p_1(\alpha) - p_2(\alpha) = 0$. Pela terceira propriedade de polinômio minimal, $|p_1|^\circ = |p_2|^\circ$. Seja $n := |p_1|^\circ$ e sejam $p_1 = +_{i=0}^{n} a_i x^i$ e $p_2 = +_{i=0}^{n} b_i x^i$. Pela segunda propriedade de polinômio minimal, $a_n = b_n = 1$. Então $a_n - b_n = 0$ e

$$(p_1 - p_2)(x) = \mathop{+}_{i=0}^{n-1} (a_i - b_i) x^i,$$

e conclui-se que $|p_1 - p_2|^\circ < n$. Se $p_1 \neq p_2$, existe $i \in [n]$ tal que $a_i \neq b_i$. Então seja $k := \max\{i \in [n] \mid a_i \neq b_i\}$. Assim, para todo $i > k$, $a_i = b_i$, o que implica $a_i - b_i = 0$, e temos que $|p_1 - p_2|^\circ = k$ e

$$(p_1 - p_2)(x) = \mathop{+}_{i=0}^{k} (a_i - b_i) x^i.$$

Como $\boldsymbol{C}$ é corpo e $a_k - b_k \neq 0$, existe $(a_k - b_k)^{-1} \in C$. Definindo

$$p := (a_k - b_k)^{-1}(p_1 - p_2) = \mathop{+}_{i=0}^{k-1} (a_k - b_k)^{-1}(a_i - b_i) x^i + x^k,$$

segue que $p \in C[x]^*$, $|p|^\circ = k$, $p(\alpha) = (a_k - b_k)^{-1}(p_1 - p_2)(\alpha) = (a_k - b_k)^{-1}(p_1(\alpha) - p_2(\alpha)) = 0$ e p é mônico. Mas $|p|^\circ = k < n = |p_1|^\circ = |p_2|^\circ$, o que contradiz a minimalidade do grau de p_1 e p_2. Portanto $p_1 = p_2$ e está provada a unicidade. ■

⊢ **Proposição 7.78.** *Seja $\boldsymbol{C}$ um corpo e $p \in C[x]^*$. Se p é mônico e redutível, então existem $p_1, p_2 \in C[x]^*$ tais que $p = p_1 p_2$ e p_1 e p_2 são mônicos.*

□ *Demonstração.* Como p é redutível, existem $p'_1, p'_2 \in C[x]^*$ tais que $p = p'_1 p'_2$. Sejam $n := |p_1|^\circ$, $p_1 = +_{i=0}^{n} a_i x^i$, e $m := |p_2|^\circ$, $p_2 = +_{i=0}^{m} b_i x^i$. Pela definição de produto, sabemos que $a_n b_m = 1$, pois p é mônico. Como $\boldsymbol{C}$ é um corpo, existem $(a_n)^{-1}, (b_m)^{-1} \in C$. Definindo

$$p_1 := (a_n)^{-1} p'_1 = \mathop{+}_{i=0}^{n-1} (a_n)^{-1} a_i x^i + x^n \quad \text{e} \quad p_2 := (b_m)^{-1} p'_2 = \mathop{+}_{i=0}^{m-1} (b_m)^{-1} b_i x^i + x^m,$$

segue que $p_1, p_2 \in C[x]^*$ são mônicos e que

$$p = p_1' p_2' = (a_n)p_1(b_m)p_2 = (a_n b_m)p_1 p_2 = p_1 p_2.$$

■

⊢ **Proposição 7.79.** *Sejam $\boldsymbol{C}$ um corpo, $\boldsymbol{E}$ uma extensão de $\boldsymbol{C}$, $\alpha \in E$ um elemento algébrico sobre $\boldsymbol{C}$ e $p \in C[x]^*$ um polinômio mônico tal que $p(\alpha)$. Então p é o polinômio minimal de α sobre $\boldsymbol{C}$ se, e somente se, p é irredutível em $\boldsymbol{C}[x]$.*

□ *Demonstração.* Suponhamos que p é polinômio minimal de α sobre $\boldsymbol{C}$. Se p não é irredutível em $C[x]$, como p é mônico, então existem $p_1, p_2 \in C[x]^*$ mônicos tais que $0 < |p_i|^\circ < |p|^\circ$ para todo $i \in \{1, 2\}$. Como $p(\alpha) = p_1(\alpha)p_2(\alpha) = 0$ e $\boldsymbol{C}$ é corpo, segue que $p_1(\alpha) = 0$ ou $p_2(\alpha) = 0$. No primeiro caso, $p_1 \in C[x]^*$ é um polinômio mônico, $p_1(\alpha) = 0$ e $|p_1|^\circ < |p|^\circ$, o que contradiz a minimalidade de p. No segundo caso, $p_2 \in C[x]^*$ é um polinômio mônico, $p_2(\alpha) = 0$ e $|p_2|^\circ < |p|^\circ$, o que também contradiz a minimalidade de p, e temos um absurdo. Portanto p é irredutível em $\boldsymbol{C}[x]$.

Reciprocamente, suponhamos que p é irredutível em $\boldsymbol{C}[x]$. Por hipótese, $p(\alpha) = 0$ e p é mônico. Seja $p_\alpha \in C[x]^*$ o polinômio minimal de α sobre $\boldsymbol{C}$. Então $|p_\alpha|^\circ \leq |p|^\circ$. Como p_α é mônico, exitem $q, r \in C[x]$ tais que $p = qp_\alpha + r$. Como $p(\alpha) = p_\alpha(\alpha) = 0$, então $r(\alpha) = p(\alpha) - q(\alpha)p_\alpha(\alpha) = 0$. Se $r \neq 0$, então $|r|^\circ < |p_\alpha|^\circ$. Seja $n := |r|^\circ$. Como $r(\alpha) = 0$, segue que $|r|^\circ > 0$. Então seja $r(x) = +_{i=0}^{n} a_i x^i$. Como $\boldsymbol{C}$ é corpo, existe $(a_n)^{-1} \in C$. Assim, definindo

$$p' := (a_n)^{-1} r = \mathop{+}_{i=0}^{n-1} (a_n)^{-1} a_i x^i + x^n,$$

e segue que $p' \in C[x]^*$, $p'(\alpha) = (a_n)^{-1} p'(\alpha) = 0$ e p' é mônico. Mas $|p'|^\circ = n = |r|^\circ < |p_\alpha|^\circ$, o que contradiz a minimalidade do grau de p_α. Então $r \neq 0$. Se $r = 0$, então $p = qp_\alpha$. Mas p é irredutível, e $p_\alpha \in C[x]^*$, o que implica $q \in C$. Como p e p_α são mônicos, segue que $q = 1$ e, portanto, $p = p_\alpha$. ■

7.4.3 Extras

⊢ **Teorema 7.80.** *Seja $k \subseteq K$ uma extensão de corpos, $\overline{k}$ fecho algébrico de k. Então as seguintes condições são equivalentes:*

1. *DIAGRAMA COMUTATIVO*
2. *K é corpo de raízes sobre k de uma família $(f_i)_{i \in I}$ de polinômios em $k[x] \smallsetminus k$;*
3. *Se $f \in k[x] \smallsetminus k$ é irredutível em $k[x]$ com raiz α, então $f(x) = c(x - \alpha_1) \cdots (x - \alpha_n)$ em $k[x]$, com $\alpha_1 = \alpha$ e $c \in k \smallsetminus \{0\}$.*

$$\begin{tikzcd} k \arrow[r, "\cdot"] \arrow[d, "\cdot"'] & \overline{k} \\ K \arrow[ur, "\sigma"'] & \end{tikzcd}$$

:⊢ **Definição 7.74.** Uma extensão de corpos $k \subseteq K$ é uma *extensão normal* se ela satisfaz uma das três condições do teorema acima.

OBS: Se $k \subseteq \overline{k}$ é uma extensão algébrica, então todo $\beta \in \overline{k}$ é algébrico sobre k vale para $\beta \in K$, então $k \subseteq K$ é extensão algébrica.

Se $k \subseteq K$ é extensão algébrica, então $k \subseteq K \subseteq \overline{K}$ são extensões algébricas, então $k \subseteq \overline{K}$ é extensão algébrica. Então $\overline{k} \sim \overline{K}$ é fecho algébrico de k.

⊢ **Proposição 7.81.** *Seja $k \subseteq K$ um extensão algébrica e $\sigma : K \to K$ um homomorfismo de corpos que satisfaz $\sigma|_k = id|_k$. Então σ é um isomorfismo de corpos.*

:⊢ **Definição 7.75.** Seja $E \subseteq F$ uma extensão algébrica e $\sigma : E \to L$ um homomorfismo de corpos tal que L é algebricamente fechado, $\sigma(E) \subseteq L$ é uma extensão algébrica (L é fecho algébrico de $\sigma(E)$)

$$S_\sigma := \{\mu : \mu : F \to L \text{ homomorfismo de corpos}, \mu|_E = \sigma\}.$$

⊢ **Lema 7.82.** *$|S_\sigma|$ depende de $E \subseteq F$, mas não depende de σ nem de L.*

:⊢ **Definição 7.76.** Seja $E \subseteq F$ uma extensão algébrica. O *grau de separabilidade* da extensão é $[F : E]_S := |S_\sigma|$. (Pode escolher $l = \overline{E}$ e σ inclusão.)

⊢ **Teorema 7.83.** *1. $E \subseteq F \subseteq K$ extensões algébricas, então*

$$[K : E]_S = [K : F]_S [F : E]_S$$

2. $E \subseteq F$ extensão finita (logo algébrica), então

$$[F : E]_S \leq [F : E]$$

:⊢ **Definição 7.77.** Seja $E \subseteq K$ uma extensão finita. Ela é *separável* se $[K : E]_S = [K : E]$.

⊢ **Corolário 7.84.** *Sejam $E \subseteq F \subseteq K$ extensões de corpos, $[K : E] < \infty$, $E \subseteq K$ separável. Então $E \subseteq F$ e $FsubseteqK$ são separáveis.*

□ *Demonstração.*

$$[K : F]_S [F : E]_S = [K : E]_S = [K : E] = [K : F][F : E].$$

Como $[F : E]_S \leq [F : E]$ e $[K : F]_S \leq [K : F]$, segue o corolário. ■

7.5 Ação de anel

Consideramos brevemente ações de anéis em grupos comutativos. Na definição usamos o fato de que, se $\boldsymbol{G}$ é um grupo comutativo, $\mathcal{H}(\boldsymbol{G})$ é um anel com as operações puxadas para o espaço de funções, a soma pontual sendo a soma do anel e a composição de função sendo o produto. De fato, se X é um conjunto e $\boldsymbol{G}$ um grupo comutativo, o conjunto $\mathcal{F}(X, G)$ de funções de X para G, com a soma pontual e a composição, é um anel, e basta mostrar que quando $X = G$, o conjunto $\mathcal{H}(\boldsymbol{G})$ de endomorfismos de grupo de $\boldsymbol{G}$ é um subanel de $\mathcal{F}(G)$. Demonstramos isso a seguir.

7.5.1 Anel de endomorfismos de grupo

$\vdash$ **Proposição 7.85.** *Sejam X um conjunto e $\boldsymbol{G}$ um grupo comutativo. Então*

$$\boldsymbol{\mathcal{F}}(X, \boldsymbol{G}) := (\mathcal{F}(X, G), +, -, 0, \circ, \mathrm{I})$$

é um anel[1].

$\square$ *Demonstração.* 1. $(\mathcal{F}(X, G), +, -, 0)$ é um grupo comutativo.

1.1. Para todos $f_0, f_1, f_2 \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}((f_0 + f_1) + f_2)(x) &= (f_0 + f_1)(x) + f_2(x)\\ &= (f_0(x) + f_1(x)) + f_2(x)\\ &= f_0(x) + (f_1(x) + f_2(x))\\ &= f_0(x) + (f_1 + f_2)(x)\\ &= (f_0 + (f_1 + f_2))(x);\end{aligned}$$

1.2. Para todos $f \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}(0 + f)(x) &= 0(x) + f(x)\\ &= 0 + f(x)\\ &= f(x)\\ &= f(x) + 0\\ &= f(x) + 0(x)\\ &= (f + 0)(x);\end{aligned}$$

[1] No sentido em que a multiplicação $\circ$ não precisa ser comutativa.

1.3. Para todos $f \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}((-f) + f)(x) &= (-f)(x) + f(x) \\ &= -f(x) + f(x) \\ &= 0 \\ &= f(x) + (-f(x)) \\ &= f(x) + (-f)(x) \\ &= (f + (-f))(x);\end{aligned}$$

1.4. Para todos $f_0, f_1 \in \mathcal{F}(X, G)$ e $x \in X$,

$$(f_0 + f_1)(x) = f_0(x) + f_1(x) = f_1(x) + f_0(x) = (f_1 + f_0)(x).$$

2. $(\mathcal{F}(X, G), \circ, \mathrm{I})$ é um monoide.

2.1. Para todos $f_0, f_1, f_2 \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}((f_0 \circ f_1) \circ f_2)(x) &= (f_0 \circ f_1)(f_2(x)) \\ &= (f_0(f_1(f_2(x))) \\ &= f_0((f_1 \circ f_2)(x)) \\ &= (f_0 \circ (f_1 \circ f_2))(x);\end{aligned}$$

2.2. Para todos $f \in \mathcal{F}(X, G)$ e $x \in X$,

$$(\mathrm{I} \circ f)(x) = \mathrm{I}(f(x)) = f(x) = f(\mathrm{I}(x)) = (f \circ \mathrm{I})(x).$$

3. A composição $\circ$ é distributiva à esquerda e à direita sobre a soma pontual $+$.

3.1. Para todos $f, f_0, f_1 \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}(f \circ (f_0 + f_1))(x) &= f((f_0 + f_1)(x)) \\ &= f(f_0(x) + f_1(x)) \\ &= f(f_0(x)) + f(f_1(x)) \\ &= (f \circ f_0)(x) + (f \circ f_1)(x);\end{aligned}$$

3.2. Para todos $f_0, f_1, f \in \mathcal{F}(X, G)$ e $x \in X$,

$$\begin{aligned}((f_0 + f_1) \circ f)(x) &= (f_0 + f_1)(f(x)) \\ &= f_0(f(x)) + f_1(f(x)) \\ &= (f_0 \circ f)(x) + (f_1 \circ f)(x).\end{aligned}$$

■

$\vdash$ **Proposição 7.86.** *Seja $\boldsymbol{G}$ um grupo comutativo. Então*

$$\boldsymbol{\mathcal{H}}(\boldsymbol{G}) := (\mathcal{H}(\boldsymbol{G}), +, -, 0, \circ, \mathrm{I})$$

é um anel[2].

$\square$ *Demonstração.* Como $\mathcal{H}(\boldsymbol{G}) \subseteq \mathcal{F}(G)$, basta mostrar que $\mathcal{H}(\boldsymbol{G})$ é subanel de $\boldsymbol{\mathcal{F}}(\boldsymbol{G})$.

1. $(\mathcal{H}(\boldsymbol{G}), +, -, 0)$ é subgrupo.

   1.1. Para todos $h_0, h_1 \in \mathcal{H}(\boldsymbol{G})$ e $g_0, g_1 \in G$,

$$\begin{aligned}(h_0 + h_1)(g_0 + g_1) &= h_0(g_0 + g_1) + h_1(g_0 + g_1)\\ &= h_0(g_0) + h_0(g_1) + h_1(g_0) + h_1(g_1)\\ &= h_0(g_0) + h_1(g_0) + h_0(g_1) + h_1(g_1)\\ &= (h_0 + h_1)(g_0) + (h_0 + h_1)(g_1);\end{aligned}$$

   1.2. Para todos $h \in \mathcal{H}(\boldsymbol{G})$ e $g_0, g_1 \in G$,

$$\begin{aligned}(-h)(g_0 + g_1) &= -h(g_0 + g_1)\\ &= -(h(g_0) + h(g_1))\\ &= -h(g_0) - h(g_1)\\ &= (-h)(g_0) + (-h)(g_1).\end{aligned}$$

2. $(\mathcal{H}(\boldsymbol{G}), \circ, \mathrm{I})$ é submonoide.

   2.1. Para todos $h_0, h_1 \in \mathcal{H}(\boldsymbol{G})$ e $g_0, g_1 \in G$,

$$\begin{aligned}(h_0 \circ h_1)(g_0 + g_1) &= h_0(h_1(g_0 + g_1))\\ &= h_0(h_1(g_0) + h_1(g_1))\\ &= h_0(h_1(g_0)) + h_0(h_1(g_1))\\ &= (h_0 \circ h_1)(g_0) + h_0 \circ h_1)(g_1);\end{aligned}$$

   2.2. Para todos $g_0, g_1 \in G$,

$$\mathrm{I}(g_0 + g_1) = g_0 + g_1 = \mathrm{I}(g_0) + \mathrm{I}(g_1) \blacksquare$$

$:\vdash$ **Definição 7.78.** Sejam $\boldsymbol{A}$ um anel e $\boldsymbol{G}$ um grupo comutativo. Uma *ação* de $\boldsymbol{A}$ sobre $\boldsymbol{G}$ é um homomorfismo de anel

$$\begin{aligned}\cdot\colon A &\longrightarrow \mathcal{H}(\boldsymbol{G})\\ a &\longmapsto a\cdot\colon G \longrightarrow G\\ &\qquad\qquad g \longmapsto a \cdot g.\end{aligned}$$

Denota-se $\cdot\colon \boldsymbol{A} \curvearrowright \boldsymbol{G}$. Diz-se que o anel $\boldsymbol{A}$ *age* no grupo $\boldsymbol{G}$ e denota-se $\boldsymbol{A} \curvearrowright \boldsymbol{G}$ se, e somente se, existe ação de $\boldsymbol{A}$ em $\boldsymbol{G}$.

[2] No sentido em que a multiplicação $\circ$ não precisa ser comutativa.

Se denotamos $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ e $\boldsymbol{G} = (G, +, -, \mathbf{0})$, essa definição é equivalente às seguintes propriedades.

1. Para todo $a \in A$, a função $a\cdot\colon G \longrightarrow G$ é um homomorfismo de grupo:
   1.1. Para todos $a \in A$ e $g, g' \in G$,
   $$a \cdot (g + g') = a \cdot g + a \cdot g'.$$
2. $\cdot\colon \boldsymbol{A}^{+} \longrightarrow (\mathcal{H}(\boldsymbol{G}), +, -, \mathbf{0})$ é homomorfismo de grupo:
   2.1. Para todos $a, a' \in A$ e $g \in G$,
   $$(a + a') \cdot m = a \cdot g + a' \cdot g;$$
3. $\cdot\colon \boldsymbol{A}^{\times} \longrightarrow (\mathcal{H}(\boldsymbol{G}), \circ, \mathrm{I})$ é homomorfismo de monoide:
   3.1. Para todos $a, a' \in A$ e $g \in G$,
   $$(a \times a') \cdot g = a \cdot (a' \cdot g);$$
   3.2. Para todo $g \in G$,
   $$1 \cdot g = g.$$

Capítulo 8

Módulos

8.1 Módulos e submódulos

:⊢ **Definição 8.1.** Seja $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ um anel. Um *módulo* sobre $\boldsymbol{A}$ é uma lista $\boldsymbol{M} = (M, \boldsymbol{+}, \boldsymbol{-}, \boldsymbol{0}, \boldsymbol{\cdot})$ em que $\boldsymbol{M}^+ := (M, \boldsymbol{+}, \boldsymbol{-}, \boldsymbol{0})$ é um grupo comutativo e $\boldsymbol{\cdot}\colon \boldsymbol{A} \curvearrowright \boldsymbol{M}^+$ é uma ação de anel ($\boldsymbol{\cdot}\colon \boldsymbol{A} \longrightarrow \boldsymbol{\mathcal{H}}(\boldsymbol{M}^+)$ é um homomorfismo de anel); ou seja,

1. Para todos $a \in A$ e $m, m' \in M$,
$$a \cdot (m + m') = a \cdot m + a \cdot m'.$$
2. Para todos $a, a' \in A$ e $m \in M$,
$$(a + a') \cdot m = a \cdot m + a' \cdot m;$$
3. Para todos $a, a' \in A$ e $m \in M$,
$$(a \times a') \cdot m = a \cdot (a' \cdot m);$$
4. Para todo $m \in M$,
$$1 \cdot m = m.$$

Os símbolos '$\times$' da multiplicação de $\boldsymbol{A}$ e '$\boldsymbol{\cdot}$' da ação de $\boldsymbol{A}$ sobre $\boldsymbol{M}$ serão suprimidos (e parênteses desnecessários relacionados a elas também), e os símbolos '$\boldsymbol{+}$', '$\boldsymbol{-}$' e '$\boldsymbol{0}$' das operações de $\boldsymbol{M}$ não serão diferenciados em notação dos símbolos '$+$', '$-$' e '0' das operações de $\boldsymbol{A}$.

Na definição usamos o fato de que $\boldsymbol{\mathcal{H}}(\boldsymbol{M}^+) = (\mathcal{H}(\boldsymbol{M}^+), +, -, 0, \circ, \mathrm{I})$ é um anel com as operações puxadas para o espaço de funções, a soma pontual sendo a soma do anel e a composição de função sendo o produto. Isso foi feito em §7.5.

▶ **Exemplo 8.1.** *Sejam* $\boldsymbol{A} = (A, +, -, 0, \times, 1)$ *um anel e* $I \trianglelefteq A$ *um ideal. Então* $\boldsymbol{I} = (I, +, -, 0, \cdot)$ *é um módulo sobre* $\boldsymbol{A}$*, em que* $\boldsymbol{I}^+ = (I, +, -, 0)$ *é o subgrupo aditivo de* $\boldsymbol{A}$ *e*

$$\begin{aligned} \cdot\colon A &\longrightarrow \mathcal{H}(I) \\ a &\longmapsto a\cdot\colon I \longrightarrow I \\ & \qquad\qquad i \longmapsto ai \end{aligned}$$

é a ação multiplicativa induzida pela multiplicação $\times$ *de* $\boldsymbol{A}$*. Note que, para cada* $a \in A$*, a função* $a\cdot\colon I \longrightarrow I$ *está bem definida pois* I *é um ideal, o que implica que, para todo* $i \in I$*,* $ai \in I$*.*

▶ **Exemplo 8.2.** *Seja* $\boldsymbol{G} = (G, +, -, 0)$ *um grupo comutativo. Então*

$$(G, +, -, 0, \cdot)$$

é um módulo sobre $\mathbb{Z}$*, em que*

$$\begin{aligned} \cdot\colon \mathbb{Z} &\longrightarrow \mathcal{H}(G) \\ n &\longmapsto n\cdot\colon G \longrightarrow G \\ & \qquad\qquad g \longmapsto ng = \mathop{+}_{i\in[n]} g. \end{aligned}$$

▷ **Exercício 8.1.** *Seja* $\boldsymbol{M}$ *um módulo sobre um anel* $\boldsymbol{A}$*.*

1. *Para todo* $a \in A$*,* $a0 = 0$*;*
2. *Para todo* $m \in M$*,* $0m = 0$*;*
3. *Para todos* $a \in A$ *e* $m \in M$*,* $-(am) = (-a)m = a(-m)$*.*

:⊢ **Definição 8.2.** Seja $\boldsymbol{M}$ um módulo sobre um anel $\boldsymbol{A}$. Um submódulo de $\boldsymbol{M}$ sobre $\boldsymbol{A}$ é uma lista $\boldsymbol{S} = (S, +, -, 0, \cdot)$ tal que

1. $\boldsymbol{S}^+ := (S, +, -, 0)$ é um subgrupo de $\boldsymbol{M}^+$;
2. Para todos $a \in A$ e $m \in S$, $am \in S$.

8.2 Homomorfismo de módulo

:⊢ **Definição 8.3.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ módulos sobre um anel $\boldsymbol{A}$. Um *homomorfismo de módulo* de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é uma função $h\colon M \longrightarrow M'$ tal que

1. (Aditividade) A função $h\colon \boldsymbol{M}^+ \longrightarrow \boldsymbol{M'}^+$ é um homomorfismo de grupo: para todos $m_0, m_1 \in M$,

$$h(m_0 + m_1) = h(m_0) + h(m_1);$$

2. (Homogeneidade) Para todos $a \in A$, $m \in M$,

$$h(am) = ah(m).$$

Denota-se $h\colon \boldsymbol{M} \longrightarrow \boldsymbol{M'}$. O conjunto dos homomorfismos de módulo de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é denotado $\mathcal{H}(\boldsymbol{M}, \boldsymbol{M'})$. Um *endomorfismo* de módulo de $\boldsymbol{M}$ é um homomorfismo de módulo de $\boldsymbol{M}$ para $\boldsymbol{M}$. O conjunto dos endomorfismos de módulo de $\boldsymbol{M}$ é denotado $\mathcal{H}(\boldsymbol{M})$.

▷ **Exercício 8.2** (Composição de homomorfismos)**.** *Sejam $\boldsymbol{M}$, $\boldsymbol{M'}$ e $\boldsymbol{M''}$ módulos sobre um anel $\boldsymbol{A}$, $h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M'})$ e $h' \in \mathcal{H}(\boldsymbol{M'}, \boldsymbol{M''})$ homomorfismos. Então $h' \circ h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M''})$.*

8.2.1 Módulo de homomorfismos de módulo

Por definição de homomorfismo de módulo, temos que

$$\mathcal{H}(\boldsymbol{M}, \boldsymbol{M'}) \subseteq \mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M'}^+),$$

ou seja, que todo homomorfismo de módulo é em particular um homomorfismo de grupo entre as estruturas de grupo $\boldsymbol{M}^+$ e $\boldsymbol{M'}^+$ de seus respectivos módulos. O conjunto de homomorfismos de grupo $\mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M'}^+)$ é um grupo com a adição, subtração e função nula puxadas das respectivas operações de $\boldsymbol{M'}^+$. Dotado dessa estrutura de grupo, o conjunto $\mathcal{H}(\boldsymbol{M}, \boldsymbol{M'})$ é de fato um subgrupo de $\mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M'}^+)$, que denotamos por

$$\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M'})^+ := (\mathcal{H}(\boldsymbol{M}, \boldsymbol{M'}), +, -, 0).$$

⊢ **Proposição 8.1.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ módulos sobre um anel $\boldsymbol{A}$.*

$$\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M'})^+ \leq \boldsymbol{\mathcal{H}}(\boldsymbol{M}^+, \boldsymbol{M'}^+).$$

□ *Demonstração.* 1. (Não vacuidade) A função nula

$$\begin{aligned} 0\colon M &\longrightarrow M' \\ m &\longmapsto 0 \end{aligned}$$

é um homomorfismo de módulo de $\boldsymbol{M}$ para $\boldsymbol{M'}$:

1.1. (Aditividade) Segue direto de $0 \in \mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M'}^+)$;
1.2. (Homogeneidade) Para todos $a \in A$, $m \in M$,

$$h(am) = 0 = a0 = ah(m).$$

2. (Fechamento) Para todos $h, h' \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M'})$, $h + h' \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M'})$:

2.1. (Aditividade) Segue direto de $h + h' \in \mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M}'^+)$;
2.2. (Homogeneidade) Para todos $a \in A$, $m \in M$,

$$\begin{aligned}(h + h')(am) &= h(am) + h'(am)\\ &= ah(m) + ah'(m)\\ &= a(h(m) + h'(m))\\ &= a(h + h')(m).\end{aligned}$$

3. (Invertibilidade) Para todo $h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$, $-h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$:
   3.1. (Aditividade) Segue direto de $-h \in \mathcal{H}(\boldsymbol{M}^+, \boldsymbol{M}'^+)$;
   3.2. (Homogeneidade) Para todos $a \in A$, $m \in M$,

$$(-h)(am) = -h(am) = -(ah(m)) = a(-h(m)).$$

■

:⊢ **Definição 8.4.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ módulos sobre um anel $\boldsymbol{A}$. O *módulo de homomorfismos* de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é a lista

$$\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}') := (\mathcal{H}(\boldsymbol{M}, \boldsymbol{M}'), +, -, 0, \cdot)$$

em que

1. $\mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$ é o conjunto dos homomorfismo de módulo de $\boldsymbol{M}$ para $\boldsymbol{M}'$;
2. $\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}')^+ = (\mathcal{H}(\boldsymbol{M}, \boldsymbol{M}'), +, -, 0)$ é o grupo de homomorfismos de módulo, com as operações definidas pontualmente, induzidas das operações de $\boldsymbol{M}'$;
3. $\cdot$ é a ação induzida

$$\begin{aligned}\cdot\colon A &\longrightarrow \mathcal{H}(\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}')^+)\\ a &\longmapsto a\cdot\colon \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}') \longrightarrow \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')\\ &\qquad\qquad h \longmapsto a \cdot h\colon M \longrightarrow M'\\ &\qquad\qquad\qquad\qquad g \longmapsto a \cdot h(g).\end{aligned}$$

Note que a ação é induzida pela ação de $\boldsymbol{M}'$, que é a ação que aparece na expressão $a \cdot h(g)$. Na expressão $\mathcal{H}(\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}')^+)$, o '$\boldsymbol{\mathcal{H}}$' interno se refere ao grupo de homomorfismos do módulo, subgrupo do grupo de homomorfismos do grupo comutativo $\boldsymbol{M}^+$ para o grupo comutativo $\boldsymbol{M}'^+$, enquanto o '$\mathcal{H}$' externo se refere ao conjunto de endomorfismos do grupo $\mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$.

⊢ **Proposição 8.2.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ módulos sobre um anel $\boldsymbol{A}$. O módulo de homomorfismos de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é um módulo sobre $\boldsymbol{A}$.*

□ *Demonstração.* Foi provado em 8.1 que $\boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}')^+$ é grupo comutativo, pois é subgrupo de $\boldsymbol{\mathcal{H}}(\boldsymbol{M}^+, \boldsymbol{M}'^+)$.

Devemos provar que $\cdot\colon \boldsymbol{A} \curvearrowright \boldsymbol{\mathcal{H}}(\boldsymbol{M}, \boldsymbol{M}')^+$ é ação de anel:

1. Para todos $a \in A$, $h, h' \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$ e $m \in M$,

$$\begin{aligned}(a \cdot (h + h'))(m) &= a \cdot (h + h')(m) \\ &= a \cdot (h(m) + h'(m)) \\ &= a \cdot h(m) + a \cdot h'(m) \\ &= (a \cdot h + a \cdot h')(m);\end{aligned}$$

2. Para todos $a, a' \in A$, $h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$ e $m \in M$,

$$\begin{aligned}((a + a') \cdot h)(m) &= (a + a') \cdot h(m) \\ &= a \cdot h(m) + a' \cdot h(m) \\ &= (a \cdot h + a' \cdot h)(m);\end{aligned}$$

3. Para todos $a, a' \in A$, $h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$ e $m \in M$,

$$\begin{aligned}((aa') \cdot h)(m) &= (aa') \cdot h(m) \\ &= a \cdot (a' \cdot h(m)) \\ &= a \cdot ((a' \cdot h)(m)) \\ &= (a \cdot (a' \cdot h))(m);\end{aligned}$$

4. Para todos $h \in \mathcal{H}(\boldsymbol{M}, \boldsymbol{M}')$ e $m \in M$,

$$(1 \cdot h)(m) = 1 \cdot h(m) = h(m).$$

■

Capítulo 9

Espaços lineares

9.1 Espaço e subespaço lineares

:⊢ **Definição 9.1.** Seja $\boldsymbol{C} = (C, +, -, 0, \times, 1)$ um corpo. Um *espaço linear* sobre $\boldsymbol{C}$ é um módulo $(\boldsymbol{V}, \cdot)$ sobre $\boldsymbol{C}$; ou seja,

1. Para todo $c \in C$, a função $c\cdot\colon \boldsymbol{V} \to \boldsymbol{V}$ é um homomorfismo de grupo:
   1.1. Para todos $v_0, v_1 \in M$,
   $$c \cdot (v_0 + v_1) = c \cdot v_0 + c \cdot v_1.$$
2. $\cdot\colon C \times V \to V$ é uma ação de corpo (ou de anel):
   2.1. Para todos $c_0, c_1 \in C$ e $v \in V$,
   $$(c_0 + c_1) \cdot v = c_0 \cdot v + c_1 \cdot v;$$
   2.2. Para todos $c_0, c_1 \in C$ e $v \in V$,
   $$(c_0 c_1) \cdot v = c_0 \cdot (c_1 \cdot v);$$
   2.3. Para todo $v \in V$,
   $$1 \cdot v = v;$$

Os símbolos da operação $\cdot$ de $\boldsymbol{C}$ e da ação $\cdot$ serão suprimidos (e parênteses desnecessários relacionados a elas também), e os símbolos das operações $+, -$ de $\boldsymbol{C}$ e $+, -$ de $\boldsymbol{V}$ não serão diferenciadas em notação. Um espaço linear será denotado como $\boldsymbol{V}$ quando não for relevante explicitar a ação $\cdot$.

⊢ **Proposição 9.1.** *Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. Para todos $\boldsymbol{v} \in V$ e $c \in C$,*

1. $c\boldsymbol{v} = \mathbf{0} \Leftrightarrow c = 0$ *ou* $\boldsymbol{v} = \mathbf{0}$;
2. $-(c\boldsymbol{v}) = (-c)\boldsymbol{v} = c(-\boldsymbol{v})$;
3. $c\boldsymbol{v} = (-c)(-\boldsymbol{v})$.

□ *Demonstração.* Sejam $\boldsymbol{v} \in V$ e $c \in C$.

1. Primeiro, notemos que

$$\begin{aligned} 0\boldsymbol{v} &= 0\boldsymbol{v} + \mathbf{0} \\ &= 0\boldsymbol{v} + (0\boldsymbol{v} - 0\boldsymbol{v}) \\ &= (0\boldsymbol{v} + 0\boldsymbol{v}) - 0\boldsymbol{v} \\ &= (0 + 0)\boldsymbol{v} - 0\boldsymbol{v} \\ &= 0\boldsymbol{v} - 0\boldsymbol{v} \\ &= \mathbf{0}. \end{aligned}$$

   Agora, notemos que

$$\begin{aligned} c\mathbf{0} &= c\mathbf{0} + \mathbf{0} \\ &= c\mathbf{0} + (c\mathbf{0} - c\mathbf{0}) \\ &= (c\mathbf{0} + c\mathbf{0}) - c\mathbf{0} \\ &= c(\mathbf{0} + \mathbf{0}) - c\mathbf{0} \\ &= c\mathbf{0} - c\mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

   Portanto, se $c = 0$ ou $\boldsymbol{v} = \mathbf{0}$, então $c\boldsymbol{v} = \mathbf{0}$. Agora, suponhamos que $c\boldsymbol{v} = \mathbf{0}$. Se $c \neq 0$, como $\boldsymbol{C}$ é corpo, segue da demonstração anterior que

$$\boldsymbol{v} = c^{-1}c\boldsymbol{v} = c^{-1}\mathbf{0} = \mathbf{0}.$$

2. Basta notar que

$$\begin{aligned} -(c\boldsymbol{v}) &= -(c\boldsymbol{v}) + \mathbf{0} \\ &= -(c\boldsymbol{v}) + (0\boldsymbol{v}) \\ &= -(c\boldsymbol{v}) + (c - c)\boldsymbol{v} \\ &= -(c\boldsymbol{v}) + (c\boldsymbol{v} + (-c)\boldsymbol{v}) \\ &= (-(c\boldsymbol{v}) + c\boldsymbol{v}) + (-c)\boldsymbol{v} \\ &= \mathbf{0} + (-c)\boldsymbol{v} \\ &= (-c)\boldsymbol{v} \end{aligned}$$

e que

$$\begin{aligned}
-(c\boldsymbol{v}) &= -(c\boldsymbol{v}) + \mathbf{0} \\
&= -(c\boldsymbol{v}) + (c\mathbf{0}) \\
&= -(c\boldsymbol{v}) + c(\boldsymbol{v} - \boldsymbol{v}) \\
&= -(c\boldsymbol{v}) + (c\boldsymbol{v} + c(-\boldsymbol{v})) \\
&= (-(c\boldsymbol{v} + c\boldsymbol{v}) + c(-\boldsymbol{v}) \\
&= \mathbf{0} + c(-\boldsymbol{v}) \\
&= c(-\boldsymbol{v}).
\end{aligned}$$

3. Do item anterior, segue que

$$c\boldsymbol{v} = (-(-c))\boldsymbol{v} = (-c)(-\boldsymbol{v}). \qquad \blacksquare$$

$\vdash$ **Proposição 9.2.** *Seja* $\boldsymbol{C}$ *um corpo e* n *um natural positivo. Então* $(C^n, +, \cdot)$, *em que*

$$\begin{aligned}
\cdot \colon C \times C^n &\longrightarrow C^n \\
(c, (c_1, \ldots, c_n)) &\longmapsto (c \cdot c_1, \ldots, c \cdot c_n),
\end{aligned}$$

é um espaço vetorial sobre $\boldsymbol{C}$.

$\square$ *Demonstração.* Claramente $(C^n, +)$ é um grupo comutativo com elemento neutro $(0, \ldots, 0)$. Note que, para todos $(c_1, \ldots, c_n) \in C^n$ e $c, c' \in C$,

$$1 \cdot (c_1, \ldots, c_n) = (1 \cdot c_1, \ldots, 1 \cdot c_n) = (c_1, \ldots, c_n)$$

e

$$\begin{aligned}
(c \cdot c') \cdot (c_1, \ldots, c_n) &= ((c \cdot c') \cdot c_1, \ldots, (c \cdot c') \cdot c_n) \\
&= ((c \cdot (c' \cdot c_1), \ldots, (c \cdot (c' \cdot c_n)) \\
&= c \cdot (c' \cdot c_1, \ldots, c' \cdot c_n) \\
&= c \cdot (c' \cdot (c_1, \ldots, c_n)).
\end{aligned}$$

Ainda, note que, para todos $(c_1, \ldots, c_n), (c'_1, \ldots, c'_n) \in C^n$ e $c, c' \in C$,

$$\begin{aligned}
c \cdot ((c_1, \ldots, c_n) + (c'_1, \ldots, c'_n)) &= c \cdot (c_1 + c'_1, \ldots, c_n + c'_n) \\
&= (c \cdot (c_1 + c'_1), \ldots, c \cdot (c_n + c'_n)) \\
&= (c \cdot c_1 + c \cdot c'_1, \ldots, c \cdot c_n + c \cdot c'_n) \\
&= (c \cdot c_1, \ldots, c \cdot c_n) + (c \cdot c'_1, \ldots, c \cdot c'_n) \\
&= c \cdot (c_1, \ldots, c_n) + c \cdot (c'_1, \ldots, c'_n)
\end{aligned}$$

e

$$\begin{aligned}(c+c')\cdot(c_1,\ldots,c_n) &= ((c+c')c_1,\ldots,(c+c')c_n)\\ &= (c\cdot c_1+c'\cdot c_1,\ldots,c\cdot c_n+c'\cdot c_n)_{i\in I}\\ &= (c\cdot c_1,\ldots,c\cdot c_n)+(c'\cdot c_1,\ldots,c'\cdot c_n)\\ &= c\cdot(c_1,\ldots,c_n)+c'\cdot(c_1,\ldots,c_n).\end{aligned}$$ ■

Para generalizar esse resultado, lembremos que o produto de uma família $(C_i)_{i\in I}$ de conjuntos é $\prod_{i\in I} C_i$ e, quando $C_i = C$, temos que $\prod_{i\in I} C_i = C^I$ e os elementos de C^I são funções $c = (c_i)_{i\in I}$ de I em C.

⊢ **Proposição 9.3.** *Sejam I um conjunto e $\boldsymbol{C}$ um corpo. Então $\boldsymbol{C^I} = (C^I, +, \cdot)$, em que*

$$\begin{aligned}+\colon C^I &\longrightarrow C^I\\ (\boldsymbol{c},\boldsymbol{c}') &\longmapsto (c_i+c_i')_{i\in I}\end{aligned}$$

e

$$\begin{aligned}\cdot\colon C\times C^I &\longrightarrow C^I\\ (a,\boldsymbol{c}) &\longmapsto (a\cdot c_i)_{i\in I}\end{aligned}$$

é um espaço vetorial sobre $\boldsymbol{C}$.

Esse exemplo pode ser ainda mais generalizado como a seguir.

⊢ **Proposição 9.4** (Espaço de Funções)**.** *Sejam X um conjunto e $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Então $\boldsymbol{V^X} = (V^X, +, \cdot)$, em que*

$$\begin{aligned}+\colon V^X\times V^X &\longrightarrow V^X\\ (\boldsymbol{f}_1,\boldsymbol{f}_2) &\longmapsto \boldsymbol{f}_1+\boldsymbol{f}_2\colon X\longrightarrow V\\ &\qquad\qquad x\longmapsto \boldsymbol{f}_1(x)+\boldsymbol{f}_2(x).\end{aligned}$$

e

$$\begin{aligned}\cdot\colon C\times V^X &\longrightarrow V^X\\ (c,\boldsymbol{f}) &\longmapsto c\boldsymbol{f}\colon X\longrightarrow V\\ &\qquad\qquad x\longmapsto c\boldsymbol{f}(x),\end{aligned}$$

é um espaço vetorial sobre $\boldsymbol{C}$.

□ *Demonstração.* Primeiro, sabemos que $(V^X, +)$ é um grupo comutativo com identidade $0\colon V^X \times V^X \to V^X$ definida por $0(x) = 0$. Devemos então mostrar que $\cdot\colon C \times V^X \to V^X$ satisfaz os itens da definição de espaço vetorial. Primeiro, seja $\boldsymbol{f} \in V^X$. Então, para todo $x \in X$, $(1\boldsymbol{f})(x) = 1\boldsymbol{f}(x) = \boldsymbol{f}(x)$, o que mostra que $1\boldsymbol{f} = \boldsymbol{f}$. Agora, sejam $c_1, c_2 \in C$ e $\boldsymbol{f} \in V^X$. Então, para todo $x \in X$,

$$((c_1c_2)\boldsymbol{f})(x) = (c_1c_2)\boldsymbol{f}(x) = c_1(c_2\boldsymbol{f}(x)) = c_1(c_2\boldsymbol{f})(x) = (c_1(c_2\boldsymbol{f}))(x),$$

o que mostra que $(c_1c_2)\boldsymbol{f} = c_1(c_2\boldsymbol{f})$.

Por fim, devemos mostrar as propriedades distributivas. Sejam $c \in C$ e $\boldsymbol{f}_1, \boldsymbol{f}_2 \in V^X$. Então, para todo $x \in X$,

$$\begin{aligned}(c(\boldsymbol{f}_1 + \boldsymbol{f}_2))(x) &= c(\boldsymbol{f}_1 + \boldsymbol{f}_2)(x)\\ &= c(\boldsymbol{f}_1(x) + \boldsymbol{f}_2(x))\\ &= c\boldsymbol{f}_1(x) + c\boldsymbol{f}_2(x)\\ &= (c\boldsymbol{f}_1)(x) + (c\boldsymbol{f}_2)(x)\\ &= (c\boldsymbol{f}_1 + c\boldsymbol{f}_2)(x),\end{aligned}$$

o que mostra que $c(\boldsymbol{f}_1 + \boldsymbol{f}_2) = c\boldsymbol{f}_1 + c\boldsymbol{f}_2$. Agora, sejam $c_1, c_2 \in C$ e $\boldsymbol{f} \in V^X$. Então, para todo $x \in X$,

$$\begin{aligned}((c_1 + c_2)\boldsymbol{f})(x) &= (c_1 + c_2)\boldsymbol{f}(x)\\ &= c_1\boldsymbol{f}(x) + c_2\boldsymbol{f}(x)\\ &= (c_1\boldsymbol{f})(x) + (c_2\boldsymbol{f})(x)\\ &= (c_1\boldsymbol{f} + c_2\boldsymbol{f})(x),\end{aligned}$$

o que mostra que $(c_1 + c_2)\boldsymbol{f} = c_1\boldsymbol{f} + c_2\boldsymbol{f}$. Assim, concluímos que $(V^X, +, \cdot)$ é um espaço vetorial sobre $\boldsymbol{C}$. ■

:⊢ **Definição 9.2.** Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Um *subespaço vetorial* de $\boldsymbol{V}$ é um conjunto não vazio $W \subseteq V$ tal que

1. $\forall \boldsymbol{w}_1, \boldsymbol{w}_2 \in W \qquad \boldsymbol{w}_1 + \boldsymbol{w}_2 \in W$;
2. $\forall c \in C\ \forall \boldsymbol{w} \in W \qquad c\boldsymbol{w} \in W$.

⊢ **Proposição 9.5.** *Seja $\boldsymbol{V} = (V, +, \cdot)$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Então um conjunto não vazio $W \subseteq V$ é um subespaço vetorial de $\boldsymbol{V}$ se, e somente se, $\boldsymbol{W} = (W, +|_{W\times W}, \cdot|_{C\times W})$ é um espaço vetorial sobre $\boldsymbol{C}$.*

⊢ **Proposição 9.6.** *Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e W um subespaço vetorial de $\boldsymbol{V}$. Então*

1. $\boldsymbol{0} \in W$;

2. *$\{\mathbf{0}\}$ e V são subespaços vetoriais de $\boldsymbol{V}$.*

□ *Demonstração.* 1. Como W não é vazio, seja $\boldsymbol{w} \in W$. Então $0\boldsymbol{w} = \mathbf{0} \in W$.

2. Seja $W = \{\mathbf{0}\}$. Se $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, $\boldsymbol{w}_1 = \mathbf{0}$ e $\boldsymbol{w}_2 = \mathbf{0}$, e segue que $\boldsymbol{w}_1 + \boldsymbol{w}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Ainda, para todo $c \in C$, segue que $c\boldsymbol{w}_1 = c\mathbf{0} = \mathbf{0} \in W$. Seja $W = V$. Como $\boldsymbol{V}$ é espaço vetorial, então V é subespaço vetorial de $\boldsymbol{V}$ pala proposição anterior. ■

⊢ **Proposição 9.7.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $(W_i)_{i\in I}$ uma família de subespaços vetoriais de $\boldsymbol{V}$. Então*

$$W := \bigcap_{i\in I} W_i$$

é um subespaço vetorial de $\boldsymbol{V}$.

□ *Demonstração.* Como, para todo $i \in I$, $\mathbf{0} \in W_i$, segue que $\mathbf{0} \in W$ e, portanto, W não é vazio. Agora, sejam $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$ e $c \in C$. Então, para todo $i \in I$, $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W_i$ e, como W_i é subespaço vetorial de $\boldsymbol{V}$, segue que $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W_i$ e que $c\boldsymbol{w}_1 \in W_i$. Logo $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W$ e $c\boldsymbol{w}_1 \in W$, o que mostra que W é subespaço vetorial de $\boldsymbol{V}$. ■

⊢ **Proposição 9.8.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $\{W_i\}_{i\in I}$ uma cadeia de subespaços vetoriais de $\boldsymbol{V}$ (ou seja, para todos $I, j \in I$, $W_I \subseteq W_j$ ou $W_j \subseteq W_i$). Então*

$$W := \bigcup_{i\in I} W_i$$

é um subespaço vetorial de $\boldsymbol{V}$.

□ *Demonstração.* Como, para todo $i \in I$, $\mathbf{0} \in W_i$, pois W_i é subespaço vetorial de $\boldsymbol{V}$, segue que $\mathbf{0} \in W$ e, portanto, W não é vazio. Agora, sejam $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$. Então existem $i, j \in I$ tais que $\boldsymbol{w}_1 \in W_i$ e $\boldsymbol{w}_2 \in W_j$. Nesse caso, $W_i \subseteq W_j$ ou $W_j \subseteq W_i$. Sem perda de generalidade, suponhamos o primeiro caso. Então segue que $\boldsymbol{w}_1 \in W_j$ e, portanto, $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W_j$, o que mostra que $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W$. Agora, seja $c \in C$ e notemos que, como W_i é subespaço vetorial de $\boldsymbol{V}$, segue que $c\boldsymbol{w}_1 \in W_i$. Logo $c\boldsymbol{w}_1 \in W$, o que mostra que W é subespaço vetorial de $\boldsymbol{V}$. ■

:⊢ **Definição 9.3.** Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$, $W \subseteq V$ e $(W_i)_{i\in I}$ uma indexação do conjunto de todos subespaços vetoriais de $\boldsymbol{V}$ dos quais W é subconjunto. O *subespaço vetorial gerado por* W em $\boldsymbol{V}$ é o subespaço vetorial

$$\langle W \rangle := \bigcap_{i\in I} W_i.$$

Nesse caso, dizemos que W é um *conjunto gerador* de $\langle W \rangle$ ou que W gera $\langle W \rangle$.

$\vdash$ **Proposição 9.9.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Então $\langle\emptyset\rangle = \{\mathbf{0}\}$.*

$\square$ *Demonstração.* Como $\{\mathbf{0}\}$ é um subespaço vetorial de V e $\emptyset \subseteq \{\mathbf{0}\}$, segue que, se $\boldsymbol{v} \in \langle\emptyset\rangle$, então $\boldsymbol{v} \in \{\mathbf{0}\}$, o que implica $\boldsymbol{v} = \mathbf{0}$ e, portanto, que $\langle\emptyset\rangle = \{\mathbf{0}\}$. $\blacksquare$

9.2 Combinação linear de vetores

$:\vdash$ **Definição 9.4.** Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$ um conjunto finito tal que $W = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$. Uma *combinação linear* de W em $\boldsymbol{V}$ é um vetor $\boldsymbol{v} \in V$ tal que existem $c_1, \ldots, c_n \in C$ satisfazendo

$$\boldsymbol{v} = \mathop{\mathbin{+}}_{i=1}^{n} c_i \boldsymbol{w}_i.$$

Se W é um conjunto infinito, uma *combinação linear* de W é uma combinação linear de um subconjunto finito de W.

O vetor $\mathbf{0}$ é combinação linear de qualquer conjunto, pois é a soma vazia.

$\vdash$ **Teorema 9.10.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$ não vazio. Então $\langle W\rangle$ é o conjunto de todas as combinações lineares de W em $\boldsymbol{V}$.*

$\square$ *Demonstração.* Consideremos, primeiro, o caso em que $W = \emptyset$. Nesse caso, $\langle W\rangle = \{\mathbf{0}\}$, e a única combinação linear de W é a soma vazia $\mathbf{0}$, o que mostra a igualdade dos conjuntos.

Agora, assumamos que $W \neq \emptyset$ e seja $(W_j)_{j\in J}$ uma indexação do conjunto de todos subespaços de $\boldsymbol{V}$ que contêm W. Primeiro, mostraremos que uma combinação linear de W em $\boldsymbol{V}$ está em $\langle W\rangle$. Seja $\boldsymbol{v} := +_{i=1}^{n} c_i\boldsymbol{w}_i$ uma combinação linear de W em $\boldsymbol{V}$. Para todo $j \in J$, W_j é um subespaço vetorial de $\boldsymbol{V}$. Portanto, para todo $i \in [n]$, segue que $c_i\boldsymbol{w_i} \in W_j$ e, então, que $\boldsymbol{v} \in W_j$. Logo $\boldsymbol{v} \in \langle W\rangle$.

Reciprocamente, mostraremos que o conjunto de todas combinações lineares de W em $\boldsymbol{V}$ é um subespaço vetorial de $\boldsymbol{V}$. Primeiro, notemos que $\mathbf{0}$ é uma combinação linear de W, pois, para todo $\boldsymbol{w} \in W$, vale $\mathbf{0} = 0\boldsymbol{w}$. Agora, sejam $\boldsymbol{v}_1 = +_{i=1}^{n} c_i\boldsymbol{w}_i$ e $\boldsymbol{v}_2 = +_{i=1}^{m} c'_i\boldsymbol{w}'_i$ combinações lineares de W em $\boldsymbol{V}$ e $c \in C$. Então, se definirmos, para todo $i \in [m]$, $\boldsymbol{w}_{n+i} := \boldsymbol{w}'_i$ e $c_{n+i} := c'_i$ e, para todo $i \in [n]$, $\overline{c}_i := cc_i$, segue que

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 = \mathop{\mathbin{+}}_{i=1}^{n} c_i\boldsymbol{w}_i + \mathop{\mathbin{+}}_{i=1}^{m} c'_i\boldsymbol{w}'_i = \mathop{\mathbin{+}}_{i=1}^{n+m} c_i\boldsymbol{w}_i$$

e

$$c\boldsymbol{v}_1 = \mathop{\mathbin{+}}_{i=1}^{n} (cc_i)\boldsymbol{w}_i = \mathop{\mathbin{+}}_{i=1}^{n} \overline{c}_i\boldsymbol{w}_i$$

são combinações lineares de W em $\boldsymbol{V}$, o que implica que o conjunto de todas combinações lineares de W em $\boldsymbol{V}$ é um subespaço de $\boldsymbol{V}$. Assim, como $\langle W \rangle$ é subconjunto de todo conjunto que é subespaço vetorial de $\boldsymbol{V}$ contendo W, segue que o conjunto de todas combinações lineares de W em $\boldsymbol{V}$ é igual ao subespaço gerado por W. ■

⊢ **Proposição 9.11.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$, $W \subseteq V$ e $\boldsymbol{v} \in \langle W \rangle$. Então existem $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in W$ distintos e $c_1, \ldots, c_n \in C$ tais que*

$$\boldsymbol{v} = \mathop{\mathbin{+}}_{i=1}^{n} c_i \boldsymbol{w}_i.$$

□ *Demonstração.* Como $\boldsymbol{v} \in \langle W \rangle$, existem $\boldsymbol{w}'_1, \ldots, \boldsymbol{w}'_m \in W$ e $c'_1, \ldots, c'_m \in C$ tais que $\boldsymbol{v} = +_{i=1}^{m} c'_i \boldsymbol{w}'_i$. Vamos particionar o conjunto dos índices $[m]$ com a seguinte relação de equivalência: para todo $i, j \in [m]$, $i \sim j$ se, e somente se, $\boldsymbol{w}'_i = \boldsymbol{w}'_j$. Essa relação é de equivalência pois a igualdade de vetores é uma relação de equivalência. Agora, seja n o número de classes de equivalências dessa relação. Para cada $i \in [n]$, seja $j \in P_i$ e definimos os vetores $\boldsymbol{w}_i := \boldsymbol{w}'_j$. Notemos que os vetores $\boldsymbol{w}_i$ estão bem definidos, não dependem do j, pois, se $k \in P_i$, então $\boldsymbol{w}_i = \boldsymbol{w}'_j = \boldsymbol{w}'_k$. Ainda, definimos os coeficientes $c_i := +_{j \in P_i} c'_j$. Desse modo, segue que

$$\boldsymbol{v} = \mathop{\mathbin{+}}_{i=1}^{m} c'_i \boldsymbol{w}'_i = \mathop{\mathbin{+}}_{i=1}^{n} \mathop{\mathbin{+}}_{j \in P_i} c'_j \boldsymbol{w}'_j = \mathop{\mathbin{+}}_{i=1}^{n} \mathop{\mathbin{+}}_{j \in P_i} c'_j \boldsymbol{w}_i = \mathop{\mathbin{+}}_{i=1}^{n} c_i \boldsymbol{w}_i.$$

Por fim, notemos que os $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ são distintos por definição, já que, se $\boldsymbol{w}_i = \boldsymbol{w}_j$ para $i, j \in [n]$, então existem $k, l \in [m]$ tais que $k \in P_i$, $l \in P_j$ e $\boldsymbol{w}_i = \boldsymbol{w}'_k$, $\boldsymbol{w}_j = \boldsymbol{w}'_l$. Mas isso implica $\boldsymbol{w}'_k = \boldsymbol{w}'_l$, o que implica $P_i = P_j$ e, portanto, $i = j$. ■

:⊢ **Definição 9.5** (Dependência Linear)**.** Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$. Dizemos que W é *linearmente dependente* em $\boldsymbol{V}$ se existem $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in W$ distintos e $c_1, \ldots, c_n \in C$ não nulos tais que

$$\boldsymbol{0} = \mathop{\mathbin{+}}_{i=1}^{n} c_i \boldsymbol{w}_i.$$

Caso contrário, dizemos que W é *linearmente independente* em $\boldsymbol{V}$.

⊢ **Proposição 9.12.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$. Então W é linearmente dependente se, e somente se, existe $\boldsymbol{w} \in W$ que é combinação linear de $W \smallsetminus \{\boldsymbol{w}\}$ em $\boldsymbol{V}$.*

□ *Demonstração.* Suponhamos que W é linearmente dependente. Então existem vetores $\boldsymbol{w}'_1, \ldots, \boldsymbol{w}'_n \in W$ distintos e $c'_1, \ldots, c'_n \in C$ não nulos tais que

$$\boldsymbol{0} = \mathop{\mathbin{+}}_{i=1}^{n} c'_i \boldsymbol{w}'_i.$$

Como $c'_1, \ldots, c'_n$ são não nulos, então existe $j \in [n]$ tal que $c'_j \neq 0$. Definindo $\boldsymbol{w}_i := \boldsymbol{w}'_i$ se $1 \leq i < j$ e $\boldsymbol{w}_i := \boldsymbol{w}'_{i-1}$ se $j < i \leq n$, e $c_i := (c'_j)^{-1}(-c'_i)$ para todo $1 \leq i < j$ ou $j < i \leq n$, segue que

$$\boldsymbol{w}'_j = \mathop{\bigplus}_{i=1}^{j-1} (c'_j)^{-1}(-c'_i)\boldsymbol{w}'_i + \mathop{\bigplus}_{i=j+1}^{n} (c'_j)^{-1}(-c'_i)\boldsymbol{w}'_i = \mathop{\bigplus}_{i=1}^{n-1} c_i\boldsymbol{w}_i.$$

Por tanto, como $\boldsymbol{w}_i \in W \smallsetminus \{\boldsymbol{w}'_j\}$ e $c_i \in C$ para todo $i \in [n-1]$, $\boldsymbol{w}'_j$ é combinação linear de $W \smallsetminus \{\boldsymbol{w}'_j\}$ em $\boldsymbol{V}$.

Reciprocamente, suponhamos que existe $\boldsymbol{w} \in W$ que é combinação linear de $W \smallsetminus \{\boldsymbol{w}\}$ em $\boldsymbol{V}$. Então existem $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in W \smallsetminus \{\boldsymbol{w}\}$ distintos e $c_1, \ldots, c_n \in C$ tais que

$$\boldsymbol{w} = \mathop{\bigplus}_{i=1}^{n} c_i\boldsymbol{w}_i.$$

Definindo $\boldsymbol{w}_{n+1} := \boldsymbol{w}$ e $c_{n+1} := -1$, segue que

$$\boldsymbol{0} = \mathop{\bigplus}_{i=1}^{n} c_i\boldsymbol{w}_i - \boldsymbol{w} = \mathop{\bigplus}_{i=1}^{n+1} c_i\boldsymbol{w}_i.$$

Então, como $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{n+1}$ são distintos e $c_{n+1} = -1 \neq 0$, segue que W é linearmente dependente. ■

⊢ **Proposição 9.13.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$. Então*

1. *$\emptyset$ é linearmente independente em $\boldsymbol{V}$;*
2. *Se $\boldsymbol{0} \in W$, então W é linearmente dependente em $\boldsymbol{V}$;*
3. *Se $W = \{\boldsymbol{v}\} \neq \{\boldsymbol{0}\}$, então W é linearmente independente em $\boldsymbol{V}$.*
4. *Sejam $\boldsymbol{v}, \boldsymbol{v}' \neq 0$. $W = \{\boldsymbol{v}, \boldsymbol{v}'\}$ é linearmente dependente se, e somente se, existe $c \in C \smallsetminus \{0\}$ tal que $\boldsymbol{v}' = c\boldsymbol{v}$.*

□ *Demonstração.*

1. Suponha, por absurdo, que $\emptyset$ não é linearmente independente. Então existem $\boldsymbol{w_1}, \ldots, \boldsymbol{w_n} \in \emptyset$ distintos e $c_1, \ldots, c_n \in C$ não nulos tais que
$$\boldsymbol{0} = \mathop{\bigplus}_{i=1}^{n} c_i\boldsymbol{w_i}.$$
Mas $\boldsymbol{w_1}, \ldots, \boldsymbol{w_n} \in \emptyset$ é um absurdo.
2. Seja $c \in C \smallsetminus \{0\}$. Então, como $\boldsymbol{0} = c\boldsymbol{0}$, segue que W é linearmente dependente em $\boldsymbol{V}$.
3. Se $\boldsymbol{0} = c\boldsymbol{v}$, como $\boldsymbol{v} \neq \boldsymbol{0}$, segue que $c = 0$, o que mostra que W é linearmente independente em $\boldsymbol{V}$.

4. Se W é linearmente dependente, então existem $c, c' \in C \setminus \{0\}$ tais que $0 = c\boldsymbol{v} + c'\boldsymbol{v}'$, o que implica que

$$\boldsymbol{v}' = -\frac{c}{c'}\boldsymbol{v}.$$

Reciprocamente, se existe $c \in C \setminus \{0\}$ tal que $\boldsymbol{v}' = c\boldsymbol{v}$, então

$$0 = -c\boldsymbol{v} + c\boldsymbol{v} = -c\boldsymbol{v} + \boldsymbol{v}',$$

o que mostra que W é linearmente dependente.

■

⊢ **Proposição 9.14.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W \subseteq V$. Então W é linearmente independente em $\boldsymbol{V}$ se, e somente se, para toda combinação linear $\boldsymbol{v} = +_{i=1}^{n} c_i\boldsymbol{w}_i \neq \boldsymbol{0}$ de W em $\boldsymbol{V}$ tal que $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ são distintos e não nulos, então $c_1, \ldots, c_n$ são únicos.*

□ *Demonstração.* Primeiro, suponhamos que W é linearmente dependente em $\boldsymbol{V}$. Então existem $\boldsymbol{w}'_1, \ldots, \boldsymbol{w}'_{n'} \in W$ distintos e $c'_1, \ldots, c'_{n'} \in C$ não nulos tais que

$$\boldsymbol{0} = \mathop{+}_{i=1}^{n'} c'_i\boldsymbol{w}'_i.$$

Nesse caso, seja $\boldsymbol{v} \in \langle W \rangle$. Se $\boldsymbol{v} = \boldsymbol{0}$, então segue que
..... . Se $\boldsymbol{v} \neq \boldsymbol{0}$

■

⊢ **Proposição 9.15.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $\{W_i\}_{i \in I}$ uma cadeia de conjuntos linearmente independentes em $\boldsymbol{V}$ (ou seja, para todos $i, j \in I$, $W_i \subseteq W_j$ ou $W_j \subseteq W_i$). Então*

$$W := \bigcup_{i \in I} W_i$$

é um conjunto linearmente independente em $\boldsymbol{V}$.

□ *Demonstração.* Como, para todo $i \in I$, $\boldsymbol{0} \in W_i$, segue que $\boldsymbol{0} \in W$ e, portanto, W não é vazio. Agora, sejam $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$. Então existem $i, j \in I$ tais que $\boldsymbol{w}_1 \in W_i$ e $\boldsymbol{w}_2 \in W_j$. Nesse caso, $W_i \subseteq W_j$ ou $W_j \subseteq W_i$. Sem perda de generalidade, suponhamos o primeiro caso. Então segue que $\boldsymbol{w}_1 \in W_j$ e, portanto, $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W_j$, o que mostra que $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W$. Agora, seja $c \in C$ e notemos que, como W_i é subespaço vetorial de $\boldsymbol{V}$, segue que $c\boldsymbol{w}_1 \in W_i$. Logo $c\boldsymbol{w}_1 \in W$, o que mostra que W é subespaço vetorial de $\boldsymbol{V}$. ■

9.3 Soma de subespaços vetoriais

:⊢ **Definição 9.6.** Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $(W_i)_{i\in I}$ uma família de subespaços vetoriais de $\boldsymbol{V}$. A *soma* de $(W_i)_{i\in I}$ é o subespaço vetorial gerado pela união de W_i. Denotamos

$$\mathop{+}_{i\in I} W_i := \left\langle \bigcup_{i\in I} W_i \right\rangle.$$

Caso $(W_i)_{i\in I}$ seja uma família finita, escrevemos $W_1 + \cdots + W_n$.

:⊢ **Definição 9.7.** Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Uma *soma direta* é a soma de uma família $(W_i)_{i\in I}$ de subespaços vetoriais de $\boldsymbol{V}$ tal que $W_i \cap W_j = \{\boldsymbol{0}\}$ para todo $i, j \in I$, $i \neq j$. Nesse caso, denotamos

$$\bigoplus_{i\in I} W_i := \left\langle \bigcup_{i\in I} W_i \right\rangle.$$

Caso $(W_i)_{i\in I}$ seja uma família finita, escrevemos $W_1 \oplus \cdots \oplus W_n$.

⊢ **Proposição 9.16.** *Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W_1, \ldots, W_n$ subespaços vetoriais de $\boldsymbol{V}$ tais que $V = +_{i=1}^{n} W_i$. Então*

$$V = \bigoplus_{i=1}^{n} W_i$$

se, e somente se, para todo $\boldsymbol{v} \in V$, existem únicos $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_n \in W_n$ tais que

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n} \boldsymbol{w}_i.$$

□ *Demonstração.* Mostraremos, primeiro, que se V é soma direta de $W_1, \ldots, W_n$, então todo vetor de V é soma única de vetores de $W_1, \ldots, W_n$. A demonstração será por indução em n. O caso base é trivial, pois, se $V = W_1$, então, para todo $\boldsymbol{v} \in V$, $\boldsymbol{v} \in W_1$. Agora, suponhamos que a proposição vale para todo natural menor ou igual a n. Sejam $W_1, \ldots, W_{n+1}$ subespaços vetoriais de V tais que $V = +_{i=1}^{n+1} W_i$. Então existem $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_{n+1} \in W_{n+1}$ tais que

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n+1} \boldsymbol{w}_i.$$

Suponhamos que existam $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_{n+1} \in W_{n+1}$ tais que

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n+1} \boldsymbol{w}'_i.$$

Então

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n+1} \boldsymbol{w}_i = \mathop{+}_{i=1}^{n+1} \boldsymbol{w}'_i,$$

o que implica

$$\mathop{+}_{i=1}^{n} (\boldsymbol{w}_i - \boldsymbol{w}'_i) = \boldsymbol{w}'_{n+1} - \boldsymbol{w}_{n+1}.$$

Como, para todo $i \in [n+1]$, $w_i, w'_i \in W_i$, segue que $\boldsymbol{w}_i - \boldsymbol{w}'_i \in W_i$. Definamos $W := \bigcup_{i=1}^{n} W_i$. Assim, segue que

$$\mathop{+}_{i=1}^{n} (\boldsymbol{w}_i - \boldsymbol{w}'_i) \in \langle W \rangle$$

e

$$\boldsymbol{w}'_{n+1} - \boldsymbol{w}_{n+1} \in W_{n+1}.$$

Ainda, como V é soma direta de $W_1, \ldots, W_{n+1}$, então segue que $W \cap W_{n+1} = \{\mathbf{0}\}$. Portanto concluímos que

$$\mathop{+}_{i=1}^{n} (\boldsymbol{w}_i - \boldsymbol{w}'_i) = \boldsymbol{w}'_{n+1} - \boldsymbol{w}_{n+1} = \mathbf{0}.$$

Assim, concluímos que $\boldsymbol{w}'_{n+1} = \boldsymbol{w}_{n+1}$ e que

$$\mathop{+}_{i=1}^{n} \boldsymbol{w}_i = \mathop{+}_{i=1}^{n} \boldsymbol{w}'_i.$$

Mas notemos que

$$\langle \boldsymbol{W} \rangle = (\langle W \rangle, +|_{\langle W \rangle \times \langle W \rangle}, \cdot|_{\langle W \rangle \times \langle W \rangle})$$

é um espaço vetorial e $W_1, \ldots, W_n$ são subespaços vetoriais de $\langle \boldsymbol{W} \rangle$ tais que $\langle W \rangle = \mathop{+}_{i=1}^{n} W_i$. Portanto, pela hipótese de indução, segue que, para todo $i \in [n]$, $\boldsymbol{w}_i = \boldsymbol{w}'_i$ e, portanto, concluímos que existem únicos $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_{n+1} \in W_{n+1}$ tais que

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n+1} \boldsymbol{w}_i.$$

Suponhamos, então, que todo vetor de V é soma de únicos vetores de W_1, …, W_n. Sejam $i, j \in [n]$, $i \neq j$, e $\boldsymbol{v} \in W_i \cap W_j$. Como $\boldsymbol{v} \in V$, segue que existem únicos $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_n \in W_n$ tais que

$$\boldsymbol{v} = \mathop{+}_{k=1}^{n} \boldsymbol{w}_k.$$

Sem perda de generalidade, suponhamos $i < j$. Notemos que

$$\boldsymbol{v} = \mathop{+}_{i=1}^{n} \boldsymbol{w}_i = \mathop{+}_{k=1}^{i-1} \boldsymbol{w}_k + (\boldsymbol{w}_i + \boldsymbol{v}) + \mathop{+}_{k=i+1}^{j-1} \boldsymbol{w}_k + (\boldsymbol{w}_j - \boldsymbol{v}) + \mathop{+}_{k=j+1}^{n} \boldsymbol{w}_k.$$

Como $\boldsymbol{v} \in W_i \cap W_j$, segue que $(\boldsymbol{w}_i + \boldsymbol{v}) \in W_i$ e $(\boldsymbol{w}_j - \boldsymbol{v}) \in W_j$ e, portanto, como $\boldsymbol{w}_1 \in W_1, \ldots, \boldsymbol{w}_n \in W_n$ são únicos, segue que $(\boldsymbol{w}_i + \boldsymbol{v}) = \boldsymbol{w}_i$ e $(\boldsymbol{w}_j - \boldsymbol{v}) = \boldsymbol{w}_j$; ou seja, $\boldsymbol{v} = \boldsymbol{0}$. Logo V é soma direta de $W_1, \ldots, W_n$. ■

9.4 Bases de espaços vetoriais

:⊢ **Definição 9.8.** Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Uma *base* de $\boldsymbol{V}$ é um conjunto $B \subseteq V$ linearmente independente em $\boldsymbol{V}$ que gera V; ou seja, $V = \langle B \rangle$. Uma base de um subespaço vetorial W de $\boldsymbol{V}$ é uma base do espaço vetorial $\boldsymbol{W} = (W, +|_{W \times W}, \cdot|_{W \times W})$.

⊢ **Teorema 9.17.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Então existe base B de $\boldsymbol{V}$ e, se L é um conjunto linearmente independente em $\boldsymbol{V}$, existe uma base B de $\boldsymbol{V}$ tal que $L \subseteq B$.*

□ *Demonstração.* A afirmação de que todo espaço vetorial tem uma base é consequência da segunda afirmação porque, tomando $L = \emptyset$, sabemos que L é linearmente independente e, portanto, existe base B de $\boldsymbol{V}$ que contém $\emptyset$. Demonstraremos a segunda afirmação.

Sejam L um conjunto linearmente independente em $\boldsymbol{V}$ e P o conjunto dos subconjuntos $S \subseteq V$ tais que $L \subseteq S$ e S é linearmente independente em $\boldsymbol{V}$. Então $(P, \subseteq)$ é um conjunto parcialmente ordenado com a contenção de conjuntos usual. Agora, seja $(S_i)_{i \in I}$ uma cadeia de $(P, \subseteq)$. Consideremos o conjunto $S := \bigcup_{i \in I} S_i$. Como $L \subseteq S_i$ para todo $i \in I$, então $L \subseteq S$. Devemos mostrar que S é um conjunto linearmente independente em $\boldsymbol{V}$. Para isso, seja $M \subseteq S$ subconjunto finito de S. Como $(S_i)_{i \in I}$ é uma cadeia, existe $i \in I$ tal que $M \subseteq S_i$. Mas, como S_i é linearmente independente, então M também o é e, portanto, S é linearmente independente. Logo S é um limitante superior da cadeia. Concluímos, portanto, que existe um elemento maximal B de $(P, \subseteq)$ que, por definição de P, é linearmente independente e $L \subseteq B$.

Vamos mostrar que B é base de $\boldsymbol{V}$. Devemos mostrar que B gera V, ou seja, que $V = \langle B \rangle$. Seja $\boldsymbol{v} \in V$ e suponhamos, por absurdo, que $\boldsymbol{v} \notin \langle B \rangle$. Então, em particular, $\boldsymbol{v} \notin B$; logo $B \subset B \cup \{\boldsymbol{v}\}$. Concluiremos que $B \cup \{\boldsymbol{v}\}$ é linearmente independente, o que contradiz a maximalidade de B. Seja S um subconjunto finito de $B \cup \{\boldsymbol{v}\}$. Se $\boldsymbol{v} \notin S$, então $S \subseteq B$ e, portanto, é linearmente independente, pois B o é; se $\boldsymbol{v} \in S$, sejam $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} := S \smallsetminus \{\boldsymbol{v}\} \subseteq B$ e $c, c_1, \ldots, c_n \in C$ tais que

$$c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n - c\boldsymbol{v} = \boldsymbol{0}.$$

Como $\boldsymbol{v} \notin \langle B \rangle$, então $c = 0$, pois, caso contrário, teríamos

$$\boldsymbol{v} = \frac{c_1}{c}\boldsymbol{v}_1 + \cdots + \frac{c_n}{c}\boldsymbol{v}_n.$$

Assim, segue que $c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n = \mathbf{0}$. Mas $S \setminus \{\boldsymbol{v}\} \subseteq B$ é linearmente independente, pois B o é, o que implica que $c_1 = \cdots = c_n = 0$ e, portanto, S é linearmente independente. Com isso, concluímos que $B \cup \{\boldsymbol{v}\}$ é linearmente independente, pois todo subconjunto finito é, e isso contradiz a maximalidade de B. Por esse absurdo, segue que $\boldsymbol{v} \in \langle B \rangle$ e, portanto, que $V = \langle B \rangle$. Concluímos que B é uma base de $\boldsymbol{V}$ que contém L. ■

⊢ **Proposição 9.18.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $W, W' \subseteq V$ conjuntos finitos. Se W é linearmente independente em $\boldsymbol{V}$ e W' gera V, então $\#(W) \leq \#(W')$.*

□ *Demonstração.* Se $W = \emptyset$, então $0 = \#(W') \leq \#(W)$. Caso contrário, seja $\#(W) = n$ e $(\boldsymbol{w}_i)_{i \in [n]}$ uma indexação de W. Suponhamos, por absurdo, que $W' = \emptyset$. Então, como W' gera V e $\langle W' \rangle = \{\mathbf{0}\}$, segue que $V = \{\mathbf{0}\}$, o que é absurdo, pois isso implica que $W = \{\mathbf{0}\}$, que é um conjunto linearmente dependente. Então $W' \neq \emptyset$. Seja $\#(W') = m$ e $(\boldsymbol{w}'_i)_{i \in [m]}$ uma indexação de W'. Queremos mostrar que $n \leq m$. Suponhamos, por absurdo, que $m < n$. Como W é linearmente independente, então, para todo $i \in [n]$, $\boldsymbol{w}_i \neq \mathbf{0}$. Como W' gera V, existem $c_1, \ldots, c_m \in C$ tais que

$$\boldsymbol{w}_1 = \mathop{+}_{i=1}^{m} c_i \boldsymbol{w}'_i,$$

e os $c_1, \ldots, c_m \in C$ não são todos nulos pois, caso contrário, teríamos $\boldsymbol{w}_1 = \mathbf{0}$. Assim, suponhamos, sem perda de generalidade, que $c_1 \neq 0$. Então

$$\boldsymbol{w}'_1 = c_1^{-1}\boldsymbol{w}_1 - \mathop{+}_{i=2}^{m} c_1^{-1} c_i \boldsymbol{w}'_i.$$

Seja $W_1 := \{\boldsymbol{w}_1, \boldsymbol{w}'_2, \ldots, \boldsymbol{w}'_m\}$. Como W' gera V e todo elemento de W' pode ser escrito como combinação linear de W_1, W_1 gera V. Assim, analogamente, podemos escrever $\boldsymbol{w}_2$ como combinação linear de W_1, como $\boldsymbol{w}_2 \neq \mathbf{0}$, segue que nem todo os coeficientes da combinação linear são nulos. Mais ainda, se somente o coeficiente de $\boldsymbol{w}_1$ é não nulo, então $\boldsymbol{w}_2$ é múltiplo de $\boldsymbol{w}_1$, o que contradiz a independência linear de W. Portanto, deve existir um coeficiente dos $\boldsymbol{w}'_2, \ldots, \boldsymbol{w}'_m$ não nulo. Assim, sem perda de generalidade, suponhamos que o coeficiente de $\boldsymbol{w}'_2$ é não nulo. Então, como no caso anterior, $\boldsymbol{w}'_2$ pode ser escrito como combinação linear de $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}'_3, \ldots, \boldsymbol{w}'_m$ e segue que o conjunto $W_2 := \{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}'_3, \ldots, \boldsymbol{w}'_m\}$ gera V. Repetindo o processo, que termina porque $m < n$ são finitos, achamos o conjunto $W_m := \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$, que gera V e é um subconjunto próprio de W, pois

$m < n$. Mas isso implica que $\boldsymbol{w}_{m+1} \in W$ é uma combinação linear de W_m em $\boldsymbol{V}$, o que implica que W é linearmente dependente, uma contradição. Logo $m \leq n$. ■

⊢ **Teorema 9.19.** *Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. Se $B, B' \subseteq V$ são bases de $\boldsymbol{V}$, então $\#(B) = \#(B')$.*

□ *Demonstração.* Primeiro, vamos mostrar que não ocorre o caso de uma base ser um conjunto finito e outra ser um conjunto infinito. Suponhamos, sem perda de generalidade, que B é um conjunto finito com $\#(B) = n$, $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\} := B$, e B' é um conjunto infinito. Seja $i \in [n]$. Como $\boldsymbol{b}_i \in V$ e B' gera V, segue que existem $\boldsymbol{b}_{i,1}, \ldots, \boldsymbol{b}_{i,n_i} \in B'$ e $c_{i,1}, \ldots, c_{i,n_i} \in C$ tais que

$$\boldsymbol{b}_i = \mathop{\mathbin{+}}_{j=1}^{n_i} c_{i,j}\boldsymbol{b}_{i,j}.$$

Notemos que o conjunto de todos esses $\boldsymbol{b}_{i,j}$ é $B'' := \bigcup_{i=1}^{n}\{\boldsymbol{b}_{i,j} : j \in [n_i]\}$, que é um subconjunto finito de B' e, portanto, um subconjunto próprio. Assim, como $B'' \subset B'$, existe $\boldsymbol{b} \in B' \smallsetminus B''$. Como $\boldsymbol{b} \in V$ e B é base, segue que existem $c_1, \ldots, c_n \in C$ tais que $\boldsymbol{b} = +_{i=1}^{n} c_i\boldsymbol{b}_i$. Mas então

$$\boldsymbol{b} = \mathop{\mathbin{+}}_{i=1}^{n} c_i\boldsymbol{b}_i = \mathop{\mathbin{+}}_{i=1}^{n} c_i \mathop{\mathbin{+}}_{j=1}^{n_i} c_{i,j}\boldsymbol{b}_{i,j} = \mathop{\mathbin{+}}_{i=1}^{n} \mathop{\mathbin{+}}_{j=1}^{n_i} c_i c_{i,j}\boldsymbol{b}_{i,j},$$

o que mostra que $\boldsymbol{b} \in B'$ pode ser escrito como uma combinação linear de $B' \smallsetminus \{\boldsymbol{b}\}$ em $\boldsymbol{V}$; ou seja, B' não é linearmente independente, o que é um absurdo. Assim, existem dois casos a serem considerados; o primeiro em que ambas as bases são conjuntos finitos e o outro em que ambas são conjuntos infinitos.

Suponhamos, no primeiro caso, que B e B' são conjuntos finitos com $\#(B) = n$ e $\#(B') = m$. Como B é linearmente independente e B' gera V, segue que $\#(B) \leq \#(B')$. Reciprocamente, como B' é linearmente independente e B gera V, segue que $\#(B') \leq \#(B)$. Assim, segue que $\#(B) = \#(B')$. Agora, suponhamos que B e B' são conjuntos infinitos.

TERMINAR ■

:⊢ **Definição 9.9.** Sejam $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$ e $B \subseteq V$ uma base $\boldsymbol{V}$. A *dimensão* de $\boldsymbol{V}$ é o número ordinal $\dim \boldsymbol{V} := \#(B)$.

9.5 Funções lineares

:⊢ **Definição 9.10.** Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$. Uma função *linear* de $\boldsymbol{V}$ para $\boldsymbol{W}$ é uma função $L\colon V \to W$ que satisfaz

1. (Aditividade) Para todos $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$,
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2);$$
2. (Homogeneidade) Para todos $c \in C, \boldsymbol{v} \in V$,
$$L(c\boldsymbol{v}) = cL(\boldsymbol{v}).$$

Denota-se $L\colon \boldsymbol{V} \to \boldsymbol{W}$. O conjunto das funções lineares de $\boldsymbol{V}$ para $\boldsymbol{W}$ é denotado $\mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$ e o conjunto das funções lineares de $\boldsymbol{V}$ para $\boldsymbol{V}$ é denotado $\mathcal{L}(\boldsymbol{V})$.

É imediato da definição que as duas propriedades são equivalentes à seguinte propriedade

1. (Linearidade) Para todos $c_1, c_2 \in C$ e $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$,
$$L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1L(\boldsymbol{v}_1) + c_2L(\boldsymbol{v}_2).$$

$\vdash$ **Proposição 9.20.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$ e $L \in \mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$. Então*

1. *(Linearidade generalizada) Para todos $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$ e $c_1, \ldots, c_n \in C$,*
$$L\left(\mathop{+}_{i=1}^{n} c_i\boldsymbol{v}_i\right) = \mathop{+}_{i=1}^{n} c_iL(\boldsymbol{v}_i).$$
2. *$L(\mathbf{0}) = \mathbf{0}$;*
3. *$L(-\boldsymbol{v}) = -L(\boldsymbol{v})$.*

$\vdash$ **Proposição 9.21.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$. O espaço linear $\mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$ é um subespaço linear de $\boldsymbol{W}^{\boldsymbol{V}}$.*

$\square$ *Demonstração.* Primeiro, sejam $L_1, L_2 \in \mathcal{L}(\boldsymbol{V}; \boldsymbol{W})$. Então, para todos $c \in C$ e $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$,
$$\begin{aligned}(L_1 + L_2)(\boldsymbol{v}_1 + c\boldsymbol{v}_2) &= L_1(\boldsymbol{v}_1 + c\boldsymbol{v}_2) + L_2(\boldsymbol{v}_1 + c\boldsymbol{v}_2)\\ &= L_1(\boldsymbol{v}_1) + cL_1(\boldsymbol{v}_2) + L_2(\boldsymbol{v}_1) + cL_2(\boldsymbol{v}_2)\\ &= L_1(\boldsymbol{v}_1) + L_2(\boldsymbol{v}_1) + cL_1(\boldsymbol{v}_2) + cL_2(\boldsymbol{v}_2)\\ &= (L_1 + L_2)(\boldsymbol{v}_1) + c(L_1 + L_2)(\boldsymbol{v}_2).\end{aligned}$$

Agora, sejam $c' \in C$ e $L \in \mathcal{L}(\boldsymbol{V}; \boldsymbol{W})$. Então
$$\begin{aligned}(c'L)(\boldsymbol{v}_1 + c\boldsymbol{v}_2) &= c'L(\boldsymbol{v}_1 + c\boldsymbol{v}_2)\\ &= c'(L(\boldsymbol{v}_1) + cL(\boldsymbol{v}_2))\\ &= c'L(\boldsymbol{v}_1) + c'cL(\boldsymbol{v}_2)\\ &= (c'L)(\boldsymbol{v}_1) + c(c'L)(\boldsymbol{v}_2).\end{aligned}$$

Portanto concluímos que $\mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$ é um subespaço linear de $\boldsymbol{W}^{\boldsymbol{V}}$. $\blacksquare$

$\vdash$ **Proposição 9.22.** *Sejam $\boldsymbol{V}_1$, $\boldsymbol{V}_2$ e $\boldsymbol{V}_3$ espaços lineares sobre um corpo $\boldsymbol{C}$. Se $L_1 \in \mathcal{L}(\boldsymbol{V}_1, \boldsymbol{V}_2)$, $L_2 \in \mathcal{L}(\boldsymbol{V}_2, \boldsymbol{V}_3)$, então $L_2 \circ L_1 \in \mathcal{L}(\boldsymbol{V}_1, \boldsymbol{V}_3)$.*

$\square$ *Demonstração.* Sejam $c \in C$ e $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Então

$$\begin{aligned}(L_2 \circ L_1)(\boldsymbol{v}_1 + c\boldsymbol{v}_2) &= L_2(L_1(\boldsymbol{v}_1 + c\boldsymbol{v}_2)) \\ &= L_2(L_1(\boldsymbol{v}_1) + cL_1(\boldsymbol{v}_2)) \\ &= L_2(L_1(\boldsymbol{v}_1)) + cL_2(L_1(\boldsymbol{v}_2)) \\ &= (L_2 \circ L_1)(\boldsymbol{v}_1) + c(L_2 \circ L_1)(\boldsymbol{v}_2),\end{aligned}$$

o que mostra que $L_2 \circ L_1$ é linear. $\blacksquare$

$\vdash$ **Proposição 9.23.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$ e $L \in \mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$. Se L é invertível, então $L^{-1} \in \mathcal{L}(\boldsymbol{W}, \boldsymbol{V})$.*

$\square$ *Demonstração.* Seja $L \in \mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$. Se L é invertível, $L^{-1} \in V^W$ e, para todos $c \in C$ e $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, existem $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ tais que $L(\boldsymbol{v}_1) = \boldsymbol{w}_1$ e $L(\boldsymbol{v}_2) = \boldsymbol{w}_2$ e segue que

$$\begin{aligned}L^{-1}(\boldsymbol{w}_1 + c\boldsymbol{w}_2) &= L^{-1}(L(\boldsymbol{v}_1) + cL(\boldsymbol{v}_2)) \\ &= L^{-1}(L(\boldsymbol{v}_1 + c\boldsymbol{v}_2)) \\ &= \boldsymbol{v}_1 + c\boldsymbol{v}_2 \\ &= L^{-1}(\boldsymbol{w}_1) + cL^{-1}(\boldsymbol{w}_2),\end{aligned}$$

o que mostra que L^{-1} é linear. $\blacksquare$

$\vdash$ **Proposição 9.24.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$, $B_V = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ base de $\boldsymbol{V}$, $B_W = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ base de $\boldsymbol{W}$ e $L \in \mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$. Então, para todo $\boldsymbol{v} \in V$, existem únicos $c_1, \ldots, c_m \in C$ tais que*

$$L(\boldsymbol{v}) = \mathop{\mathbin{+}}_{i=1}^{m} c_j \boldsymbol{w}_j.$$

$\square$ *Demonstração.* Primeiro demonstraremos a existência. Sabemos que, como B_V é base de V, então existem únicos $a_1, \ldots, a_n \in C$ tais que $\boldsymbol{v} = +_{i=1}^{n} a_i \boldsymbol{v}_i$. Mas, como L é linear, então

$$L(\boldsymbol{v}) = L\left(\mathop{\mathbin{+}}_{i=1}^{n} a_i \boldsymbol{v}_i\right) = \mathop{\mathbin{+}}_{i=1}^{n} a_i L(\boldsymbol{v}_i).$$

Agora, como B_W é base de $\boldsymbol{W}$, para cada $i \in \{1, \ldots, n\}$ existem únicos

$$b_{i1}, \ldots, b_{im} \in C$$

tais que $L(\boldsymbol{v}_i) = +_{j=1}^{m} b_{ij}\boldsymbol{w}_j$. Assim, definindo $c_j := +_{i=1}^{n} a_i b_{ij}$, segue que

$$L(\boldsymbol{v}) = \mathop{+}_{i=1}^{n} a_i \left(\mathop{+}_{j=1}^{m} b_{ij}\boldsymbol{w}_j \right) = \mathop{+}_{j=1}^{m} \left(\mathop{+}_{i=1}^{n} a_i b_{ij} \right) \boldsymbol{w}_j = \mathop{+}_{j=1}^{m} c_j \boldsymbol{w}_j.$$

■

9.6 Produto e coproduto de espaços vetoriais

9.6.1 Produto

:⊢ **Definição 9.11.** Seja $(\boldsymbol{V_i})_{i\in I} = ((V_i, +_i, \cdot_i))_{i\in I}$ uma família de espaços vetoriais sobre um corpo $\boldsymbol{C}$. O *produto categórico* de $(\boldsymbol{V_i})_{i\in I}$ é a tripla

$$\prod_{i\in I} \boldsymbol{V_i} = (V, +, \cdot),$$

em que $V = \prod_{i\in I} V_i$,

$$\begin{aligned} +\colon V \times V &\longrightarrow V \\ (v_1, v_2) &\longmapsto ((v_1)_i +_i (v_2)_i)_{i\in I} \end{aligned}$$

e

$$\begin{aligned} \cdot\colon C \times V &\longrightarrow V \\ (c, v) &\longmapsto (c \cdot_i v_i)_{i\in I}. \end{aligned}$$

⊢ **Proposição 9.25.** *Seja $(\boldsymbol{V_i})_{i\in I} = ((V_i, +_i, \cdot_i))_{i\in I}$ uma família de espaços vetoriais sobre um corpo $\boldsymbol{C}$. Então $\prod_{i\in I} \boldsymbol{V_i} = (V, +, \cdot)$ é um espaço vetorial sobre $\boldsymbol{C}$.*

□ *Demonstração.*

1. $(V, +)$ é um grupo pois tem a mesma operação do produto de grupos (6.26).
2. Seja $v \in V$. Então
$$1v = (1v_i)_{i\in I} = (v_i)_{i\in I} = v.$$
Sejam $c_1, c_2 \in C$ e $v \in V$. Então
$$\begin{aligned} (c_1c_2)v &= ((c_1c_2)v_i)_{i\in I} \\ &= (c_1(c_2v_i))_{i\in I} \\ &= c_1(c_2v_i)_{i\in I} \\ &= c_1(c_2v). \end{aligned}$$

3. (Distributividades) Sejam $c \in C$ e $v, v' \in V$. Então

$$\begin{aligned} c(v+v') &= c(v_i + v'_i)_{i\in I} \\ &= (c(v_i + v'_i))_{i\in I} \\ &= (cv_i + cv'_i)_{i\in I} \\ &= (cv_i)_{i\in I} + (cv'_i)_{i\in I} \\ &= cv + cv'. \end{aligned}$$

Sejam $c_1, c_2 \in C$ e $v \in V$. Então

$$\begin{aligned} (c_1+c_2)v &= ((c_1+c_2)v_i)_{i\in I} \\ &= (c_1v_i + c_2v_i)_{i\in I} \\ &= (c_1v_i)_{i\in I} + (c_2v_i)_{i\in I} \\ &= c_1v + c_2v. \end{aligned}$$

■

⊢ **Proposição 9.26.** *Seja* $(\boldsymbol{V_i})_{i\in I} = ((V_i, +_i, \cdot_i))_{i\in I}$ *uma família de espaços vetoriais sobre um corpo* $\boldsymbol{C}$*. Para todo* $i \in I$*, a projeção canônica* $\pi_i : \prod_{i\in I} \boldsymbol{V_i} \to \boldsymbol{V_i}$ *é uma função linear.*

□ *Demonstração.* Sejam $c \in C$ e $v, v' \in \prod_{i\in I} V_i$. Então

$$\pi_i(v + cv') = \pi_i((v_i + cv'_i)_{i\in I}) = v_i + cv'_i = \pi_i(v) + c\pi_i(v').$$

■

⊢ **Proposição 9.27** (Propriedade Universal)**.** *Sejam* $(\boldsymbol{V_i})_{i\in I}$ *uma família de espaços vetoriais sobre um corpo* $\boldsymbol{C}$*,* $\boldsymbol{X}$ *um espaço vetorial sobre* $\boldsymbol{C}$ *e, para todo* $i \in I$*,* $L_i : \boldsymbol{X} \to \boldsymbol{V_i}$ *uma função linear. Então existe única função linear* $L : \boldsymbol{X} \to \prod_{i\in I} \boldsymbol{V_i}$ *tal que, para todo* $i \in I$*,* $\pi_i \circ L = L_i$ *(o diagrama comuta).*

$$\begin{tikzcd} & \prod\limits_{i\in I} \boldsymbol{V_i} \arrow[d, "\pi_i"] \\ \boldsymbol{X} \arrow[ur, dashed, "L"] \arrow[r, "L_i"'] & \boldsymbol{V_i} \end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} L\colon X &\longrightarrow \prod_{i\in I} V_i \\ x &\longmapsto (L_i(x))_{i\in I}. \end{aligned}$$

Da propriedade universal para conjuntos, L é a única função de X para $\times_{i\in I} \boldsymbol{V_i}$ tal que, para todo $i \in I$, $\pi_i \circ L = L_i$. Falta mostrar que L é função linear. Sejam $c \in C$ e $x_1, x_2 \in V$. Então

$$\begin{aligned} L(x_1 + cx_2) &= (L_i(x_1 + cx_2))_{i\in I} \\ &= (L_i(x_1) + cL_i(x_2))_{i\in I} \\ &= (L_i(x_1))_{i\in I} + c(L_i(x_2))_{i\in I} \\ &= L(x_1) + cL(x_2). \end{aligned}$$

■

9.6.2 Coproduto (soma)

:⊢ **Definição 9.12.** Seja $(\boldsymbol{V_i})_{i\in I} = ((V_i, +_i, \cdot_i))_{i\in I}$ uma família de espaços vetoriais. A *soma categórica* de $(\boldsymbol{V_i})_{i\in I}$ é

$$\bigsqcup_{i\in I} \boldsymbol{V_i} = (V, +, \cdot),$$

em que

$$V = \left\{ v = (v_i)_{i\in I} \in \bigsqcap_{i\in I} V_i \;\middle|\; \#\left(\operatorname{supp}(v)\right) < \#\left(\mathbb{N}\right) \right\},$$

$$\begin{aligned} +\colon V \times V &\longrightarrow V \\ (v, v') &\longmapsto (v_i +_i v'_i)_{i\in I} \end{aligned}$$

e

$$\begin{aligned} \cdot\colon C \times V &\longrightarrow V \\ (c, v) &\longmapsto (cv_i)_{i\in I}. \end{aligned}$$

Observe que, se $\#\left(I\right) < \#\left(\mathbb{N}\right)$, então $\bigsqcap_{i\in I} \boldsymbol{V_i} = \bigsqcup_{i\in I} \boldsymbol{V_i}$.

⊢ **Proposição 9.28** (Propriedade Universal)**.** *Sejam $(\boldsymbol{V_i})_{i\in I}$ uma família de espaços vetoriais sobre um corpo $\boldsymbol{C}$, $\boldsymbol{X}$ um espaço vetorial sobre $\boldsymbol{C}$ e, para todo $i \in I$, $L_i : \boldsymbol{V_i} \to \boldsymbol{X}$ uma função linear. Então existe única função linear $L : \boldsymbol{X} \to \bigsqcup_{i\in I} \boldsymbol{V_i}$ tal que, para todo $i \in I$, $L \circ \iota_i = L_i$ (o diagrama comuta).*

$$\begin{tikzcd} \boldsymbol{V_i} \arrow[r, "L_i"] \arrow[d, "\iota_i"'] & \boldsymbol{X} \\ \bigsqcup_{i\in I} \boldsymbol{V_i} \arrow[ur, dashed, "L"'] & \end{tikzcd}$$

O coproduto de espaços vetoriais é também chamado de *soma* ou *soma direta* e denotado

$$\bigoplus_{i\in I} \boldsymbol{V_i}.$$

9.7 Projeções lineares

:⊢ **Definição 9.13.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. Uma *projeção linear* em $\boldsymbol{V}$ é uma função linear $p\colon V \to V$ idempotente: $p^2 = p$.

⊢ **Proposição 9.29.** *Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$ e $p\colon V \to V$ uma projeção linear.*

1. $p|_{p(V)} = \mathrm{I}|_{p(V)}$;
2. $(\mathrm{I} - p)\colon V \to V$ *é uma projeção linear em* $\boldsymbol{V}$;
3. $(\mathrm{I} - p)^{-1}(0) = p(V)$ *e* $(\mathrm{I} - p)(V) = p^{-1}(0)$;
4. $V = p(V) \oplus p^{-1}(0) = p(V) \oplus (\mathrm{I} - p)(V)$;
5. *Para todo* $\lambda \in C \smallsetminus \{0, 1\}$,
$$(\lambda\mathrm{I} - p)^{-1} = \frac{1}{\lambda}\mathrm{I} + \frac{1}{\lambda(1-\lambda)}p$$
portanto $\mathrm{Esp}(p) \subseteq \{0, 1\}$;
6. p *é invertível se, e somente se,* $p = \mathrm{I}$.
7. *Se* $p \neq 0$, *seu polinômio mínimo é* $x^2 - x = x(x-1)$, *logo* p *é diagonalizável pois tem raízes distintas.*

□ *Demonstração.*

1. Seja $v \in p(V)$. Então existe $v' \in V$ tal que $v = p(v')$, logo
$$p(v) = p(p(v')) = p^2(v') = p(v') = v.$$
2. A função $\mathrm{I} - p$ é linear pois é uma diferença de funções lineares. Basta mostrar que ela é idempotente. Basta notar que
$$(\mathrm{I} - p)^2 = \mathrm{I} - p - p + p^2 = \mathrm{I} - p - p + p = \mathrm{I} - p.$$
3. Seja $v \in (\mathrm{I} - p)^{-1}(0)$. Então $(\mathrm{I} - p)(v) = 0$, logo $v = p(v)$, o que implica que $v \in p(V)$. Reciprocamente, seja $v \in p(V)$. Então do item 1 segue que $p(v) = v$, logo $(\mathrm{I} - p)(v) = v - p(v) = v - v = 0$.
A outra relação segue de $(\mathrm{I} - p)$ ser projeção e $p = \mathrm{I} - (\mathrm{I} - p)$.
4. Seja $v \in p(V) \cap p^{-1}(0)$. Então $p(v) = 0$ e $p(v) = v$, o que implica que $v = p(v) = 0$. Assim temos que $p(V) \cap p^{-1}(0) = \{0\}$.
Seja $v \in V$. Então $p(v) \in p(V)$ e $(\mathrm{I}-p)(v) \in p^{-1}(0)$ (pois $(\mathrm{I}-p)(V) = p^{-1}(0)$), logo $v = p(v) + (\mathrm{I} - p)(v)$, o que mostra que $V = p(V) + p^{-1}(V)$.

5. Note que
$$\frac{1}{1-\lambda} - \frac{1}{\lambda} - \frac{1}{\lambda(1-\lambda)} = \frac{\lambda - (\lambda - 1) - 1}{\lambda(1-\lambda)} = 0,$$
logo
$$\begin{aligned}(\lambda \mathrm{I} - p) \circ \left(\frac{1}{\lambda}\mathrm{I} + \frac{1}{\lambda(1-\lambda)}p\right) &= \mathrm{I} + \frac{1}{1-\lambda}p - \frac{1}{\lambda}p - \frac{1}{\lambda(1-\lambda)}p^2 \\ &= \mathrm{I} + \left(\frac{1}{1-\lambda} - \frac{1}{\lambda} - \frac{1}{\lambda(1-\lambda)}\right)p \\ &= \mathrm{I}.\end{aligned}$$
e
$$\left(\frac{1}{\lambda}\mathrm{I} + \frac{1}{\lambda(1-\lambda)}p\right) \circ (\lambda \mathrm{I} - p) = \mathrm{I} - \frac{1}{\lambda}p + \frac{1}{1-\lambda}p - \frac{1}{\lambda(1-\lambda)}p^2 = \mathrm{I}.$$
6. Se p é invertível, $V = p(V)$, logo
$$p = p|_V = p|_{p(V)} = \mathrm{I}|_{p(V)} = \mathrm{I}|_V = \mathrm{I}.$$
A recíproca é evidente.
7. Evidente.

■

9.8 Funções multilineares

:⊢ **Definição 9.14.** Sejam $\boldsymbol{V_0}, \cdots, \boldsymbol{V_{k-1}}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$ e $k \in \mathbb{N}$. Uma função *k-linear* de $(\boldsymbol{V_0}, \ldots, \boldsymbol{V_{k-1}})$ para $\boldsymbol{W}$ é uma função
$$L\colon V_0 \times \cdots \times V_{k-1} \to W$$
qua satisfaz

1. (Multilinearidade) Para todos $i \in [k]$ e
$$(v_0, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{k-1}) \in V_0 \times \cdots \times V_{i-1} \times V_{i+1} \times \cdots \times V_{k-1},$$
a função
$$\begin{aligned} L(v_0, \ldots, v_{i-1}, \cdot, v_{i+1}, \ldots, v_{k-1})\colon \boldsymbol{V_i} &\longrightarrow \boldsymbol{W} \\ v &\longmapsto L(v_0, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_{k-1}) \end{aligned}$$
é uma função linear.

O conjunto dessas funções é denotado

$$\mathcal{L}(V_0, \ldots, V_{k-1}; W)$$

e, quando todos os espaços $\boldsymbol{V_i}$ são iguais, denota-se

$$\mathcal{L}^k(V, W) := \mathcal{L}(\underbrace{V, \ldots, V}_{k}; W)$$

$\vdash$ **Proposição 9.30.** *Sejam* $\boldsymbol{V_0}, \cdots, \boldsymbol{V_{k-1}}$ *e* $\boldsymbol{W}$ *espaços lineares sobre um corpo* $\boldsymbol{C}$ *e, para cada* $i \in [k]$, $d_i \in \mathbb{N}$ *a dimensão e* $b^{(i)} = (b_j^{(i)})_{j \in [d_i]}$ *uma base ordenada de* $\boldsymbol{V_i}$. *Toda função k-linear* $L \in \mathcal{L}(V_0, \ldots, V_{k-1}; W)$ *está determinada pelos seus valores em* $(b_{j_0}^{(0)}, \ldots, b_{j_{k-1}}^{(k-1)})_{(j_0, \ldots, j_{k-1}) \in [d_0] \times \cdots \times [d_{k-1}]}$.

$\square$ *Demonstração.* Sejam $v_0, \ldots, v_{m-1} \in V_k$, $c^0, \ldots, c^{m-1} \in C$, e, para todo $i \in [k] \smallsetminus \{k\}$, $v_i' \in V_i$. Como consequência da propriedade de linearidade generalizada para funções lineares,

$$L\left(v_0', \ldots, \mathop{+}_{i \in [m]} c^i v_i, \ldots, v_{k-1}'\right) = \mathop{+}_{i \in [m]} c^i L\left(v_0', \ldots, v_i, \ldots, v_{k-1}'\right).$$

Sendo assim, para cada $i \in [k]$, sejam $v_i \in V_i$ e $v_{(i)}^0, \ldots, v_{(i)}^{d_i} \in C$ os coeficientes de v_i na base $b^{(i)}$, de modo que

$$v_i = \mathop{+}_{j \in [d_i]} v_{(i)}^j b_j^{(i)}.$$

Pela linearidade em cada entrada, temos que

$$\begin{aligned}
L(v_0, \ldots, v_{k-1}) &= L\left(\mathop{+}_{j_0 \in [d_0]} v_{(0)}^{j_0} b_{j_0}^{(0)}, \ldots, \mathop{+}_{j_{k-1} \in [d_{k-1}]} v_{(k-1)}^{j_{k-1}} b_{j_{k-1}}^{(k-1)}\right) \\
&= \mathop{+}_{j_0 \in [d_0]} v_{(0)}^{j_0} L\left(b_{j_0}^{(0)}, \ldots, \mathop{+}_{j_{k-1} \in [d_{k-1}]} v_{(k-1)}^{j_{k-1}} b_{j_{k-1}}^{(k-1)}\right) \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
&= \mathop{+}_{j_0 \in [d_0]} \cdots \mathop{+}_{j_{k-1} \in [d_{k-1}]} v_{(0)}^{j_0} \cdots v_{(k-1)}^{j_{k-1}} L\left(b_{j_0}^{(0)}, \ldots, b_{j_{k-1}}^{(k-1)}\right) \\
&= \mathop{+}_{(j_0, \ldots, j_{k-1}) \in [d_0] \times \cdots \times [d_{k-1}]} v_{(0)}^{j_0} \cdots v_{(k-1)}^{j_{k-1}} L\left(b_{j_0}^{(0)}, \ldots, b_{j_{k-1}}^{(k-1)}\right).
\end{aligned}$$

Portanto a função L está determinada pelos valores que tem nos elementos

$$(b_{j_0}^{(0)}, \ldots, b_{j_{k-1}}^{(k-1)})_{(j_0, \ldots, j_{k-1}) \in [d_0] \times \cdots \times [d_{k-1}]}.$$

$\blacksquare$

$\vdash$ **Proposição 9.31.** *Sejam* $\boldsymbol{V_0}, \cdots, \boldsymbol{V_{k-1}}$ *e* $\boldsymbol{W}$ *espaços lineares sobre um corpo* $\boldsymbol{C}$*. Então*

$$\mathcal{L}(\boldsymbol{V_0}, \ldots, \boldsymbol{V_{k-1}}; \boldsymbol{W}) := (\mathcal{L}(V_0, \ldots, V_{k-1}; W), +, \cdot),$$

em que $+$ *e* $\cdot$ *são a soma e o produto escalar pontuais induzidos por* $\boldsymbol{W}$*, é um espaço linear sobre* $\boldsymbol{C}$*.*

9.8.1 Simetria, antissimetria e alternância

$:\vdash$ **Definição 9.15.** Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$. Uma função k-linear de $\boldsymbol{V}$ para $\boldsymbol{W}$

1. *simétrica* é uma função $f \in \mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$ tal que, para toda permutação $p \in \mathfrak{S}_k$ e todos v_0, …, $v_{k-1} \in V$,

$$f(v_{p(0)}, \ldots, v_{p(k-1)}) = f(v_0, \ldots, v_{k-1});$$

   O conjunto das funções k-lineares simétricas é denotado $\mathcal{S}^k(\boldsymbol{V}, \boldsymbol{W})$.
2. *antissimétrica* é uma função $f \in \mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$ tal que, para toda permutação $p \in \mathfrak{S}_k$ e todos $v_0, \ldots, v_{k-1} \in V$,

$$f(v_{p(0)}, \ldots, v_{p(k-1)}) = \pm(p) f(v_0, \ldots, v_{k-1});$$

3. *alternada* é uma função $f \in \mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$ tal que, para todos $v_0, \ldots, v_{k-1} \in V$ linearmente dependentes,

$$f(v_0, \ldots, v_{k-1}) = 0.$$

   O conjunto das funções k-lineares alternadas é denotado $\mathcal{A}^k(\boldsymbol{V}, \boldsymbol{W})$.

$\vdash$ **Proposição 9.32.** *Sejam* $\boldsymbol{V}$ *e* $\boldsymbol{W}$ *espaços lineares sobre um corpo* $\boldsymbol{C}$ *e* $f \in \mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$*.*

1. *A função* f *é alternada se, e somente se, para todos* v_0, …, $v_{k-1} \in V$ *tais que existem* $i, j \in [k]$ *distintos satisfazendo* $v_i = v_j$*,*

$$f(v_0, \ldots, v_{k-1}) = 0.$$

2. *Se* f *é alternada, então é antissimétrica. Se* $\operatorname{car}(\boldsymbol{C}) \neq 2$ *e* f *é antissimétrica, então é alternada.*
3. *Se* $\operatorname{car}(\boldsymbol{C}) = 2$*, então* f *é antissimétrica se, e somente se, é simétrica.*

□ *Demonstração.* 1. Se f é alternada, então, para todos $v_0, \ldots, v_{k-1} \in V$ tais que $v_i = v_j$ para dois $i, j \in [k]$ distintos, o conjunto $\{v_0, \ldots, v_{k-1}\}$ é linearmente dependente, portanto $f(v_0, \ldots, v_{k-1}) = 0$. Reciprocamente, suponha que f satisfaz a propriedade e sejam $v_0, \ldots, v_{k-1} \in V$ linearmente dependentes. Então existe $i \in [k]$ tal que v_i é combinação linear dos outros v_j: existem $c_j \in C$, com $j \in [k] \smallsetminus \{i\}$, tais que

$$v_i = \mathop{+}_{j \in [k] \smallsetminus \{i\}} c_j v_j.$$

Assim, segue da k-linearidade e da propriedade enunciada que

$$\begin{aligned} f(v_0, \ldots, v_{k-1}) &= f\left(v_0, \ldots, \mathop{+}_{j \in [k] \smallsetminus \{i\}} c_j v_j, \ldots, v_{k-1}\right) \\ &= \mathop{+}_{j \in [k] \smallsetminus \{i\}} c_j f(v_0, \ldots, v_j, \ldots, v_{k-1}) \\ &= \mathop{+}_{j \in [k] \smallsetminus \{i\}} c_j 0 = 0. \end{aligned}$$

2. Suponha f alternada e sejam $v_0, \ldots, v_{k-1} \in V$. Então segue da k-linearidade e da alternância de f que

$$\begin{aligned} 0 =& f(v_0, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_{k-1}) \\ =& f(v_0, \ldots, v_i, \ldots, v_i, \ldots, v_{k-1}) + f(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_{k-1}) \\ &+ f(v_0, \ldots, v_j, \ldots, v_i, \ldots, v_{k-1}) + f(v_0, \ldots, v_j, \ldots, v_j, \ldots, v_{k-1}) \\ =& f(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_{k-1}) + f(v_0, \ldots, v_j, \ldots, v_i, \ldots, v_{k-1}), \end{aligned}$$

portanto

$$f(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_{k-1}) = -f(v_0, \ldots, v_j, \ldots, v_i, \ldots, v_{k-1}).$$

Como toda permutação $p \in \mathfrak{S}_k$ é um produto de $N \in \mathbb{N}$ inversões, e como $\pm(p) = (-1)^N$, segue por indução que, para toda permutação $p \in \mathfrak{S}_k$,

$$f(v_{p(0)}, \ldots, v_{p(k-1)}) = (-1)^N f(v_0, \ldots, v_{k-1}) = \pm(p) f(v_0, \ldots, v_{k-1}).$$

Suponha que $\operatorname{car}(\boldsymbol{C}) \neq 2$. Sejam $f \in \mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$ antissimétrica e v_0, …, $v_{k-1} \in V$ tais que $v_i = v_j$ para dois $i, j \in [k]$ distintos. Considerando a permutação $(i \quad j) \in \mathfrak{S}_k$, segue da antissimetria de f e de $\pm((i \quad j)) = -1$ que

$$\begin{aligned} f(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_{k-1}) &= f(v_0, \ldots, v_j, \ldots, v_i, \ldots, v_{k-1}) \\ &= -f(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_{k-1}), \end{aligned}$$

portanto
$$2f(v_0, \ldots, v_{k-1}) = 0.$$
Como $\mathrm{car}(\boldsymbol{C}) \neq 2$, segue que $f(v_0, \ldots, v_{k-1}) = 0$. Do item 1 segue que f é alternada.

3. Se $\mathrm{car}(\boldsymbol{C}) = 2$, então $-1 = 1$. Isso implica que, para qualquer permutação p, $\pm(p) = 1$.

■

Um exemplo de uma função multilinear que é antissimétrica mas não é alternada em para um corpo de característica 2 é o seguinte. Seja $\mathbb{Z}_2 = \{0, 1\}$ o corpo de característica 2 e considere a função $f\colon \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2$ dada por

$$\begin{aligned} f\colon \mathbb{Z}_2 \times \mathbb{Z}_2 &\longrightarrow \mathbb{Z}_2 \\ (0,0) &\longmapsto 0 \\ (0,1) &\mapsto 0 \\ (1,0) &\mapsto 0 \\ (1,1) &\mapsto 1 \end{aligned}$$

Pode-se verificar que essa função é bilinear e antissimétrica, mas não é alternada porque $f(1,1) = 1 \neq 0$.

De modo mais geral, para qualquer função $p\colon [k] \to [k]$, não necessariamente bijetiva, considerando o caráter $\pm(p)$ que vale 0 se p não é uma permutação, e 1 ou -1 conforme a paridade da permutação p, as definições de formas antissimétricas e alternadas podem ser unificadas: a de formas antissimétricas pode ser mantida como foi feita e, para o caso de formas alternadas, devido à propriedade 1 da proposição anterior, pode-se enunciar: para qualquer $p\colon [k] \to [k]$,

$$f(v_{p(0)}, \ldots, v_{v_{k-1}}) = \pm(p) f(v_0, \ldots, v_{k-1}).$$

O caso em que p é bijetiva recai na definição de formas antissimétricas, e o caso em que p não é bijetiva recai na definição de formas alternadas, mais precisamente da propriedade 1 da proposição anterior, que é equivalente à definição de forma alternada.

⊢ **Proposição 9.33.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{W}$ espaços lineares sobre um corpo $\boldsymbol{C}$. Os espaços $\mathcal{S}^k(\boldsymbol{V}, \boldsymbol{W})$ e $\mathcal{A}^k(\boldsymbol{V}, \boldsymbol{W})$ são subespaços lineares de $\mathcal{L}^k(\boldsymbol{V}, \boldsymbol{W})$.*

9.8.2 Espaços de funções multilineares

⊢ **Proposição 9.34.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ e $\boldsymbol{V}''$ espaços lineares sobre um corpo $\boldsymbol{C}$. Os espaços são linearmente isomorfos.*

$$\mathcal{L}(\boldsymbol{V}, \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}'')) \simeq \mathcal{L}(\boldsymbol{V}, \boldsymbol{V}'; \boldsymbol{V}'').$$

□ *Demonstração.* Denotaremos uma função $f \in \mathcal{L}(\boldsymbol{V}, \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}''))$ por

$$\begin{aligned} f\colon V &\longrightarrow \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}'') \\ v &\longmapsto f_v\colon V' \longrightarrow V'' \\ & \qquad\qquad v' \longmapsto f_v(v'). \end{aligned}$$

Definimos as funções

$$\begin{aligned} I\colon \mathcal{L}(\boldsymbol{V}, \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}'')) &\longrightarrow \mathcal{L}(\boldsymbol{V}, \boldsymbol{V}'; \boldsymbol{V}'') \\ f &\longmapsto I(f)\colon V \times V' \longrightarrow V'' \\ & \qquad\qquad (v, v') \longmapsto f_v(v') \end{aligned}$$

e

$$\begin{aligned} J\colon \mathcal{L}(\boldsymbol{V}, \boldsymbol{V}'; \boldsymbol{V}'') &\longrightarrow \mathcal{L}(\boldsymbol{V}, \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}'')) \\ f &\longmapsto J(f)\colon V \longrightarrow \mathcal{L}(\boldsymbol{V}', \boldsymbol{V}'') \\ & \qquad\qquad v \longmapsto J(f)_v\colon V' \longrightarrow V'' \\ & \qquad\qquad\qquad\qquad v' \longmapsto f(v, v'). \end{aligned}$$

É simples mostrar que I e J estão bem definidas, são lineares e uma é inversa da outra. ■

9.9 Formas multilineares

:⊢ **Definição 9.16.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. Uma *forma k-linear* em $\boldsymbol{V}$ é uma função $f \in \mathcal{L}^k(\boldsymbol{V}; \boldsymbol{C})$, ou seja, um funcional k-linear em $(\boldsymbol{V}, \ldots, \boldsymbol{V})$. O conjunto das formas k-lineares simétricas é denotado $\mathcal{S}^k(\boldsymbol{V})$ e o das alternadas é denotado $\mathcal{A}^k(\boldsymbol{V})$.

Temos que $\mathcal{S}^1(\boldsymbol{V}) = \mathcal{A}^1(\boldsymbol{V}) = \mathcal{L}^k(\boldsymbol{V})$ e $\mathcal{S}^0(\boldsymbol{V}) = \mathcal{A}^0(\boldsymbol{V}) = \mathcal{L}^0(\boldsymbol{V}) = C$.

9.9.1 Produto tensorial de formas multilineares

:⊢ **Definição 9.17.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$, $f \in \mathcal{L}^k(V)$, $f' \in \mathcal{L}^{k'}(V)$ e $v_0, \ldots, v_{k+k'-1} \in V$. O *produto tensorial* de f e f' em $(v_0, \ldots, v_{k+k'-1})$ é

$$(f \otimes f')(v_0, \ldots, v_{k+k'-1}) := f(v_0, \ldots, v_{k-1}) f'(v_k, \ldots, v_{k+k'-1}).$$

Sejam $n \in \mathbb{N}$ e $(f_i)_{i \in [n]}$ formas multilineares em $\boldsymbol{V}$. O *produto tensorial* dessas formas é definido recursivamente

$$\bigotimes_{i \in [n]} f_i := \begin{cases} f_0, & n = 1 \\ \left(\bigotimes_{i \in [n-1]} f_i \right) \otimes f_{n-1}, & n > 1. \end{cases}$$

$\vdash$ **Proposição 9.35.** *Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$.*

1. *Para todas formas $f \in \mathcal{L}^k(\boldsymbol{V})$ e $f' \in \mathcal{L}^{k'}(\boldsymbol{V})$, a função $f \otimes f'$ é uma forma $k + k'$-linear.*
2. *(Bilinearidade) A função*

$$\begin{aligned} \otimes \colon \mathcal{L}^k(\boldsymbol{V}) \times \mathcal{L}^{k'}(\boldsymbol{V}) &\longrightarrow \mathcal{L}^{k+k'}(\boldsymbol{V}) \\ (f, f') &\longmapsto f \otimes f' \end{aligned}$$

   é bilinear.
3. *(Associatividade) Para todas formas $f \in \mathcal{L}^k(\boldsymbol{V})$, $f' \in \mathcal{L}^{k'}(\boldsymbol{V})$ e $f'' \in \mathcal{L}^{k''}(\boldsymbol{V})$,*

$$(f \otimes f') \otimes f'' = f \otimes (f' \otimes f'').$$

9.9.2 Produto alternado de formas multilineares

$:\vdash$ **Definição 9.18.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$, $f \in \mathcal{A}^k(\boldsymbol{V})$, $f' \in \mathcal{A}^{k'}(\boldsymbol{V})$ e $v_0, \ldots, v_{k+k'-1} \in \boldsymbol{V}$. O *produto alternado* (ou *exterior*) de f e f' em $(v_0, \ldots, v_{k+k'-1})$ é

$$(f \wedge f')(v_0, \ldots, v_{k+k'-1}) := \frac{1}{k!k'!} \mathop{+}_{p \in \mathfrak{S}_{k+k'}} \pm(p)(f \otimes f')(v_{p(i)})_{i \in [k+k']}.$$

Sejam $n \in \mathbb{N}$ e $(f_i)_{i \in [n]}$ formas multilineares em $\boldsymbol{V}$. O *produto alternado* dessas formas é definido recursivamente

$$\bigwedge_{i \in [n]} f_i := \begin{cases} f_0, & n = 1 \\ \left(\bigwedge_{i \in [n-1]} f_i \right) \wedge f_{n-1}, & n > 1. \end{cases}$$

$\vdash$ **Proposição 9.36.** *Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$.*

1. *Para todas formas $f \in \mathcal{L}^k(\boldsymbol{V})$ e $f' \in \mathcal{L}^{k'}(\boldsymbol{V})$, a função $f \wedge f'$ é uma forma $k + k'$-linear alternada.*
2. *(Bilinearidade) A função*

$$\begin{aligned} \wedge \colon \mathcal{A}^k(\boldsymbol{V}) \times \mathcal{A}^{k'}(\boldsymbol{V}) &\longrightarrow \mathcal{A}^{k+k'}(\boldsymbol{V}) \\ (f, f') &\longmapsto f \wedge f' \end{aligned}$$

   é bilinear.
3. *(Associatividade) Para todas formas alternadas $f \in \mathcal{A}^k(\boldsymbol{V})$, $f' \in \mathcal{A}^{k'}(\boldsymbol{V})$ e $f'' \in \mathcal{A}^{k''}(\boldsymbol{V})$,*

$$(f \wedge f') \wedge f'' = f \wedge (f' \wedge f'').$$

4. Para todas formas $f \in \mathcal{A}^k(\boldsymbol{V})$ e $f' \in \mathcal{A}^{k'}(\boldsymbol{V})$

$$f' \wedge f = (-1)^{kk'} f \wedge f'.$$

□ *Demonstração.* 1. Seja $v_0, \ldots, v_{k+k'-1} \in V$ e $\bar{p} \in \mathfrak{S}_{k+k'}$. Então

$$\begin{aligned}
(f \wedge f')(v_{\bar{p}(i)})_{i\in[k+k']} &= \frac{1}{k!k'!} \mathop{+}_{p\in\mathfrak{S}_{k+k'}} \pm(p)(f \otimes f')(v_{p(\bar{p}(i))})_{i\in[k+k']} \\
&= \frac{1}{k!k'!} \mathop{+}_{p\in\mathfrak{S}_{k+k'}} \pm(\bar{p}^{-1}p)(f \otimes f')(v_{p(i)})_{i\in[k+k']} \\
&= \frac{1}{k!k'!} \mathop{+}_{p\in\mathfrak{S}_{k+k'}} \pm(\bar{p}^{-1})\pm(p)(f \otimes f')(v_{p(i)})_{i\in[k+k']} \\
&= \pm(\bar{p}^{-1})\frac{1}{k!k'!} \mathop{+}_{p\in\mathfrak{S}_{k+k'}} \pm(p)(f \otimes f')(v_{p(i)})_{i\in[k+k']} \\
&= \pm(\bar{p})(f \wedge f')(v_i)_{i\in[k+k']}.
\end{aligned}$$

2.

■

⊢ **Proposição 9.37.** *Sejam $\boldsymbol{V}$ um espaço linear de dimensão finita d sobre um corpo $\boldsymbol{C}$, $(b_i)_{i\in[d]}$ um base ordenada de $\boldsymbol{V}$ e $({b_i}^*)_{i\in[d]}$ a base dual de $\boldsymbol{V}^*$. Então $\mathcal{A}^k(\boldsymbol{V})$ é um espaço linear sobre $\boldsymbol{C}$ de dimensão $\binom{d}{k}$ e o conjunto*

$$\left\{ b^*_{i_0} \wedge \cdots \wedge b^*_{i_{k-1}} \;\middle|\; i_0 < \cdots < i_{k-1} \in [d] \right\}$$

de formas alternadas k-lineares em $\boldsymbol{V}$ é uma base para $\mathcal{A}^k(\boldsymbol{V})$.

Em particular, isso mostra que formas d-lineares num espaço de dimensão d são todas múltiplos umas das outras, pois $\binom{d}{d} = 1$. Podemos fixar o valor de uma das formas como 1 na base canônica e chamá-la de *determinante*.

⊢ **Definição 9.19.** Seja $d \in \mathbb{N}$. O *determinante* em $\mathbb{R}^d$ é a forma d-linear

$$\det \cdot, \cdots, \cdot := {\mathrm{i}_0}^* \wedge \ldots \wedge {\mathrm{i}_{d-1}}^*.$$

Como comentado, essa é a única forma d-linear em $\mathbb{R}^d$ tal que

$$\det \mathrm{i}_0, \ldots, \mathrm{i}_{d-1} = 1.$$

Na seção seguinte, definiremos o conceito de determinante de modo mais geral e independente de base.

9.9.3 Formas puxadas e determinante

:⊢ **Definição 9.20.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ espaços lineares sobre um corpo $\boldsymbol{C}$, $L\colon V \to V'$ uma função linear e $f \in \mathcal{L}^k(V')$. A *forma k-linear puxada* por L de f é a função

$$\begin{aligned} L^* f\colon V^k &\longrightarrow C \\ (v_0, \ldots, v_{k-1}) &\longmapsto f(L(v_0), \ldots, L(v_{k-1})). \end{aligned}$$

A função *k-adjunta* induzida por L é a função linear

$$\begin{aligned} L^*\colon \mathcal{L}^k(V') &\longrightarrow \mathcal{L}^k(V) \\ f &\longmapsto L^* f. \end{aligned}$$

A notação é ambígua porque, para cada $k \in [d+1]$, a função L^* é uma função diferente. De fato, para $f \in \mathcal{L}^k(V')$,

$$L^* f = f \circ L^{\otimes k},$$

o que quer dizer que

$$L^* = (L^{\otimes k})^\star$$

⊢ **Proposição 9.38.** *Sejam $\boldsymbol{V}$, $\boldsymbol{V}'$ e $\boldsymbol{V}''$ espaços lineares sobre um corpo $\boldsymbol{C}$ e $L\colon V \to V'$ e $L'\colon V' \to V''$ funções lineares. Então*

$$(L' \circ L)^* = L^* \circ L'^*.$$

□ *Demonstração.* Seja $f \in \mathcal{L}^k(V)$ tal que $f \neq 0$. Então

$$\begin{aligned} (L' \circ L)^* f(v_0, \ldots, v_{k-1}) &= f((L' \circ L)(v_0), \ldots, (L' \circ L)(v_{k-1})) \\ &= f(L'(L(v_0)), \ldots, L'(L(v_{k-1}))) \\ &= L'^* f(L(v_0), \ldots, L(v_{k-1})) \\ &= L^*(L'^* f)(v_0, \ldots, v_{k-1}) \\ &= (L^* \circ L'^*) f(v_0, \ldots, v_{k-1}). \end{aligned}$$

■

Essas funções k-adjuntas estão também definidas se restringimos os espaços de funcionais $\mathcal{L}^k$ para espaços de funcionais alternados $\mathcal{A}^k$. Se considerarmos um espaço linear d-dimensional $\boldsymbol{V}$, o espaço $\mathcal{A}^d(\boldsymbol{V})$ é um espaço 1-dimensional, o que implica que todas funções lineares são multiplicações por escalares. Isso nos permite definir o determinante de uma função linear intrinsecamente como o escalar que representa seu d-adjunto.

:⊢ **Definição 9.21.** Sejam $\boldsymbol{V}$ um espaço linear d-dimensional sobre um um corpo $\boldsymbol{C}$, $L\colon V \to V$ uma função linear e $L^*\colon \mathcal{A}^d(V) \to \mathcal{A}^d(V)$. O *determinante* de L é o escalar $\det L \in C$ tal que, para toda forma $f \in \mathcal{A}^d(\boldsymbol{V})$,

$$L^* f = \det L f.$$

Isso nos dá por definição a igualdade

$$f(L(v_0), \ldots, L(v_{d-1})) = \det L f(v_0, \ldots, v_{d-1}).$$

⊢ **Proposição 9.39.** *Sejam $\boldsymbol{V}$ um espaço linear d-dimensional sobre um um corpo $\boldsymbol{C}$ e $L, L' \in \mathcal{L}(V, V)$. Então*

$$\det L' \circ L = \det L \det L'.$$

□ *Demonstração.* Seja $f \in \mathcal{A}^n(V)$ tal que $f \neq 0$. Então

$$\det L' \circ L f = (L' \circ L)^* f = (L^* \circ L'^*) f = \det L \det L' f.$$

■

⊢ **Proposição 9.40.** *Sejam $\boldsymbol{V}$ um espaço linear d-dimensional sobre um um corpo $\boldsymbol{C}$ e $L\colon V \to V$ uma função linear. Então L é invertível se, e somente se,* $\det L \neq 0$.

□ *Demonstração.* Basta notar que L é invertível se, e somente se, leva base em base e $f \in \mathcal{A}^d(V)$ é nula em um conjunto linearmente independente de vetores. ■

9.9.4 Formas bilineares simétricas e formas quadráticas

:⊢ **Definição 9.22.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. Uma *forma quadrática* em $\boldsymbol{V}$ é uma função $[\![\cdot]\!] : V \longrightarrow C$ que satisfaz

1. (2-Homogeneidade) Para todos $c \in C$ e $v \in V$, $[\![cv]\!] = c^2 [\![v]\!]$;
2. (Bilinearidade polar) A função

$$\begin{aligned} b\colon V \times V &\longrightarrow C \\ (v, v') &\longmapsto [\![v + v']\!] - [\![v]\!] - [\![v']\!] \end{aligned}$$

é uma forma bilinear.

Um *espaço quadrático* é um par $(\boldsymbol{V}, [\![\cdot]\!])$ em que $\boldsymbol{V}$ é um espaço linear (de dimensão finita) e $[\![\cdot]\!]$ uma forma quadrática em $\boldsymbol{V}$.

$\vdash$ **Proposição 9.41.** *Seja* $\boldsymbol{V}$ *um espaço linear sobre um corpo* $\boldsymbol{C}$ *de característica diferente de* 2.

1. *Se* $[\![\cdot]\!]$ *uma forma quadrática em* $\boldsymbol{V}$*, a função*

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon V \times V &\longrightarrow C \\ (v, v') &\longmapsto \frac{[\![v + v']\!] - [\![v]\!] - [\![v']\!]}{2} \end{aligned}$$

   é uma 2*-forma simétrica e, para todo* $v \in V$, $[\![v]\!] = \langle v, v \rangle$.
2. *Se* $\langle \cdot, \cdot \rangle$ *uma* 2*-forma simétrica em* $\boldsymbol{V}$*. A função*

$$\begin{aligned} [\![\cdot]\!] \colon V &\longrightarrow C \\ v &\longmapsto \langle v, v \rangle \end{aligned}$$

   é uma forma quadrática e, para todos $v, v' \in V$, $\langle v, v' \rangle = \frac{[\![v+v']\!]-[\![v]\!]-[\![v']\!]}{2}$.

$\square$ *Demonstração.* 1. Primeiro mostramos que $\langle \cdot, \cdot \rangle$ é bilinear simétrica.

- (Bilinearidade) Segue da bilinearidade de $b(v, v') = [\![v + v']\!] - [\![v]\!] - [\![v']\!]$ que $\langle \cdot, \cdot \rangle = \frac{1}{2} b$ é bilinear.
- (Simetria) Para todos $v, v' \in V$,

$$\langle v, v' \rangle = \frac{[\![v + v']\!] - [\![v]\!] - [\![v']\!]}{2} = \frac{[\![v' + v]\!] - [\![v']\!] - [\![v]\!]}{2} = \langle v', v \rangle .$$

Por fim, segue da 2-homogeneidade de $[\![\cdot]\!]$ que, para todo $v \in V$,

$$\begin{aligned} \langle v, v \rangle &= \frac{[\![v + v]\!] - [\![v]\!] - [\![v]\!]}{2} \\ &= \frac{[\![2v]\!] - 2\,[\![v]\!]}{2} \\ &= \frac{4\,[\![v]\!] - 2\,[\![v]\!]}{2} \\ &= \frac{2\,[\![v]\!]}{2} \\ &= [\![v]\!] . \end{aligned}$$

2. Primeiro mostramos que $[\![\cdot]\!]$ é forma quadrática.

- (2-Homogeneidade) Para todos $c \in C$ e $v \in V$, segue da bilinearidade de $\langle \cdot, \cdot \rangle$ que

$$[\![cv]\!] = \langle cv, cv \rangle = c^2 \langle v, v \rangle = c^2 [\![v]\!] .$$

- (Bilinearidade polar) Para todos $v, v' \in V$, segue da bilinearidade e simetria de $\langle \cdot, \cdot \rangle$ que

$$\begin{aligned} b(v, v') &= [\![v + v']\!] - [\![v]\!] - [\![v']\!] \\ &= \langle v + v', v + v' \rangle - \langle v, v \rangle - \langle v', v' \rangle \\ &= \langle v, v \rangle + \langle v, v' \rangle + \langle v', v \rangle + \langle v', v' \rangle - \langle v, v \rangle - \langle v', v' \rangle \\ &= \langle v, v' \rangle + \langle v, v' \rangle \\ &= 2 \langle v, v' \rangle , \end{aligned}$$

  portanto b é bilinear.

Por fim, segue da igualdade anterior que, para todos $v, v' \in V$,

$$\langle v, v' \rangle = \frac{[\![v + v']\!] - [\![v]\!] - [\![v']\!]}{2}. \qquad \blacksquare$$

9.9.4.1 2-Formas reais e os espaços $\mathbb{R}^{p,q}$

Definimos agora espaços reais munidos de formas bilineares simétricas não-degeneradas. Todas formas bilineares simétricas não-degeneradas em $\mathbb{R}^d$ são equivalentes a uma dessas formas.

:⊢ **Definição 9.23.** Sejam $p, q \in \mathbb{N}$. A 2-*forma simétrica canônica de assinatura* (p, q) é a 2-forma

$$\begin{aligned} \langle \cdot, \cdot \rangle_{p,q} \colon \mathbb{R}^{p+q} \times \mathbb{R}^{p+q} &\longrightarrow \mathbb{R} \\ (x, x') &\longmapsto \mathop{\mathbin{+}}_{k=0}^{p-1} x_k x'_k - \mathop{\mathbin{+}}_{k=p}^{p+q-1} x_k x'_k \end{aligned}$$

Para ver que essa função é bilinear, basta ver que na base canônica $(\mathrm{i}_k)_{k \in [p+q]}$, ela é dada por

$$\langle \mathrm{i}_k, \mathrm{i}_{k'} \rangle_{p,q} = \begin{cases} \delta_{k,k'} & 0 \leq k, k' \leq p - 1 \\ -\delta_{k,k'} & p \leq k, k' \leq p + q - 1 \\ 0 & \text{c.c.}. \end{cases}$$

e que a partir disso obtemos

$$
\begin{aligned}
\langle x, x' \rangle_{p,q} &= \left\langle \mathop{\mathbin{+}}_{k=0}^{p-1} x_k \mathrm{i}_k + \mathop{\mathbin{+}}_{k=p}^{p+q-1} x_k \mathrm{i}_k, \mathop{\mathbin{+}}_{k'=0}^{p-1} x_{k'} \mathrm{i}_{k'} + \mathop{\mathbin{+}}_{k'=p}^{p+q-1} x_{k'} \mathrm{i}_{k'} \right\rangle_{p,q} \\
&= \left\langle \mathop{\mathbin{+}}_{k=0}^{p-1} x_k \mathrm{i}_k, \mathop{\mathbin{+}}_{k'=0}^{p-1} x_{k'} \mathrm{i}_{k'} \right\rangle_{p,q} + \left\langle \mathop{\mathbin{+}}_{k=0}^{p-1} x_k \mathrm{i}_k, \mathop{\mathbin{+}}_{k'=p}^{p+q-1} x_{k'} \mathrm{i}_{k'} \right\rangle_{p,q} \\
&+ \left\langle \mathop{\mathbin{+}}_{k=p}^{p+q-1} x_k \mathrm{i}_k, \mathop{\mathbin{+}}_{k'=0}^{p-1} x_{k'} \mathrm{i}_{k'} \right\rangle_{p,q} + \left\langle \mathop{\mathbin{+}}_{k=p}^{p+q-1} x_k \mathrm{i}_k, \mathop{\mathbin{+}}_{k'=p}^{p+q-1} x_{k'} \mathrm{i}_{k'} \right\rangle_{p,q} \\
&= \mathop{\mathbin{+}}_{k=0}^{p-1} \mathop{\mathbin{+}}_{k'=0}^{p-1} x_k x_{k'} \left\langle \mathrm{i}_k, \mathrm{i}_{k'} \right\rangle_{p,q} + \mathop{\mathbin{+}}_{k=0}^{p-1} \mathop{\mathbin{+}}_{k'=p}^{p+q-1} x_k x_{k'} \left\langle \mathrm{i}_k, \mathrm{i}_{k'} \right\rangle_{p,q} \\
&+ \mathop{\mathbin{+}}_{k=p}^{p+q-1} \mathop{\mathbin{+}}_{k'=0}^{p-1} x_k x_{k'} \left\langle \mathrm{i}_k, \mathrm{i}_{k'} \right\rangle_{p,q} + \mathop{\mathbin{+}}_{k=p}^{p+q-1} \mathop{\mathbin{+}}_{k'=p}^{p+q-1} x_k x_{k'} \left\langle \mathrm{i}_k, \mathrm{i}_{k'} \right\rangle_{p,q} \\
&= \mathop{\mathbin{+}}_{k=0}^{p-1} x_k x'_k - \mathop{\mathbin{+}}_{k=p}^{p+q-1} x_k x'_k.
\end{aligned}
$$

A forma quadrática associada a $\langle \cdot, \cdot \rangle_{p,q}$ é a forma

$$
[\![x]\!]_{p,q} = \mathop{\mathbin{+}}_{k=0}^{p-1} {x_k}^2 - \mathop{\mathbin{+}}_{k=p}^{p+q-1} {x_k}^2.
$$

O espaço linear $\mathbb{R}^{p+q}$ dotado dessa forma é denotado $\mathbb{R}^{p,q}$.

9.10 Produto tensorial de espaços lineares

:⊢ **Definição 9.24.** Sejam X um conjunto, $\boldsymbol{C}$ um corpo e $f : X \to C$ uma função. O *suporte* de f é o conjunto

$$\operatorname{supp}(f) := f^{-1}\left(\{0\}^{\complement}\right) = \{x \in X \mid f(x) \neq 0\}.$$

:⊢ **Definição 9.25.** Sejam I um conjunto e $\boldsymbol{C}$ um corpo. O *espaço linear livre em* I sobre $\boldsymbol{C}$ é o conjunto

$$\mathcal{F}(I) := \left\{v \in C^I \;\middle|\; \#\left(\operatorname{supp}(v)\right) < \#\left(\mathbb{N}\right)\right\}.$$

Os elementos de $\mathcal{F}(I)$ são as *combinações lineares formais* de elementos de I sobre $\boldsymbol{C}$.

A *inclusão* de I em $\mathcal{F}(I)$ é a função

$$\begin{aligned} \iota\colon I &\longrightarrow \mathcal{F}(I) \\ v &\longmapsto \delta_v\colon I \longrightarrow C \\ & \qquad\qquad i \longmapsto \begin{cases} 1, & i = v \\ 0, & i \neq v. \end{cases} \end{aligned}$$

Denotaremos os elementos δ_v por v quando não houver necessidade de diferenciá-los.

Note que $\mathcal{F}(I) = \bigsqcup_{i \in I} \boldsymbol{C}$ e, portanto,

$$\boldsymbol{\mathcal{F}(I)} := (\mathcal{F}(I)), +, \cdot),$$

em que $+$ e $\cdot$ são a soma e o produto escalar pontuais induzidos por $\boldsymbol{W}$, é um espaço linear sobre $\boldsymbol{C}$ com base $\{\delta_v \mid v \in I\}$. Por isso, definido $c_i := v(i) \in C$, uma função $v : I \to C$ de suporte finito é uma soma

$$v = \mathop{\boldsymbol{+}}_{i \in I} c_i \delta_i = \mathop{\boldsymbol{+}}_{i \in I} c_i v_i,$$

que é uma soma finita porque $\operatorname{supp}(v)$ é finito, ou seja, somente uma quantidade finita dos c_i é não nulo.

⊢ **Proposição 9.42** (Propriedade Caracterítica dos Espaços Lineares Livres)**.** *Sejam I um conjunto, $\boldsymbol{C}$ um corpo e $\boldsymbol{V}$ um espaço linear sobre $\boldsymbol{C}$. Para toda função $f : I \to V$, existe uma única função linear $\bar{f} : \mathcal{F}(I) \to V$ tal que $\bar{f} \circ \iota = f$ (o diagrama comuta).*

$$\begin{tikzcd}
\boldsymbol{\mathcal{F}(I)} \arrow[dr, dashed, "\bar{f}"] & \\
I \arrow[u, "\iota"] \arrow[r, "f"'] & \boldsymbol{V}
\end{tikzcd}$$

□ *Demonstração.* Basta usar a propriedade característica de coproduto de espaços lineares, ou notar o que uma função $\bar{f}$ pode ser construída definindo seus valores em $\delta_v \in \mathcal{F}(I)$. Para todo $v = +_{i \in I} c_i v_i \in \mathcal{F}(I)$, define-se

$$\bar{f}\left(\mathop{+}_{v_i \in I} c_i v_i\right) := \mathop{+}_{v_i \in I} c_i f(v_i)$$

A função é única pois está definida da base δ_v. ■

:⊢ **Definição 9.26.** Sejam $\boldsymbol{C}$ um corpo e $V_0, \ldots, V_{n-1}$ conjuntos. Consideremos o conjunto $\mathcal{R}$ gerado por vetores de $\mathcal{F}(V_0 \times \cdots \times V_{n-1})$ da forma

$$(v_0, \ldots, v_i + v_i', \ldots, v_{n-1}) - (v_0, \ldots, v_i, \ldots, v_{n-1}) - (v_0, \ldots, v_i', \ldots, v_{n-1}),$$

$$(v_0, \ldots, cv_i, \ldots, v_{n-1}) - c(v_0, \ldots, v_i, \ldots, v_{n-1}),$$

em que $v_i, v_i' \in V_i$ para todo $i \in [n]$ e $c \in C$. O *produto tensorial* de $V_0, \ldots, V_{n-1}$ é o espaço linear

$$\boldsymbol{V_0} \otimes \cdots \otimes \boldsymbol{V_{n-1}} := \mathcal{F}(\boldsymbol{V_0} \times \cdots \times \boldsymbol{V_{n-1}}) \big/ \boldsymbol{\mathcal{R}}$$

A classe de equivalência de $(v_0, \ldots, v_{n-1}) \in V_0 \times \cdots \times V_{n-1}$ em $V_0 \otimes \cdots \otimes V_{n-1}$ é denotada $v_0 \otimes \cdots \otimes v_{n-1}$ e o *mapa tensorial canônico* é a função $\otimes := \pi \circ \iota \colon V_0 \times \cdots \times V_{n-1} \to V_0 \otimes \cdots \otimes V_{n-1}$, em que $\pi \colon \mathcal{F}(V_0 \times \cdots \times V_{n-1}) \to \mathcal{R}$ é a projeção do quociente de espaços lineares.

⊢ **Proposição 9.43** (Propriedade Característica de Produto Tensorial)**.** *Sejam* $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{n-1}, \boldsymbol{W}$ *espaços lineares sobre um corpo* $\boldsymbol{C}$*.*

1. *O mapa tensorial canônico*

$$\otimes \colon \boldsymbol{V_0} \times \cdots \times \boldsymbol{V_{n-1}} \to \boldsymbol{V_0} \otimes \cdots \otimes \boldsymbol{V_{n-1}}$$

é uma função multilinear.

2. *Para toda função multilinear* $L\colon \boldsymbol{V_0} \times \cdots \times \boldsymbol{V_{n-1}} \to \boldsymbol{W}$*, existe única função linear* $\bar{L}\colon \boldsymbol{V_0} \otimes \cdots \otimes \boldsymbol{V_{n-1}} \to \boldsymbol{W}$ *tal que* $\bar{L} \circ \otimes = L$ *(o diagrama comuta).*

$$\begin{tikzcd}
\boldsymbol{V_0} \otimes \cdots \otimes \boldsymbol{V_{n-1}} \arrow[dr, dashed, "\bar{L}"] & \\
\boldsymbol{V_0} \times \cdots \times \boldsymbol{V_{n-1}} \arrow[u, "\otimes"] \arrow[r, "L"'] & \boldsymbol{W}
\end{tikzcd}$$

□ *Demonstração.* 1. Vale por definição.

2. Pela propriedade característica de espaços lineares livres, existe única função linear $\tilde{L} : \mathcal{F}(V_0 \times \cdots \times V_{n-1}) \to W$ tal que $\tilde{L} \circ \iota = L$. Mas como L é multilinear, o subespaço $\mathcal{R}$ está contido no núcleo de $\tilde{L}$, pois, para todos $v_i, v_i' \in V_i$, com $i \in [n]$, e $c \in C$,

$$\begin{aligned}
\tilde{L}(v_0, \ldots, v_i + v_i', \ldots, v_{n-1}) &= L(v_0, \ldots, v_i + v_i', \ldots, v_{n-1}) \\
&= L(v_0, \ldots, v_i, \ldots, v_{n-1}) + L(v_0, \ldots, v_i', \ldots, v_{n-1}) \\
&= \tilde{L}(v_0, \ldots, v_i, \ldots, v_{n-1}) + \tilde{L}(v_0, \ldots, v_i', \ldots, v_{n-1}) \\
&= \tilde{L}((v_0, \ldots, v_i, \ldots, v_{n-1}) + (v_0, \ldots, v_i', \ldots, v_{n-1}))
\end{aligned}$$

e

$$\begin{aligned}
\tilde{L}(v_0, \ldots, cv_i, \ldots, v_{n-1}) &= L(v_0, \ldots, cv_i, \ldots, v_{n-1}) \\
&= cL(v_0, \ldots, v_i, \ldots, v_{n-1})) \\
&= c\tilde{L}(v_0, \ldots, v_i, \ldots, v_{n-1}) \\
&= \tilde{L}(c(v_0, \ldots, v_i, \ldots, v_{n-1})).
\end{aligned}$$

Isso implica que existe função linear $\bar{L} : V_0 \otimes \cdots \otimes V_{n-1} \to W$ que satisfaz $\bar{L} \circ \pi = \tilde{L}$. Como $\otimes = \pi \circ \iota$, segue que

$$\bar{L} \circ \otimes = \bar{L} \circ \pi \circ \iota = \tilde{L} \circ \iota = L. \qquad \blacksquare$$

Algumas identidades importantes. Os espaços lineares das funções multilineares é isomorfo ao espaço das funções lineares no produto tensorial:

$$\mathcal{L}(\boldsymbol{V_0}, \ldots, \boldsymbol{V_{n-1}}; \boldsymbol{W}) \simeq \mathcal{L}(\boldsymbol{V_0} \otimes \cdots \otimes \boldsymbol{V_{n-1}}; \boldsymbol{W}).$$

⊢ **Proposição 9.44.** *Valem os seguintes isomorfismos lineares:*

1. *Para todos espaços lineares* $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{n-1}$ *e* $\boldsymbol{V}'$,
$$\mathcal{L}(\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{n-1}; \boldsymbol{V}') \simeq \mathcal{L}(\boldsymbol{V}_0 \otimes \cdots \otimes \boldsymbol{V}_{n-1}; \boldsymbol{V}').$$
2. *Para todos espaços lineares* $\boldsymbol{V}$ *e* $\boldsymbol{V}'$ *e todo* $n \in \mathbb{N}$,
$$\mathcal{L}^n(\boldsymbol{V}, \boldsymbol{V}') \simeq \mathcal{L}(\boldsymbol{V}^{\otimes n}, \boldsymbol{V}').$$
3. *Para todos espaços lineares* $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{n-1}$ *e* $\boldsymbol{V}'$,
$$\mathcal{L}(\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{n-1}; \boldsymbol{V}') \simeq \mathcal{L}(\boldsymbol{V}_0, \mathcal{L}(\boldsymbol{V}_1, \ldots, \mathcal{L}(\boldsymbol{V}_{n-1}, \boldsymbol{V}') \ldots)).$$
4. *Para todos espaços lineares* $\boldsymbol{V}$ *e* $\boldsymbol{V}'$ *e todo* $n \in \mathbb{N}$,
$$\mathcal{L}^n(\boldsymbol{V}, \boldsymbol{V}') \simeq \mathcal{L}(\boldsymbol{V}, \mathcal{L}(\boldsymbol{V}, \ldots, \mathcal{L}(\boldsymbol{V}, \boldsymbol{V}') \ldots)).$$

□ *Demonstração.* Segue facilmente de 9.43 e 9.34. ■

9.10.1 Tensores

:⊢ **Definição 9.27.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$ e $k \in \mathbb{N}$. A *potência tensorial de tipo* k (ou *potência* k*-tensorial*) de $\boldsymbol{V}$ é o espaço
$$V^{\otimes k} := \bigotimes_{i \in [k]} V = \underbrace{V \otimes \cdots \otimes V}_{k}.$$
Um k-vetor de $\boldsymbol{V}$ é um elemento de $V^{\otimes k}$ e um k-covetor de $\boldsymbol{V}$ é um elemento de $(V^*)^{\otimes k}$.

A *potência tensorial de tipo* (k, k') (ou *potência* (k, k')*-tensorial*) de $\boldsymbol{V}$ é o espaço
$$V^{\otimes (k,k')} := (V^*)^{\otimes k} \otimes V^{\otimes k'}.$$
Um (k, k')-tensor de $\boldsymbol{V}$ é um elemento de $V^{\otimes (p,q)}$.

A *álgebra tensorial* de $\boldsymbol{V}$ é o espaço
$$V^{\otimes} := \bigoplus_{(p,q) \in \mathbb{N} \times \mathbb{N}} V^{\otimes (p,q)}.$$

Temos as identificações
$$\begin{aligned} V^{\otimes (0,0)} &\simeq C \\ V^{\otimes (1,0)} &\simeq V^* \\ V^{\otimes (0,1)} &\simeq V \\ V^{\otimes (1,1)} &\simeq \mathcal{L}(V, V). \end{aligned}$$

Capítulo 10

Álgebras

10.1 Álgebra e conceitos básicos

:⊢ **Definição 10.1.** Seja $\boldsymbol{C}$ um corpo. Uma *álgebra* sobre $\boldsymbol{C}$ é uma lista

$$\boldsymbol{A} := ((A, +, -, 0, \times), \cdot)$$

tal que $((A, +, -, 0), \cdot)$ é um espaço linear sobre $\boldsymbol{C}$ e $\times \colon A \times A \longrightarrow A$ é uma função bilinear. Uma álgebra é *associativa, comutativa, alternativa* ou *antissimétrica* conforme a respectiva propriedade de $\times$.

Uma *álgebra unitária* sobre $\boldsymbol{C}$ é uma lista

$$((A, +, -, 0, \times, 1), \cdot)$$

tal $((A, +, -, 0, \times), \cdot)$ é uma álgebra sobre $\boldsymbol{C}$ e $1 \in A$ é identidade de $(A, \times)$.

Uma *álgebra de divisão associativa* sobre $\boldsymbol{C}$ é uma lista

$$((A, +, -, 0, \times, 1, \diagup), \cdot)$$

tal $((A, +, -, 0, \times, 1), \cdot)$ é uma álgebra unitária associativa sobre $\boldsymbol{C}$ e $\diagup \colon A \setminus \{0\} \longrightarrow A \setminus \{0\}$ é uma inversão de $(A, \times, 1)$.

Por simplicidade, denotamos uma álgebra por $(\boldsymbol{A}, \times)$ se queremos ressaltar a notação para a multiplicação de A.

⊢ **Proposição 10.1.** *Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. Então*

$$((\mathcal{L}(\boldsymbol{V}), +, -, 0, \circ, \mathrm{I}), \cdot)$$

é uma álgebra unitária associativa sobre $\boldsymbol{C}$.

□ *Demonstração.* Sabemos que $((\mathcal{L}(\boldsymbol{V}), +, -, 0), \cdot)$ é um espaço linear sobre $\boldsymbol{C}$. Para mostrar que é uma álgebra, devemos mostrar que a composição de funções $\circ$ é bilinear. Sejam $L, L', L'' \in \mathcal{L}(\boldsymbol{V})$ e $c \in C$. Para todo $v \in V$, segue que

$$\begin{aligned}((cL + L') \circ L'')(v) &= (cL + L')(L''(v)) \\ &= cL(L''(v)) + L'(L''(v)) \\ &= cL \circ L''(v) + L' \circ L''(v) \\ &= (cL \circ L'' + L' \circ L'')(v),\end{aligned}$$

portanto $(cL + L') \circ L'' = cL \circ L'' + L' \circ L''$.

Para todo $v \in V$, segue da linearidade de L que

$$\begin{aligned}(L \circ (cL' + L''))(v) &= L((cL' + L'')(v)) \\ &= L(cL'(v) + L''(v)) \\ &= cL(L'(v)) + L(L''(v)) \\ &= cL \circ L'(v) + L \circ L''(v) \\ &= (cL \circ L' + L \circ L'')(v),\end{aligned}$$

portanto $L \circ (cL' + L'') = cL \circ L' + L \circ L''$. A composição de funções é associativa, portanto a álgebra é associativa. Por fim, a função identidade I é identidade da composição. ■

⊢ **Proposição 10.2.** *Sejam $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$ e I um conjunto. Então $(A^I, \times)$, em que $\times \colon A^I \times A^I \to A^I$ é o produto entrada a entrada, é uma álgebra sobre $\boldsymbol{C}$. Se o produto de A é associativo ou comutativo, então o produto de A^I é, respectivamente, associativo ou comutativo, e é se A é unitária, $(1)_{in \in I}$ é identidade do produto de A^I.*

□ *Demonstração.* Sabemos que A^I é um espaço linear sobre $\boldsymbol{C}$. Basta mostrar que $\times$ é um produto bilinear. Sejam $(a_i)_{i\in I}, (a'_i)_{i\in I}, (a''_i)_{i\in I} \in A^I$ e $c \in C$. Então

$$\begin{aligned}(c(a_i)_{i\in I} + (a'_i)_{i\in I}) \times (a''_i)_{i\in I} &= (ca_i + a'_i)_{i\in I} \times (a''_i)_{i\in I} \\ &= ((ca_i + a'_i) \times a''_i)_{i\in I} \\ &= (ca_i \times a''_i + a'_i \times a''_i)_{i\in I} \\ &= c(a_i \times a''_i)_{i\in I} + (a'_i \times a''_i)_{i\in I} \\ &= c(a_i)_{i\in I} \times (a''_i)_{i\in I} + (a'_i)_{i\in I} \times (a''_i)_{i\in I}.\end{aligned}$$

A demonstração da linearidade na segunda entrada é análoga, e as demonstrações de associatividade e comutatividade e identidade são triviais. ■

10.1.1 Ação adjunta

Como de costume, denotamos a multiplicação à esquerda por um elemento fixo por $a \in A$ por

$$\begin{aligned} a\times \colon A &\longrightarrow A \\ a' &\longmapsto a \times a'. \end{aligned}$$

Pela bilinearidade da multiplicação, temos que $a\times \in \mathcal{L}(\boldsymbol{A})$ e, portanto está bem definida uma função que leva cada $a \in A$ em uma função linear $a\times$, como definimos a seguir.

:⊢ **Definição 10.2.** Sejam $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$ e $a \in A$. A *ação adjunta* de $\boldsymbol{A}$ é a função linear

$$\begin{aligned} \mathrm{ad} \colon A &\longrightarrow \mathcal{L}(\boldsymbol{A}) \\ a &\longmapsto \mathrm{ad}_a := a\times \colon A \longrightarrow A \\ & \qquad\qquad\qquad a' \longmapsto a \times a'. \end{aligned}$$

▷ **Exercício 10.1.** *Seja $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$. A álgebra $\boldsymbol{A}$ é associativa se, e somente se, para todos $a, a' \in A$,*

$$\mathrm{ad}_{a\times a'} = \mathrm{ad}_a \circ \mathrm{ad}_{a'}\,.$$

Explicitamente, a condição anterior é

$$(a \times a')\times = (a\times) \circ (a'\times).$$

A ação adjunta é uma representação quando a álgebra é associativa.

10.1.2 Colchete comutador

:⊢ **Definição 10.3.** Seja $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$. O *colchete comutador* de $(\boldsymbol{A}, \times)$ é a função

$$\begin{aligned} [\cdot, \cdot] \colon A \times A &\longrightarrow A \\ (a, a') &\longmapsto a \times a' - a' \times a. \end{aligned}$$

⊢ **Proposição 10.3.** *Seja $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$. Então*

1. *$(\boldsymbol{A}, [\cdot, \cdot])$ é uma álgebra alternativa sobre $\boldsymbol{C}$;*
2. *O produto $\times$ é comutativo se, e somente se, $[\cdot, \cdot] = 0$;*

□ *Demonstração.* 1. Primeiro, notemos que, para todos $a, a' \in A$,

$$[a, a'] = a \times a' - a' \times a = -(a' \times a - a \times a') = -[a', a].$$

portanto $[\cdot, \cdot]$ é antissimétrico. Sendo assim, para mostrar que $[\cdot, \cdot]$ é bilinear, basta mostrar que ela é linear na primeira entrada. Para todos $a, a', a'' \in A$ e $c \in C$,

$$\begin{aligned}[ca + a', a''] &= (ca + a') \times a'' - a'' \times (ca + a') \\ &= ca \times a'' + a' \times a'' - ca'' \times a - a'' \times a' \\ &= ca \times a'' - ca'' \times a + a' \times a'' - a'' \times a' \\ &= c\,[a, a''] + [a', a''].\end{aligned}$$

Por fim, para todo $a \in A$,

$$[a, a] = a \times a - a \times a = 0.$$

2. Suponhamos, primeiro, que $\times$ é comutativo. Então, para todos $a, a' \in A$,

$$[a, a'] = a \times a' - a' \times a = a \times a' - a \times a' = 0.$$

Reciprocamente, suponhamos que $[\cdot, \cdot] = 0$. Então, para todos $a, a' \in A$,

$$a \times a' = a \times a' + 0 = a \times a' + [a', a] = a \times a' + a' \times a - a \times a' = a' \times a.$$

■

10.1.3 Derivação

:⊢ **Definição 10.4.** Seja $(\boldsymbol{A}, \times)$ uma álgebra sobre um corpo $\boldsymbol{C}$. Uma *derivação* em $\boldsymbol{A}$ é uma função linear $D\colon A \longrightarrow A$ tal que

- (Derivabilidade[1]) Para todos $a, a' \in A$,

$$D(a \times a') = D(a) \times a' + a \times D(a').$$

O conjunto dessas derivações é $\mathfrak{D}(\boldsymbol{A})$.

Note que a propriedade acima nem sempre é equivalente a

$$D(a \times a') = a' \times D(a) + a \times D(a'),$$

pois o produto $\times$ nem sempre é comutativo, mas sempre é equivalente a

$$D(a \times a') = a \times D(a') + D(a) \times a',$$

pois a soma $+$ é comutativa.

[1] A propriedade de derivabilidade é também chamada de 'regra do produto' ou 'regra de Leibniz'.

⊢ **Proposição 10.4.** *Sejam* $(\boldsymbol{A}, \times)$ *uma álgebra sobre um corpo* $\boldsymbol{C}$ *e* $D\colon A \longrightarrow A$ *uma derivação em* $\boldsymbol{A}$.

1. *(Regra do produto generalizada) Para todos* $a_0, \ldots, a_{n-1} \in A$,
$$D(a_0 \cdots a_{n-1}) = \mathop{+}_{i\in[n]} a_0 \cdots D(a_i) \cdots a_{n-1};$$
2. *Se* $\times$ *é comutativa, então, para todos* $a \in A$ *e* $n \in \mathbb{N} \smallsetminus \{0\}$,
$$D(a^n) = na^{n-1}D(a);$$
3. *Se existe unidade* $1 \in A$ *do produto, então*
$$D(1) = 0.$$
4. *(Regra do produto de ordem superior) Para todos* $a, a' \in A$ *e* $n \in \mathbb{N}$,
$$D^n(aa') = \mathop{+}_{i\in[n+1]} \binom{n}{i} D^{n-i}(a)D^i(a').$$

Uma derivação pode ser equivalentemente descrita usando o comutador de composição de funções.

▷ **Exercício 10.2.** *Sejam* $\boldsymbol{A}$ *uma álgebra sobre um corpo* $\boldsymbol{C}$ *e* $D\colon A \longrightarrow A$ *uma função linear. São equivalentes*

1. D *é uma derivação;*
2. *Para todo* $a \in A$, $(D(a))\times = [D, a\times]$;
3. *Para todo* $a \in A$, $\times(D(a)) = [D, \times a]$.

⊢ **Proposição 10.5.** *Seja* $\boldsymbol{A}$ *uma álgebra sobre um corpo* $\boldsymbol{C}$. *O conjunto* $\mathfrak{D}(\boldsymbol{A})$ *das derivações em* $\boldsymbol{A}$ *é um subespaço linear de* $\mathcal{L}(\boldsymbol{A})$.

□ *Demonstração.* Para mostrar que $\mathfrak{D}(\boldsymbol{A})$ é subespaço linear de $\mathcal{L}(\boldsymbol{A})$, mostraremos que $\mathfrak{D}(\boldsymbol{A})$ é fechado pela adição e pela ação por escalar pontuais. Para todo $c \in C$ e todas derivações $D, D' \in \mathfrak{D}(\boldsymbol{A})$, a função $cD + D'$ é linear, pois D e D' são lineares; resta mostrar que ela é uma derivação. Para todos $a, a' \in A$,

$$\begin{aligned}
(cD + D')(aa') &= (cD)(aa') + D'(aa') \\
&= cD(aa') + D'(aa') \\
&= cD(a)a' + caD(a') + D'(a)a' + aD'(a') \\
&= (cD(a) + D'(a))a' + a(cD(a') + D'(a')) \\
&= (cD + D')(a)a' + a(cD + D')(a').
\end{aligned}$$

portanto $cD + D'$ é uma derivação. ■

10.2 Álgebra derivativa

As álgebras que descreveremos nesta seção são conhecidas como 'álgebras de Lie'. Elas são álgebras cujo produto é antissimétrico e satisfaz uma propriedade importante, conhecida por 'identidade de Jacobi', mas que chamamos de derivação adjunta pois ela é equivalente a dizer que a adjunta de todo elemento é uma derivação da álgebra.

$\vdash$ **Proposição 10.6.** *Sejam $\boldsymbol{A}$ uma álgebra sobre um corpo $\boldsymbol{C}$ e $a \in A$. A função adjunta* ad_a *é uma derivação em $\boldsymbol{A}$ se, e somente se,*

- *(Derivação adjunta) Para todos $a', a'' \in A$,*

$$a \times (a' \times a'') = (a \times a') \times a'' + a' \times (a \times a'').$$

A demonstração é imediata. Como comentado, essa propriedade é conhecida às vezes como 'identidade de Jacobi'. No entanto, a identidade mais conhecida como 'identidade de Jacobi' é

$$a \times (a' \times a'') + a' \times (a'' \times a) + a'' \times (a \times a') = 0,$$

que é equivalente à anterior se o produto é antissimétrico. Na maioria das vezes em que se usa essa identidade o produto é de fato antissimétrico, o que torna as duas propriedades equivalentes.

$:\vdash$ **Definição 10.5.** Seja $\boldsymbol{C}$ um corpo. Uma *álgebra derivativa* sobre $\boldsymbol{C}$ é uma álgebra

$$\boldsymbol{\mathfrak{A}} := ((\mathfrak{A}, +, -, 0, [\cdot, \cdot]), \cdot)$$

sobre $\boldsymbol{C}$ cuja multiplicação $[\cdot, \cdot] : \mathfrak{A} \times \mathfrak{A} \longrightarrow \mathfrak{A}$ satisfaz

- (Alternatividade) Para todo $a \in \mathfrak{A}$, $[a, a] = 0$;
- (Derivação adjunta) Para todo $a, a', a'' \in \mathfrak{A}$,

$$[a, [a', a'']] = [[a, a'], a''] + [a', [a, a'']] \,.$$

A multiplicação $[\cdot, \cdot]$ é o *colchete de derivação* da álgebra. A ação adjunta de $\mathfrak{A}$ é denotada

$$\begin{aligned} [\cdot| : \mathfrak{A} &\longrightarrow \mathcal{L}(\boldsymbol{\mathfrak{A}}) \\ a &\longmapsto [a| : \mathfrak{A} \longrightarrow \mathfrak{A} \\ & \qquad\qquad a' \longmapsto [a, a'] \,. \end{aligned}$$

$\vdash$ **Proposição 10.7.** *Seja $(\boldsymbol{A}, \times)$ uma álgebra associativa sobre um corpo $\boldsymbol{C}$. A álgebra $(\boldsymbol{A}, [\cdot, \cdot])$, em que $[\cdot, \cdot]$ é o colchete comutador de $\times$, é uma álgebra derivativa sobre $\boldsymbol{C}$.*

□ *Demonstração.* Já demonstramos que $(\boldsymbol{A}, [\cdot,\cdot])$ é uma álgebra alternativa, resta mostrar a derivação adjunta. Para todos $a, a', a'' \in A$, segue da associatividade de $\times$ que

$$\begin{aligned}
[a,[a',a'']] &= a \times (a' \times a'' - a'' \times a') - (a' \times a'' - a'' \times a') \times a \\
&= a \times (a' \times a'') - a \times (a'' \times a') - (a' \times a'') \times a + (a'' \times a') \times a \\
&= (a \times a') \times a'' - (a \times a'') \times a' - a' \times (a'' \times a) + a'' \times (a' \times a) \\
&= (a \times a') \times a'' - (a \times a'') \times a' - a' \times (a'' \times a) + a'' \times (a' \times a) \\
&+ a' \times (a \times a'') - (a' \times a) \times a'' + (a'' \times a) \times a' - a'' \times (a \times a') \\
&= (a \times a') \times a'' - (a' \times a) \times a'' - a'' \times (a \times a') + a'' \times (a' \times a) \\
&+ a' \times (a \times a'') - a' \times (a'' \times a) - (a \times a'') \times a' + (a'' \times a) \times a' \\
&= (a \times a' - a' \times a) \times a'' - a'' \times (a \times a' - a' \times a) \\
&+ a' \times (a \times a'' - a'' \times a) - (a \times a'' - a'' \times a) \times a' \\
&= [[a,a'],a''] + [a',[a,a'']] \,. \qquad \blacksquare
\end{aligned}$$

Isso mostra, em particular, que a álgebra de funções lineares em um espaço linear $\boldsymbol{V}$ com composição $((\mathcal{L}(\boldsymbol{V}), +, -, 0, \circ), \cdot)$ gera uma álgebra derivativa

$$((\mathcal{L}(\boldsymbol{V}), +, -, 0, [\cdot,\cdot]), \cdot),$$

em que o colchete comutador da álgebra é dado, para todas $L, L' \in \mathcal{L}(\boldsymbol{V})$, por

$$[L, L'] = L \circ L' - L' \circ L.$$

Para diferenciá-la da álgebra associativa de composição de funções lineares, a denotaremos como $\mathfrak{L}(\boldsymbol{V})$. A estrutura de espaço linear subjacente é a mesma para as duas álgebras, mas a multiplicação é diferente.

:⊢ **Definição 10.6.** Seja $\boldsymbol{V}$ um espaço vetorial sobre um corpo $\boldsymbol{C}$. A *álgebra derivativa linear* de $\boldsymbol{V}$ é a álgebra derivativa

$$\boldsymbol{\mathfrak{L}(V)} := ((\mathfrak{L}(\boldsymbol{V}), +, -, 0, [\cdot,\cdot]), \cdot)$$

em que $\mathfrak{L}(\boldsymbol{V}) := \mathcal{L}(\boldsymbol{V})$, $((\mathfrak{L}(\boldsymbol{V}), +, -, 0), \cdot)$ é o espaço linear de funções lineares e $[\cdot,\cdot]$ é o colchete comutador da composição $\circ$ de funções.

Consideremos, agora, o conjunto $\mathfrak{D}(\boldsymbol{A})$ das derivações em uma álgebra $\boldsymbol{A}$. Por 10.5, o espaço $\mathfrak{D}(\boldsymbol{A})$ é um subespaço linear de $\mathcal{L}(\boldsymbol{A})$. No entanto, $\mathfrak{D}(\boldsymbol{A})$ não é uma subálgebra de $\mathcal{L}(\boldsymbol{A})$ com o produto de composição de funções, pois, para todos $D, D' \in \mathfrak{D}(\boldsymbol{A})$ e $a, a' \in A$,

$$(D \circ D')(aa') = (D \circ D')(a)a' + a(D \circ D')(a') + D'(a)D(a') + D(a)D'(a')$$

o que significa que $D \circ D'$ nem sempre é uma derivação, pois $D'(a) \times D(a') + D(a) \times D'(a')$ nem sempre é 0. Pra resolver esse problema, devemos considerar em $\mathfrak{D}(\boldsymbol{A})$ o colchete comutador da composição de funções

⊢ **Proposição 10.8.** *Seja $\boldsymbol{A}$ uma álgebra sobre um corpo $\boldsymbol{C}$. A álgebra de derivações $(\mathfrak{D}(\boldsymbol{A}), [\cdot,\cdot])$ é uma álgebra derivativa.*

□ *Demonstração.* Por 10.5, o espaço $((\mathfrak{D}(\boldsymbol{A}), +, -, 0), \cdot)$ é um subespaço linear de $((\mathcal{L}(\boldsymbol{A}), +, -, 0), \cdot)$. Primeiro mostramos que $\mathfrak{D}(\boldsymbol{A})$ é fechado pelo colchete comutador. Denotaremos o produto $a \times a'$ por aa' para simplificar a notação. Para todas $D, D' \in \mathfrak{D}(\boldsymbol{A})$ e $a, a' \in A$,

$$\begin{aligned}
(D \circ D')(aa') &= D(D'(aa')) \\
&= D(D'(a)a' + aD'(a')) \\
&= D(D'(a)a') + D(aD'(a')) \\
&= D(D'(a))a' + D'(a)D(a') + D(a)D'(a') + aD(D'(a')) \\
&= (D \circ D')(a)a' + a(D \circ D')(a') + D'(a)D(a') + D(a)D'(a').
\end{aligned}$$

Trocando D e D', segue que

$$(D' \circ D)(aa') = (D' \circ D)(a)a' + a(D' \circ D)(a') + D(a)D'(a') + D'(a)D(a'),$$

Dessas duas expressões, segue que

$$\begin{aligned}
[D, D'](aa') &= (D \circ D')(aa') - (D' \circ D)(aa') \\
&= (D \circ D')(a)a' + a(D \circ D')(a') - (D' \circ D)(a)a' - a(D' \circ D)(a') \\
&= [D, D'](a)a' + a[D, D'](a').
\end{aligned}$$

portanto $[D, D'] \in \mathfrak{D}(\boldsymbol{A})$.

■

10.3 Álgebra involutiva

Antes de definirmos uma álgebra involutiva, definimos o que é um anel (e um corpo) involutivos.

:⊢ **Definição 10.7.** Uma *anel involutivo* é um par $(\boldsymbol{A}, {}^*)$ tal que $\boldsymbol{A}$ é um anel e ${}^*\colon A \longrightarrow A$ é uma operação unária que satisfaz

- (Involutividade) Para todo $a \in A$, $(a^*)^* = a$;
- (Aditividade) Para todos $a, a' \in A$, $(a + a')^* = a^* + (a')^*$;
- (Anti-multiplicatividade) Para todos $a, a' \in A$, $(aa')^* = (a')^* a^*$;
- (Identidade multiplicativa) $1^* = 1$.

⊢ **Proposição 10.9.** *Seja $(\boldsymbol{A}, *)$ um anel involutivo.*

1. *(Identidade aditiva)* $0^* = 0$*;*
2. *(Invertibilidade aditiva) Para todo* $a \in A$, $(-a)^* = -a^*$*;*
3. *(Invertibilidade multiplicativa) Para todo* $a \in A^{\prime}$, $a^* \in A^{\prime}$ *e* $(a^{-1})^* = (a^*)^{-1}$.

□ *Demonstração.* 1. Segue da aditividade que

$$0 = 0^* - 0^* = 0 + 0^* - 0^* = 0^* + 0^* - 0^* = 0^*.$$

2. Para todo $a \in A$, segue da aditividade e do item anterior que

$$\begin{aligned} -a^* &= -a^* + 0 \\ &= -a^* + a^* - a^* \\ &= (-a + a)^* - a^* \\ &= 0^* - a^* \\ &= 0 - a^* \\ &= a^*. \end{aligned}$$

3. Para todo $a \in A^{\prime}$, segue da multiplicatividade e da identidade multiplicativa que

$$1 = 1^* = aa^{-1^*} = a^*(a^{-1})^*$$

e

$$1 = 1^* = a^{-1}a^* = (a^{-1})^* a^*$$

portanto $a^* \in A^{\prime}$ e $(a^{-1})^* = (a^*)^{-1}$.

■

Notação. Por causa da propriedade de invertibilidade multiplicativa, denota-se, para todo $a \in A^{\prime}$,

$$a^{-*} := (a^{-1})^*.$$

► **Exemplo 10.1.** *1. Todo anel com a involução trivial dada pela identidade.*
2. Os números complexos com a conjugação complexa.
3. O anel de matrizes reais com a transposta.
4. O anel de matrizes reais com a transposta conjugada.
5. A adjunta de uma transformação linear contínua em um espaço de produto interno completo.

:⊢ **Definição 10.8.** Seja $(\boldsymbol{C}, {}^*)$ um corpo involutivo. Uma *álgebra involutiva* sobre $\boldsymbol{C}$ é uma lista $((A, +, -, 0, \times, 1, {}^*), \cdot)$ em que $((A, +, -, 0, \times, 1), \cdot)$ é um álgebra unitária sobre $\boldsymbol{C}$ e ${}^*\colon A \longrightarrow A$ é uma operação unária que satisfaz

- (Involutividade) Para todo $a \in A$, $(a^*)^* = a$;
- (Aditividade) Para todos $a, a' \in A$, $(a + a')^* = a^* + (a')^*$;
- (Anti-multiplicatividade) Para todos $a, a' \in A$, $(aa')^* = (a')^* a^*$;
- (Identidade multiplicativa) $1^* = 1$;
- (Compatibilidade) Para todos $c \in C$ e $a \in A$, $(ca)^* = c^* a^*$.

:⊢ **Definição 10.9.** Seja $\boldsymbol{A}$ uma álgebra involutiva sobre um corpo involutivo $\boldsymbol{C}$. A norma quadrática de $\boldsymbol{A}$ é a função

$$\begin{aligned} [\![\cdot]\!] \colon A &\longrightarrow A \\ x &\longmapsto [\![x]\!] := xx^*. \end{aligned}$$

⊢ **Proposição 10.10.** *Seja $\boldsymbol{A}$ uma álgebra involutiva sobre um corpo involutivo $\boldsymbol{C}$.*

1. *Para todo $c \in C$ e $x \in A$, $[\![cx]\!] = [\![c]\!]\,[\![x]\!]$;*
2. *Para todo $x \in A$, $[\![x^*]\!] = x^* x$;*
3. *Para todo $x \in A$, $[\![x]\!]^* = [\![x]\!]$;*
4. *A função $B(x, y) := [\![x + y]\!] - [\![x]\!] - [\![y]\!]$ é satisfaz $B(x, y) = xy^* + yx^*$, é simétrica e comuta com adição e inversa da adição em cada entrada.*

□ *Demonstração.* 1. Para todos $c \in C$ e $x \in A$, segue da bilinearidade da multiplicação e compatibilidade da involução que

$$[\![cx]\!] = (cx)(cx)^* = cc^* xx^* = [\![c]\!]\,[\![x]\!]\,.$$

2. Para todo $x \in A$,

$$[\![x^*]\!] = x^*(x^*)^* = x^* x;$$

3. Para todo $x \in A$,

$$[\![x]\!]^* = (xx^*)^* = (x^*)^* x^* = xx^* = [\![x]\!]\,;$$

4. Para todos $x, y \in A$,

$$\begin{aligned} B(x, y) &= [\![x + y]\!] - [\![x]\!] - [\![y]\!] \\ &= (x + y)(x + y)^* - xx^* - yy^* \\ &= x(x^* + y^*) + y(x^* + y^*) - xx^* - yy^* \\ &= xx^* + xy^* + yx^* + yy^* - xx^* - yy^* \\ &= xy^* + yx^*, \end{aligned}$$

portanto B é simétrica. Além disso, para todos $x, x', y \in A$,

$$\begin{aligned} B(x,y) &= (x+x')y^* + y(x+x')^* \\ &= xy^* + x'y^* + yx^* + y(x')^* \\ &= xy^* + yx^* + x'y^* + y(x')^* \\ &= B(x,y) + B(x',y) \end{aligned}$$

$$B(-x,y) = (-x)y^* + y(-x)^* = -xy^* - yx^* = -B(x,y).$$

portanto B comuta com adição e inversa da adição em cada entrada.

■

Produto torcido Essa construção de uma álgebra produto é conhecida por 'construção de Cayley-Dickson'. Definamos, a partir de uma álgebra involutiva $\boldsymbol{A}$, uma álgebra involutiva na soma direta A^2.

:⊢ **Definição 10.10.** Seja $\boldsymbol{A}$ uma álgebra involutiva sobre um corpo involutivo $\boldsymbol{C}$. O *produto torcido* de $\boldsymbol{A}$ (por $\boldsymbol{A}$) é

$$((A^2, +, -, 0, \times, 1, {}^*), \cdot)$$

em que

- $((A^2, +, -, 0), \cdot)$ é o produto (soma direta) como espaço linear;
- $\times$ é a *multiplicação torcida* definida por

$$\begin{aligned} \times \colon A^2 \times A^2 &\longrightarrow A^2 \\ ((x_0, x_1), (y_0, y_1)) &\longmapsto (x_0 y_0 - {y_1}^* x_1, y_1 x_0 + x_1 {y_0}^*) \end{aligned}$$

- * é a *involução torcida* definida por

$$\begin{aligned} {}^* \colon A^2 &\longrightarrow A^2 \\ (x_0, x_1) &\longmapsto ({x_0}^*, -x_1) \end{aligned}$$

- 1 é a unidade definida por $1 := (1, 0)$.

10.4 Álgebras reais

Nesta seção, descreveremos algumas álgebras sobre o corpo dos números reais $\mathbb{R}$. A álgebra real $\mathbb{R}$ é simplesmente o corpo $\mathbb{R}$. Analisaremos com detalhes os casos das álgebras dos complexos $\mathbb{R}^2$ e dos quatérnios $\mathbb{R}^4$. Descreveremos brevemente os números 2-hiperbólicos $\mathbb{R}^{1,1}$, às vezes chamados de 'complexos hiperbólicos', e os números 4-hiperbólicos $\mathbb{R}^{1,3}$.

10.4.1 Complexos $\mathbb{R}^2$

Consideremos o espaço linear $\mathbb{R}^2$ com a norma 2

$$\begin{aligned} \|\cdot\| \colon \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ x &\longmapsto ((x_0)^2 + (x_1)^2)^{\frac{1}{2}}. \end{aligned}$$

A base canônica de $\mathbb{R}^2$ será denotada por

$$\begin{aligned} \mathrm{i}_0 &:= (1,0) \\ \mathrm{i}_1 &:= (0,1). \end{aligned}$$

:⊢ **Definição 10.11.** A *multiplicação complexa* em $\mathbb{R}^2$ é a operação

$$\begin{aligned} \times \colon \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longmapsto xy := (x_0y_0 - x_1y_1, x_0y_1 + x_1y_0). \end{aligned}$$

A *unidade* de $(\mathbb{R}^2, \times)$ é

$$1_{\mathbb{R}^2} := \mathrm{i}_0 = (1,0).$$

A *inversão* de $(\mathbb{R}^2, \times, 1)$ é a operação

$$\begin{aligned} \diagup \colon \mathbb{R}^2 \smallsetminus \{0\} &\longrightarrow \mathbb{R}^2 \smallsetminus \{0\} \\ x &\longmapsto x^{-1} := \|x\|^{-2}(x_0, -x_1). \end{aligned}$$

Em geral, por não haver ambiguidade, denotaremos $1_{\mathbb{R}^2}$ simplesmente por 1 e i_1 simplesmente por i. Assim, todo $x \in \mathbb{R}^2$ pode ser denotado

$$x = x_0 1 + x_1 \mathrm{i}.$$

Como a notação acima sugere, omitiremos o símbolo de multiplicação $\times$ sempre que possível pois isso não gerará ambiguidade.

Notemos que

$$\mathrm{i}^2 = (0,1)(0,1) = (0 \times 0 - 1 \times 1, 0 \times 1 + 1 \times 0) = (-1,0) = -1.$$

⊢ **Proposição 10.11.** *A lista*

$$(\mathbb{R}^2, +, -, 0, \times, 1, \cdot)$$

é uma álgebra associativa, comutativa, unitária e invertível sobre $\mathbb{R}$.

□ *Demonstração.* Como $(\mathbb{R}^2, +, -, 0, \cdot)$ é um espaço linear sobre $\mathbb{R}$, basta mostrar as propriedades de $\times$.

- (Comutatividade) Para todos $x, y \in \mathbb{R}^2$,

$$xy = (x_0y_0 - x_1y_1, x_0y_1 + x_1y_0) = (y_0x_0 - y_1x_1, y_0x_1 + y_1x_0) = yx.$$

- (Bilinearidade) Como a multiplicação é comutativa, basta mostrar a linearidade na primeira entrada. Para todos $x, x', y \in \mathbb{R}^2$ e todo $c \in \mathbb{R}$,

$$\begin{aligned}(cx + x')y &= ((cx_0 + x'_0)y_0 - (cx_1 + x'_1)y_1, (cx_0 + x'_0)y_1 + (cx_1 + x'_1)y_0)\\ &= (cx_0y_0 + x'_0y_0 - cx_1y_1 - x'_1y_1, cx_0y_1 + x'_0y_1 + cx_1y_0 + x'_1y_0)\\ &= c(x_0y_0 - x_1y_1, x_0y_1 + x_1y_0) + (x'_0y_0 - x'_1y_1, x'_0y_1x'_1y_0)\\ &= c(xy) + x'y.\end{aligned}$$

- (Associatividade) Faremos as contas por coordenadas. Para todos $x, y, z \in \mathbb{R}^2$, segue da definição do produto que as coordenadas de xy são

$$\begin{aligned}(xy)_0 &= x_0y_0 - x_1y_1\\ (xy)_1 &= x_0y_1 + x_1y_0\end{aligned}$$

e que as coordenadas de yz são

$$\begin{aligned}(yz)_0 &= y_0z_0 - y_1z_1\\ (yz)_1 &= y_0z_1 + y_1z_0.\end{aligned}$$

Ainda da definição do produto, as coordenadas de $(xy)z$ são

$$\begin{aligned}((xy)z)_0 &= (x_0y_0 - x_1y_1)z_0 - (x_0y_1 + x_1y_0)z_1\\ ((xy)z)_1 &= (x_0y_0 - x_1y_1)z_1 + (x_0y_1 + x_1y_0)z_0\end{aligned}$$

e as coordenadas de $x(yz)$ são

$$\begin{aligned}(x(yz))_0 &= x_0(y_0z_0 - y_1z_1) - x_1(y_0z_1 + y_1z_0)\\ (x(yz))_1 &= x_0(y_0z_1 + y_1z_0) + x_1(y_0z_0 - y_1z_1).\end{aligned}$$

Para a coordenada 0, segue que

$$\begin{aligned}((xy)z)_0 &= (x_0y_0 - x_1y_1)z_0 - (x_0y_1 + x_1y_0)z_1\\ &= x_0(y_0z_0 - y_1z_1) - x_1(y_0z_1 + y_1z_0)\\ &= (x(yz))_0\end{aligned}$$

e para a coordenada 1 segue que

$$\begin{aligned}((xy)z)_1 &= (x_0y_0 - x_1y_1)z_1 + (x_0y_1 + x_1y_0)z_0\\ &= x_0(y_0z_1 + y_1z_0) + x_1(y_0z_0 - y_1z_1)\\ &= (x(yz))_1.\end{aligned}$$

Portanto

$$x(yz) = ((xy)z)_0 1 + ((xy)z)_1 \mathrm{i} = (x(yz))_0 1 + (x(yz))_1 \mathrm{i} = x(yz).$$

- (Identidade) Para todo $x \in \mathbb{R}^2$,

$$1x = (1x_0 - 0x_1, 1x_1 + 0x_0) = (x_0, x_1) = x.$$

- (Inversa) Para todo $x \in \mathbb{R}^2 \smallsetminus \{0\}$,

$$\begin{aligned} x^{-1}x &= \|x\|^{-2}(x_0, -x_1)(x_0, x_1) \\ &= \|x\|^{-2}(x_0x_0 - (-x_1)x_1, x_0(-x_1) + x_1x_0) \\ &= \|x\|^{-2}(\|x\|^2, 0) \\ &= (1, 0) \\ &= 1. \end{aligned}$$

■

Como o vetor $1 = (1, 0)$ é uma unidade, em geral o omitiremos da notação da base, denotando $x_0 1$ por x_0 e

$$x = x_0 + x_1\mathrm{i},$$

de modo que a multiplicação $\times$ nessa notação é a multiplicação usual dos números complexos

$$xy = (x_0 + x_1\mathrm{i})(y_0 + y_1\mathrm{i}) = (x_0y_0 - x_1y_1) + (x_0y_1 + x_1y_0)\mathrm{i}.$$

:⊢ **Definição 10.12.** A *conjugação* em $\mathbb{R}^2$ é a operação

$$\begin{aligned} {}^*\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto x^* := (x_0, -x_1). \end{aligned}$$

▷ **Exercício 10.3.** *1. (Involutividade) Para todo $x \in \mathbb{R}^2$, $(x^*)^* = x$;*
2. (Aditividade) Para todos $x, x' \in \mathbb{R}^2$, $(x + x')^* = x^* + (x')^*$;
3. (Identidade aditiva) $0^* = 0$;
4. (Invertibilidade aditiva) Para todo $x \in \mathbb{R}^2$, $(-x)^* = -x^*$;
5. (Anti-multiplicatividade) Para todos $x, x' \in \mathbb{R}^2$, $(xx')^* = (x')^*x^*$;
6. (Identidade multiplicativa) $1^* = 1$;

▷ **Exercício 10.4.** *Seja $x \in \mathbb{R}^2$.*

1. $\|x\|^2 = |x_0|^2 + \|\vec{x}\|^2$;
2. $\|x\|^2 = xx^*$;

A multiplicação $\times$ pode também ser vista como uma ação de $\mathbb{R}^2$ aditivo sobre $\mathbb{R}^2$ aditivo. Cada $x \in \mathbb{R}^2$ pode ser identificado com uma transformação linear $x\colon \mathbb{R}^2 \to \mathbb{R}^2$. Os elementos $x \in \mathbb{S}^1 \subseteq \mathbb{R}^2$ são as rotações: se $x \in \mathbb{S}^1$, então existe $\theta \in [0, \tau[$ tal que

$$x = \cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}.$$

Sendo assim, para todo $\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i} \in \mathbb{S}^1$, e todo $x \in \mathbb{R}^2$,

$$(\cos(\theta) + \operatorname{sen}(\theta))\mathrm{i} \times x = (\cos(\theta)x_0 - \operatorname{sen}(\theta)x_1) + (\cos(\theta)x_1 + \operatorname{sen}(\theta)x_0)\mathrm{i} = R_\theta(x).$$

Os elementos $x = x_0 \in \mathbb{R} \subseteq \mathbb{R}^2$ são as expansões e contrações. Os elementos de $\mathbb{R}^2$ podem ser decompostos como

$$x = \|x\| \left(\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}\right).$$

$\vdash$ **Proposição 10.12** (Decomposição polar)**.** *Seja* $x \in \mathbb{R}^2 \smallsetminus \{0\}$*. Então*

$$x = \|x\| \left(\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}\right),$$

em que $\theta \in [0, \tau[$ *é definido*

$$\theta := \begin{cases} \cos^{-1}\left(\frac{x_0}{\|x\|}\right), & x_1 > 0 \quad ou \quad x_0 = 1, \\ \tau - \cos^{-1}\left(\frac{x_0}{\|x\|}\right) & x_1 < 0 \quad ou \quad x_0 = -1. \end{cases}$$

$\square$ *Demonstração.* Como $x \neq 0$, $\|x\| \neq 0$, logo da igualdade $\|x\|^2 = {x_0}^2 + {x_1}^2$ segue que

$$1 = \left(\frac{x_0}{\|x\|}\right)^2 + \left(\frac{x_1}{\|x\|}\right)^2,$$

portanto $\frac{|x_0|}{\|x\|} \leq 1$, o que é equivalente a $\frac{x_0}{\|x\|} \in [-1, 1]$. Disso segue que $\phi := \cos^{-1}\left(\frac{x_0}{\|x\|}\right) \in [0, \tau \diagup 2]$, de modo que $\frac{|x_1|}{\|x\|} = (1 - (\cos(\phi))^2)^{1\diagup 2}$, o que implica que $\frac{|x_1|}{\|x\|} = \operatorname{sen}(\phi)$ ou $\frac{|x_1|}{\|x\|} = -\operatorname{sen}(\phi)$. Como $\phi \in [0, \tau \diagup 2]$, concluímos que $\operatorname{sen}(\phi) = \frac{|x_1|}{\|x\|}$.

Da definição de θ segue que $x_0 = \|x\| \cos(\theta)$ e $x_1 = \|x\| \operatorname{sen}(\theta)$. Para a demonstração disso, consideramos dois casos.

- ($x_1 > 0$ ou $x_0 = 1$) Nesse caso, $|x_1| = x_1$ e $\theta = \phi$, portanto

$$\cos(\theta) = \cos(\phi) = \frac{x_0}{\|x\|}$$

  e

$$\operatorname{sen}(\theta) = \operatorname{sen}(\phi) = \frac{|x_1|}{\|x\|} = \frac{x_1}{\|x\|}.$$

- ($x_1 < 0$ ou $x_0 = -1$) Nesse caso, $|x_1| = -x_1$ e $\theta = \tau - \phi$, portanto

$$\cos(\theta) = \cos(\tau - \phi) = \cos(-\phi) = \cos(\phi) = \frac{x_0}{\|x\|}$$

  e

$$\operatorname{sen}(\theta) = \operatorname{sen}(\tau - \phi) = \operatorname{sen}(-\phi) = -\operatorname{sen}(\phi) = -\frac{|x_1|}{\|x\|} = \frac{x_1}{\|x\|}.$$

Assim, concluímos que

$$x = x_0 + x_1\mathrm{i} = \|x\| \cos(\theta) + \|x\| \operatorname{sen}(\theta)\mathrm{i} = \|x\| \left(\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}\right). \qquad \blacksquare$$

10.4.1.1 Rotações em $\mathbb{R}^2$

$\vdash$ **Proposição 10.13.** *A função*

$$\begin{aligned} R\colon \mathbb{S}^1 &\longrightarrow \mathrm{SO}(2) \\ u &\longmapsto R_u\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ &\qquad\qquad\quad x \longmapsto ux \end{aligned}$$

é um difeomorfismo e isomorfismo de grupos.

□ *Demonstração.* A função R é diferenciável porque $R_u(x)$ é um polinômio em cada entrada. Primeiro, vamos mostrar que R de fato está bem definida, ou seja, que, para todo $u \in \mathbb{S}^1$, $R_u \in \mathrm{SO}(2)$. Seja $u \in \mathbb{S}^1$. A função R_u é linear, pois, para todos $x, x' \in \mathbb{R}^2$ e $c \in \mathbb{R}$, segue da bilinearidade da multiplicação complexa que

$$R_u(cx + c') = u(cx + c') = cux + ux' = cR_u(x) + R_u(x').$$

Isso mostra que $R_u \in \mathrm{GL}(2, \mathbb{R})$. Além disso, R_u preserva a norma, pois

$$\|R_u(x)\| = \|ux\| = \|u\| \, \|x\| = \|x\| \, .$$

Isso mostra que $R_u \in \mathrm{O}(2)$. Agora, como R é diferenciável, em particular é contínua, logo, como $\mathbb{S}^1$ é conexo, $R(\mathbb{S}^1)$ é conexo. Isso implica que $R(\mathbb{S}^1) \subseteq \mathrm{SO}(2)$.

Seja $u \in \mathbb{S}^3$ tal que $R_u = \mathrm{I}$. Então, para todo $x \in \mathbb{R}^2$, $ux = R_u(x) = x$. Em particular, $u = u1 = 1$, o que mostra que R é injetiva.

Para a sobrejetividade, primeiro notamos que, como $R(\mathbb{S}^1) \subseteq \mathrm{SO}(2)$ é compacto, pois $\mathbb{S}^1$ é compacto, e $\mathrm{SO}(2)$ é separado[2], segue que $R(\mathbb{S}^1)$ é fechado. Além disso, R é difeomorfismo local, pois o núcleo $\{1\}$ é um subgrupo discreto (portanto tem álgebra de Lie trivial), logo R é difeomorfismo local na identidade, e portanto em todo $\mathbb{S}^1$, já que um morfismo de grupos diferenciáveis[3] tem posto constante. Assim, segue que $R(\mathbb{S}^1) \subseteq \mathrm{SO}(2)$ é aberto. Como $\mathrm{SO}(2)$ é conexo, segue que $R(\mathbb{S}^1) = \mathrm{SO}(2)$ e a inversa de R é diferenciável. ■

Para calcular exatamente qual é a rotação que R_u representa em $\mathrm{SO}(2)$, lembremos que, pela decomposição polar de $\mathbb{R}^2$, todo $u \in \mathbb{S}^1$ pode ser escrito como

$$\mathrm{e}^{\theta \mathrm{i}} = \cos(\theta) + \mathrm{sen}(\theta)\mathrm{i},$$

em que $\theta \in [0, \tau]$.

A rotação de $x \in \mathbb{R}^2$ por um ângulo $\theta \in [0, \tau[$ (em torno da origem) é dada por

$$R^\theta = (\cos(\theta)x_0 - \mathrm{sen}(\theta)x_1) + (\mathrm{sen}(\theta)x_0 + \cos(\theta)x_1)\mathrm{i}.$$

[2] Hausdorff
[3] Grupos de lie

Assim segue que, para todo $x \in \mathbb{R}^2$,

$$\begin{aligned} R_{\mathrm{e}^{\theta\mathrm{i}}}(x) &= (\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i})(x_0 + x_1\mathrm{i}) \\ &= (\cos(\theta)x_0 - \operatorname{sen}(\theta)x_1) + (\operatorname{sen}(\theta)x_0 + \cos(\theta)x_1)\mathrm{i} \\ &= R^\theta(x), \end{aligned}$$

portanto $R(\mathrm{e}^{\theta\mathrm{i}}) = R^\theta$.

Isso mostra que

$$\begin{aligned} R\colon \mathbb{S}^1 &\longrightarrow \mathrm{SO}(2) \\ \mathrm{e}^{\theta\mathrm{i}} &\longmapsto R^\theta\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ &\qquad\qquad x \longmapsto \mathrm{e}^{\theta\mathrm{i}}x. \end{aligned}$$

10.4.2 Quatérnios $\mathbb{R}^4$

10.4.2.1 Multiplicação quaterniônica

Consideremos o espaço linear $\mathbb{R}^4$ com a norma 2

$$\begin{aligned} \|\cdot\|\colon \mathbb{R}^4 &\longrightarrow \mathbb{R} \\ x &\longmapsto ({x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2)^{\frac{1}{2}}. \end{aligned}$$

A base canônica de $\mathbb{R}^4$ será denotada por

$$\begin{aligned} \mathrm{i}_0 &:= (1,0,0,0) \\ \mathrm{i}_1 &:= (0,1,0,0) \\ \mathrm{i}_2 &:= (0,0,1,0) \\ \mathrm{i}_3 &:= (0,0,0,1). \end{aligned}$$

:⊢ **Definição 10.13.** A *multiplicação quaterniônica* em $\mathbb{R}^4$ é a operação

$$\begin{aligned} \times\colon \mathbb{R}^4 \times \mathbb{R}^4 &\longrightarrow \mathbb{R}^4 \\ (x,y) &\longmapsto xy := (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3, \\ &\qquad\qquad x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2, \\ &\qquad\qquad x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1, \\ &\qquad\qquad x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0). \end{aligned}$$

A *unidade* de $(\mathbb{R}^4, \times)$ é

$$1_{\mathbb{R}^4} := \mathrm{i}_0 = (1,0,0,0).$$

A *inversão* de $(\mathbb{R}^4, \times, 1)$ é a operação

$$\begin{aligned} \diagup\colon \mathbb{R}^4 \smallsetminus \{0\} &\longrightarrow \mathbb{R}^4 \smallsetminus \{0\} \\ x &\longmapsto x^{-1} := \|x\|^{-2}(x_0, -x_1, -x_2, -x_3). \end{aligned}$$

Em geral, por não haver ambiguidade, denotaremos $1_{\mathbb{R}^4}$ simplesmente por 1. Assim, todo $x \in \mathbb{R}^4$ pode ser escrito

$$x = x_0 1 + x_1 \mathrm{i}_1 + x_2 \mathrm{i}_2 + x_3 \mathrm{i}_3.$$

Como a notação acima sugere, omitiremos o símbolo de multiplicação $\times$ sempre que possível pois isso não gerará ambiguidade.

$\times$	1	i_1	i_2	i_3
1	1	i_1	i_2	i_3
i_1	i_1	-1	i_3	$-\mathrm{i}_2$
i_2	i_2	$-\mathrm{i}_3$	-1	i_1
i_3	i_3	i_2	$-\mathrm{i}_1$	-1

TABELA 10.1: Tabela de multiplicação dos quatérnios.

FIGURA 10.1: Grafo para memorizar multiplicação quaterniônica.

⊢ **Proposição 10.14.** *A lista*

$$((\mathbb{R}^4, +, -, 0, \cdot, \times, \diagup, 1), \|\cdot\|)$$

é uma álgebra[4] *(associativa) invertível normada sobre* $\mathbb{R}$.

1. *(Bilinearidade) A multiplicação quaterniônica* $\times \colon \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ *é uma função bilinear;*
2. *(Associatividade) Para todos* $x, y, z \in \mathbb{R}^4$,
$$(xy)z = x(yz);$$
3. *(Identidade) Para todo* $x \in \mathbb{R}^4$,
$$x1 = 1x = x;$$

[4]A álgebra de $\mathbb{R}^4$ com essa multiplicação costuma ser denotada por $\mathbb{H}$, mas manteremos a notação $\mathbb{R}^4$ por simplicidade de notação, pois nenhuma ambiguidade resultará disso. A letra 'H' é usada em homenagem ao matemático irlandês William Rowan Hamilton (1805 – 1865).

4. *(Inversa) Para todo* $x \in \mathbb{R}^4$,

$$x^{-1}x = 1 = xx^{-1};$$

5. *(Submultiplicatividade) Para todos* $x, y \in \mathbb{R}^4$,

$$\|xx'\| \leq \|x\| \, \|x'\| \, .$$

6. *(Unitariedade)* $\|1\| = 1$;
7. *(Invertibilidade) Para todo* $x \in \mathbb{R}^4$,

$$\left\|x^{-1}\right\| = \|x\|^{-1} \, .$$

□ *Demonstração.* Como $(\mathbb{R}^4, +, -, 0, \cdot)$ é um espaço linear sobre $\mathbb{R}$, basta mostrar as propriedades de $\times$.

- (Identidade) Para todo $x \in \mathbb{R}^4$,

$$\begin{aligned}
1x &= (1,0,0,0)(x_0, x_1, x_2, x_3) \\
&= (1x_0 - 0x_1 - 0x_2 - 0x_3)1 \\
&+ (1x_1 + 0x_0 + 0x_3 - 0x_2)\mathrm{i}_1 \\
&+ (1x_2 - 0x_3 + 0x_0 + 0x_1)\mathrm{i}_2 \\
&+ (1x_3 + 0x_2 - 0x_1 + 0x_0)\mathrm{i}_3 \\
&= (x_0, x_1, x_2, x_3) \\
&= x
\end{aligned}$$

  e

$$\begin{aligned}
x1 &= (x_0, x_1, x_2, x_3)(1,0,0,0) \\
&= (x_0 1 - x_1 0 - x_2 0 - x_3 0)1 \\
&+ (x_0 0 + x_1 1 + x_2 0 - x_3 0)\mathrm{i}_1 \\
&+ (x_0 0 - x_1 0 + x_2 1 + x_3 0)\mathrm{i}_2 \\
&+ (x_0 0 + x_1 0 - x_2 0 + x_3 1)\mathrm{i}_3 \\
&= (x_0, x_1, x_2, x_3) \\
&= x.
\end{aligned}$$

- (Bilinearidade) Para mostrar que o produto é bilinear, primeiro mostremos quais são seus valores na base $\{1, \mathrm{i}_2, \mathrm{i}_2, \mathrm{i}_3\}$. Como 1 é identidade do produto, basta calcular os valores do produto i_1, i_2 e i_3. Cálculos simples mostram que

$$\mathrm{i}_1\mathrm{i}_1 = -1 \qquad\qquad \mathrm{i}_1\mathrm{i}_2 = \mathrm{i}_3 = -\mathrm{i}_2\mathrm{i}_1$$

$$\begin{aligned} \mathrm{i}_2\mathrm{i}_2 &= -1 & \mathrm{i}_2\mathrm{i}_3 &= \mathrm{i}_1 = -\mathrm{i}_3\mathrm{i}_2 \\ \mathrm{i}_3\mathrm{i}_3 &= -1 & \mathrm{i}_3\mathrm{i}_1 &= \mathrm{i}_2 = -\mathrm{i}_1\mathrm{i}_3 \end{aligned}$$

e, sendo assim, basta mostrar que uma função bilinear $b\colon \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ com esses mesmos valores na base canônica coincide com o produto quaterniônico definido. Da bilinearidade segue que

$$\begin{aligned} b(x,y) &= b(x_0 1 + x_1\mathrm{i}_1 + x_2\mathrm{i}_2 + x_3\mathrm{i}_3, y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &= x_0 b(1, y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &+ x_1 b(\mathrm{i}_1, y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &+ x_2 b(\mathrm{i}_2, y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &+ x_3 b(\mathrm{i}_3, y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &= x_0(y_0 b(1,1) + y_1 b(1,\mathrm{i}_1) + y_2 b(1,\mathrm{i}_2) + y_3 b(1,\mathrm{i}_3)) \\ &+ x_1(y_0 b(\mathrm{i}_1,1) + y_1 b(\mathrm{i}_1,\mathrm{i}_1) + y_2 b(\mathrm{i}_1,\mathrm{i}_2) + y_3 b(\mathrm{i}_1,\mathrm{i}_3)) \\ &+ x_2(y_0 b(\mathrm{i}_2,1) + y_1 b(\mathrm{i}_2,\mathrm{i}_1) + y_2 b(\mathrm{i}_2,\mathrm{i}_2) + y_3 b(\mathrm{i}_2,\mathrm{i}_3)) \\ &+ x_3(y_0 b(\mathrm{i}_3,1) + y_1 b(\mathrm{i}_3,\mathrm{i}_1) + y_2 b(\mathrm{i}_3,\mathrm{i}_2) + y_3 b(\mathrm{i}_3,\mathrm{i}_3)) \\ &= x_0(y_0 1 + y_1\mathrm{i}_1 + y_2\mathrm{i}_2 + y_3\mathrm{i}_3) \\ &+ x_1(y_0\mathrm{i}_1 + y_1(-1) + y_2\mathrm{i}_3 + y_3(-\mathrm{i}_2)) \\ &+ x_2(y_0\mathrm{i}_2 + y_1(-\mathrm{i}_3) + y_2(-1) + y_3\mathrm{i}_1) \\ &+ x_3(y_0\mathrm{i}_3 + y_1\mathrm{i}_2 + y_2(-\mathrm{i}_1) + y_3(-1)) \\ &= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)1 \\ &+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathrm{i}_1 \\ &+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathrm{i}_2 \\ &+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathrm{i}_3 \\ &= xy. \end{aligned}$$

- (Associatividade) Faremos as contas por coordenadas. Para todos $x, y, z \in \mathbb{R}^4$, segue da definição do produto que as coordenadas de xy são

$$\begin{aligned} (xy)_0 &= x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3 \\ (xy)_1 &= x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2 \\ (xy)_2 &= x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1 \\ (xy)_3 &= x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0 \end{aligned}$$

e que as coordenadas de yz são

$$\begin{aligned} (yz)_0 &= y_0z_0 - y_1z_1 - y_2z_2 - y_3z_3 \\ (yz)_1 &= y_0z_1 + y_1z_0 + y_2z_3 - y_3z_2 \end{aligned}$$

$$
\begin{aligned}
(yz)_2 &= y_0z_2 - y_1z_3 + y_2z_0 + y_3z_1 \\
(yz)_3 &= y_0z_3 + y_1z_2 - y_2z_1 + y_3z_0.
\end{aligned}
$$

Ainda da definição do produto, as coordenadas de $(xy)z$ são

$$
\begin{aligned}
((xy)z)_0 &= (xy)_0z_0 - (xy)_1z_1 - (xy)_2z_2 - (xy)_3z_3 \\
((xy)z)_1 &= (xy)_0z_1 + (xy)_1z_0 + (xy)_2z_3 - (xy)_3z_2 \\
((xy)z)_2 &= (xy)_0z_2 - (xy)_1z_3 + (xy)_2z_0 + (xy)_3z_1 \\
((xy)z)_3 &= (xy)_0z_3 + (xy)_1z_2 - (xy)_2z_1 + (xy)_3z_0
\end{aligned}
$$

e as coordenadas de $x(yz)$ são

$$
\begin{aligned}
(x(yz))_0 &= x_0(yz)_0 - x_1(yz)_1 - x_2(yz)_2 - x_3(yz)_3 \\
(x(yz))_1 &= x_0(yz)_1 + x_1(yz)_0 + x_2(yz)_3 - x_3(yz)_2 \\
(x(yz))_2 &= x_0(yz)_2 - x_1(yz)_3 + x_2(yz)_0 + x_3(yz)_1 \\
(x(yz))_3 &= x_0(yz)_3 + x_1(yz)_2 - x_2(yz)_1 + x_3(yz)_0.
\end{aligned}
$$

Para a coordenada 0, segue que

$$
\begin{aligned}
((xy)z)_0 &= (xy)_0z_0 - (xy)_1z_1 - (xy)_2z_2 - (xy)_3z_3 \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)z_0 \\
&+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)z_1 \\
&+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)z_2 \\
&+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)z_3 \\
&= x_0(y_0z_0 - y_1z_1 - y_2z_2 - y_3z_3) \\
&+ x_1(y_0z_1 + y_1z_0 + y_2z_3 - y_3z_2) \\
&+ x_2(y_0z_2 - y_1z_3 + y_2z_0 + y_3z_1) \\
&+ x_3(y_0z_3 + y_1z_2 - y_2z_1 + y_3z_0) \\
&= x_0(yz)_0 + x_1(yz)_1 + x_2(yz)_2 + x_3(yz)_3 \\
&= (x(yz))_0;
\end{aligned}
$$

para a coordenada 1, segue que

$$
\begin{aligned}
((xy)z)_1 &= (xy)_0z_1 + (xy)_1z_0 + (xy)_2z_3 - (xy)_3z_2 \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)z_1 \\
&+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)z_0 \\
&+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)z_3
\end{aligned}
$$

$$
\begin{aligned}
&- (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)z_2 \\
&= x_0(y_0z_1 + y_1z_0 + y_2z_3 - y_3z_2) \\
&+ x_1(y_0z_0 - y_1z_1 - y_2z_2 - y_3z_3) \\
&+ x_2(y_0z_3 + y_1z_2 - y_2z_1 + y_3z_0) \\
&- x_3(y_0z_2 - y_1z_3 + y_2z_0 + y_3z_1) \\
&= x_0(yz)_1 + x_1(yz)_0 + x_2(yz)_3 - x_3(yz)_2 \\
&= (x(yz))_1;
\end{aligned}
$$

para a coordenada 2, segue que

$$
\begin{aligned}
((xy)z)_2 &= (xy)_0z_2 - (xy)_1z_3 + (xy)_2z_0 + (xy)_3z_1 \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)z_2 \\
&- (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)z_3 \\
&+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)z_0 \\
&+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)z_1 \\
&= x_0(y_0z_2 - y_1z_3 + y_2z_0 + y_3z_1) \\
&- x_1(y_0z_3 + y_1z_2 - y_2z_1 + y_3z_0) \\
&+ x_2(y_0z_0 - y_1z_1 - y_2z_2 - y_3z_3) \\
&+ x_3(y_0z_1 + y_1z_0 + y_2z_3 - y_3z_2) \\
&= x_0(yz)_2 - x_1(yz)_3 + x_2(yz)_0 + x_3(yz)_1 \\
&= (x(yz))_2;
\end{aligned}
$$

e, para a coordenada 3, segue que

$$
\begin{aligned}
((xy)z)_3 &= (xy)_0z_3 + (xy)_1z_2 - (xy)_2z_1 + (xy)_3z_0 \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)z_3 \\
&+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)z_2 \\
&- (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)z_1 \\
&+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)z_0 \\
&= x_0(y_0z_3 + y_1z_2 - y_2z_1 + y_3z_0) \\
&+ x_1(y_0z_2 - y_1z_3 + y_2z_0 + y_3z_1) \\
&- x_2(y_0z_1 + y_1z_0 + y_2z_3 - y_3z_2) \\
&+ x_3(y_0z_0 - y_1z_1 - y_2z_2 - y_3z_3) \\
&= x_0(yz)_3 + x_1(yz)_2 - x_2(yz)_1 + x_3(yz)_0 \\
&= (x(yz))_3.
\end{aligned}
$$

Portanto

$$((xy)z) = ((xy)z)_0 1 + ((xy)z)_1 \mathrm{i}_1 + ((xy)z)_2 \mathrm{i}_2 + ((xy)z)_3 \mathrm{i}_3$$

$$
\begin{aligned}
&= (x(yz))_0 1 + (x(yz))_1 \mathrm{i}_1 + (x(yz))_2 \mathrm{i}_2 + (x(yz))_3 \mathrm{i}_3 \\
&= (x(yz)).
\end{aligned}
$$

- (Inversa) Para todo $x \in \mathbb{R}^4 \setminus \{0\}$,

$$
\begin{aligned}
x^{-1}x &= \|x\|^{-2} (x_0, -x_1, -x_2, -x_3)(x_0, x_1, x_2, x_3) \\
&= \|x\|^{-2} (x_0x_0 - (-x_1)x_1 - (-x_2)x_2 - (-x_3)x_3)1 \\
&+ \|x\|^{-2} (x_0x_1 + (-x_1)x_0 + (-x_2)x_3 - (-x_3)x_2)\mathrm{i}_1 \\
&+ \|x\|^{-2} (x_0x_2 - (-x_1)x_3 + (-x_2)x_0 + (-x_3)x_1)\mathrm{i}_2 \\
&+ \|x\|^{-2} (x_0x_3 + (-x_1)x_2 - (-x_2)x_1 + (-x_3)x_0)\mathrm{i}_3 \\
&= \|x\|^{-2} ({x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2)1 \\
&= \|x\|^{-2} \|x\|^2 1 \\
&= 1
\end{aligned}
$$

e

$$
\begin{aligned}
xx^{-1} &= (x_0, x_1, x_2, x_3) \|x\|^{-2} (x_0, -x_1, -x_2, -x_3) \\
&= \|x\|^{-2} (x_0x_0 - x_1(-x_1) - x_2(-x_2) - x_3(-x_3))1 \\
&+ \|x\|^{-2} (x_0(-x_1) + x_1x_0 + x_2(-x_3) - x_3(-x_2))\mathrm{i}_1 \\
&+ \|x\|^{-2} (x_0(-x_2) - x_1(-x_3) + x_2x_0 + x_3(-x_1))\mathrm{i}_2 \\
&+ \|x\|^{-2} (x_0(-x_3) + x_1(-x_2) - x_2(-x_1) + x_3x_0)\mathrm{i}_3 \\
&= \|x\|^{-2} ({x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2)1 \\
&= \|x\|^{-2} \|x\|^2 1 \\
&= 1.
\end{aligned}
$$

- (Submultiplicatividade) Demonstramos de fato a multiplicatividade da norma, que implica a submultiplicatividade. Para todos $x, y \in \mathbb{R}^4$,

$$
\begin{aligned}
\|xy\|^2 &= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)^2 \\
&+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)^2 \\
&+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)^2 \\
&+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)^2 \\
&= (x_0y_0)^2 + (x_1y_1)^2 + (x_2y_2)^2 + (x_3y_3)^2 \\
&- 2(x_0y_0x_1y_1 + x_0y_0x_2y_2 + x_0y_0x_3y_3 \\
&\qquad - x_1y_1x_2y_2 - x_1y_1x_3y_3 - x_2y_2x_3y_3) \\
&+ (x_0y_1)^2 + (x_1y_0)^2 + (x_2y_3)^2 + (x_3y_2)^2
\end{aligned}
$$

$$
\begin{aligned}
&+ 2(x_0y_1x_1y_0 + x_0y_1x_2y_3 - x_0y_1x_3y_2 \\
&\qquad + x_1y_0x_2y_3 - x_1y_0x_3y_2 - x_2y_3x_3y_2) \\
&+ (x_0y_2)^2 + (x_1y_3)^2 + (x_2y_0)^2 + (x_3y_1)^2 \\
&- 2(x_0y_2x_1y_3 - x_0y_2x_2y_0 - x_0y_2x_3y_1 \\
&\qquad + x_1y_3x_2y_0 + x_1y_3x_3y_1 - x_2y_0x_3y_1) \\
&+ (x_0y_3)^2 + (x_1y_2)^2 + (x_2y_1)^2 + (x_3y_0)^2 \\
&+ 2(x_0y_3x_1y_2 - x_0y_3x_2y_1 + x_0y_3x_3y_0 \\
&\qquad - x_1y_2x_2y_1 + x_1y_2x_3y_0 - x_2y_1x_3y_0) \\
&= (x_0y_0)^2 + (x_1y_1)^2 + (x_2y_2)^2 + (x_3y_3)^2 \\
&+ (x_0y_1)^2 + (x_1y_0)^2 + (x_2y_3)^2 + (x_3y_2)^2 \\
&+ (x_0y_2)^2 + (x_1y_3)^2 + (x_2y_0)^2 + (x_3y_1)^2 \\
&+ (x_0y_3)^2 + (x_1y_2)^2 + (x_2y_1)^2 + (x_3y_0)^2 \\
&= {x_0}^2({y_0}^2 + {y_1}^2 + {y_2}^2 + {y_3}^2) \\
&+ {x_1}^2({y_0}^2 + {y_1}^2 + {y_2}^2 + {y_3}^2) \\
&+ {x_2}^2({y_0}^2 + {y_1}^2 + {y_2}^2 + {y_3}^2) \\
&+ {x_3}^2({y_0}^2 + {y_1}^2 + {y_2}^2 + {y_3}^2) \\
&= ({x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2)({y_0}^2 + {y_1}^2 + {y_2}^2 + {y_3}^2) \\
&= \|x\|^2 \|y\|^2 .
\end{aligned}
$$

- (Unitariedade)

$$\|1\| = (1^2 + 0^2 + 0^2 + 0^2)^{\frac{1}{2}} = 1.$$

- (Invertibilidade) Para todo $x \in \mathbb{R}^4 \smallsetminus \{0\}$,

$$
\begin{aligned}
\left\|x^{-1}\right\| &= \left(\left(\frac{x_0}{\|x\|^2}\right)^2 + \left(\frac{-x_1}{\|x\|^2}\right)^2 + \left(\frac{-x_2}{\|x\|^2}\right)^2 + \left(\frac{-x_3}{\|x\|^2}\right)^2\right)^{\frac{1}{2}} \\
&= \frac{1}{\|x\|^2}((x_0)^2 + (-x_1)^2 + (-x_2)^2 + (-x_3)^2)^{\frac{1}{2}} \\
&= \frac{\|x\|}{\|x\|^2} \\
&= \|x\|^{-1} .
\end{aligned}
$$

■

Demonstramos de fato a multiplicatividade da norma, mas a submultiplicatividade, junto às outras propriedades da norma, implica que, para todos $x, x' \in \mathbb{R}^4$,

$$\|xx'\| = \|x\| \|x'\| .$$

Como o vetor $1 = (1, 0, 0, 0)$ é uma unidade, em geral o omitiremos da notação da base, denotando $x_0 1$ por x_0 e

$$x = x_0 + x_1 \mathrm{i}_1 + x_2 \mathrm{i}_2 + x_3 \mathrm{i}_3.$$

:⊢ **Definição 10.14.** Seja $x \in \mathbb{R}^4$. A *componente escalar* de x é (o número) $x_0 := x_0 1$ e a *componente vetorial* de x é (o vetor)

$$\vec{x} := x_1 \mathrm{i}_1 + x_2 \mathrm{i}_2 + x_3 \mathrm{i}_3.$$

O *subespaço de escalares* de $\mathbb{R}^4$ é o subespaço

$$\mathbb{R}_0 := \left\{ x \in \mathbb{R}^4 \;\middle|\; \vec{x} = 0 \right\}$$

e o *subespaço de vetores* de $\mathbb{R}^4$ é o subespaço

$$\vec{\mathbb{R}}^3 := \left\{ x \in \mathbb{R}^4 \;\middle|\; x_0 = 0 \right\}.$$

A *esfera de vetores* de $\mathbb{R}^4$ é o conjunto

$$\vec{\mathbb{S}}^2 := \left\{ x \in \mathbb{S}^3 \;\middle|\; x_0 = 0 \right\} = \mathbb{S}^3 \cap \vec{\mathbb{R}}^3.$$

Isso nos permite escrever todo quatérnio $x \in \mathbb{R}^4$ unicamente como

$$x = x_0 + \vec{x}$$

e decompor $\mathbb{R}^4$ como a soma direta

$$\mathbb{R}^4 = \mathbb{R}_0 + \vec{\mathbb{R}}^3.$$

:⊢ **Definição 10.15.** A *conjugação* em $\mathbb{R}^4$ é a operação

$$\begin{aligned} {}^*\colon \mathbb{R}^4 &\longrightarrow \mathbb{R}^4 \\ x &\longmapsto x^* := x_0 - \vec{x}. \end{aligned}$$

▷ **Exercício 10.5.** *1. (Involutividade) Para todo $x \in \mathbb{R}^4$, $(x^*)^* = x$;*
2. (Aditividade) Para todos $x, x' \in \mathbb{R}^4$, $(x + x')^* = x^* + (x')^*$;
3. (Identidade aditiva) $0^* = 0$;
4. (Invertibilidade aditiva) Para todo $x \in \mathbb{R}^4$, $(-x)^* = -x^*$;
5. (Anti-multiplicatividade) Para todos $x, x' \in \mathbb{R}^4$, $(xx')^* = (x')^* x^*$;
6. (Identidade multiplicativa) $1^* = 1$;

▷ **Exercício 10.6.** *Seja $x \in \mathbb{R}^4$.*

1. $\|x\|^2 = |x_0|^2 + \|\vec{x}\|^2$;

2. $\|x\|^2 = xx^*$;
3. $x^* = \frac{1}{2}(1x(-1) + \mathrm{i}_1 x(-\mathrm{i}_1) + \mathrm{i}_2 x(-\mathrm{i}_2) + \mathrm{i}_3 x(-\mathrm{i}_3))$.

⊢ **Proposição 10.15** (Decomposição polar)**.** *Seja $x \in \mathbb{R}^4 \setminus \{0\}$. Existem únicos $\phi_0 \in [0, \tau \diagup 2]$ e $\hat{x} \in \vec{\mathbb{S}}^2$ tais que*

$$x = \|x\| \left(\cos(\phi_0) + \operatorname{sen}(\phi_0)\hat{x}\right).$$

Eles são dados por $\phi_0 := \cos^{-1}\left(\frac{x_0}{\|x\|}\right) \in [0, \tau \diagup 2]$ e $\hat{x} := \frac{\vec{x}}{\|\vec{x}\|} \in \vec{\mathbb{S}}^2$.

□ *Demonstração.* Mostraremos primeiro a decomposição. Como $x \neq 0$, $\|x\| \neq 0$, logo da igualdade $\|x\|^2 = |x_0|^2 + \|\vec{q}\|^2$ segue que

$$1 = \left(\frac{|x_0|}{\|x\|}\right)^2 + \left(\frac{\|\vec{x}\|}{\|x\|}\right)^2,$$

portanto $\frac{|x_0|}{\|x\|} \leq 1$, o que é equivalente a $\frac{x_0}{\|x\|} \in [-1, 1]$. Disso segue que $\phi_0 = \cos^{-1}\left(\frac{x_0}{\|x\|}\right) \in [0, \tau \diagup 2]$, de modo que $\frac{\|\vec{x}\|}{\|x\|} = (1 - (\cos(\phi_0))^2)^{1 \diagup 2}$, o que implica $\frac{\|\vec{x}\|}{\|x\|} = \operatorname{sen}(\phi_0)$ ou $\frac{\|\vec{x}\|}{\|x\|} = -\operatorname{sen}(\phi_0)$. Como para $\phi_0 \in [0, \tau \diagup 2]$ vale $\operatorname{sen}(\phi_0) \geq 0$, concluímos que $\frac{\|\vec{x}\|}{\|x\|} = \operatorname{sen}(\phi_0)$.

Das definições de ϕ_0 e $\hat{x}$, segue que $x_0 = \|x\| \cos(\phi_0)$ e $\vec{x} = \|x\| \operatorname{sen}(\phi_0)\hat{x}$, portanto

$$\begin{aligned} x &= x_0 + \vec{x} \\ &= \|x\| \cos(\phi_0) + \|x\| \operatorname{sen}(\phi_0)\hat{x} \\ &= \cos(\phi_0) + \operatorname{sen}(\phi_0)\frac{\vec{x}}{\|\vec{x}\|} \\ &= \|x\| \left(\cos(\phi_0) + \operatorname{sen}(\phi_0)\hat{x}\right). \end{aligned}$$

■

Para a unicidade, sejam $\phi_0' \in [0, \tau \diagup 2]$ e $\hat{x}' \in \vec{\mathbb{S}}^2$ tais que

$$x = \|x\| \left(\cos(\phi_0) + \operatorname{sen}(\phi_0)\hat{x}\right) = \|x\| \left(\cos(\phi_0') + \operatorname{sen}(\phi_0')\hat{x}'\right).$$

Então $\cos(\phi_0) = \cos(\phi_0')$ e $\operatorname{sen}(\phi_0)\hat{x} = \operatorname{sen}(\phi_0')\hat{x}'$. Da primeira igualdade segue que existe $n \in \mathbb{Z}$ tal que $\phi_0' = \phi_0 + n\tau$ ou $\phi_0' = -\phi_0 + n\tau$. Como $\phi_0, \phi_0' \in [0, \tau \diagup 2]$, segue que $\phi_0 = \phi_0'$. Isso implica que $\operatorname{sen}(\phi_0) = \operatorname{sen}(\phi_0')$, portanto da segunda igualdade segue que $\hat{x} = \hat{x}'$.

10.4.2.2 Produtos interno, escalar e vetorial

:⊢ **Definição 10.16.** O *produto interno quaterniônico* em $\mathbb{R}^4$ é a função

$$\begin{aligned}\langle\cdot,\cdot\rangle_{1,3}\colon \mathbb{R}^4\times\mathbb{R}^4 &\longrightarrow \mathbb{R}\\ (x,y) &\longmapsto \langle x,y\rangle_{1,3} := x_0y_0-(x_1y_1+x_2y_2+x_3y_3).\end{aligned}$$

O *produto escalar* em $\vec{\mathbb{R}}^3$ é a função

$$\begin{aligned}\cdot : \vec{\mathbb{R}}^3\times\vec{\mathbb{R}}^3 &\longrightarrow \mathbb{R}\\ (\vec{x},\vec{y}) &\longmapsto \vec{x}\cdot\vec{y} := x_1y_1+x_2y_2+x_3y_3.\end{aligned}$$

O *produto vetorial* em $\vec{\mathbb{R}}^3$ é a função

$$\begin{aligned}\curlywedge : \vec{\mathbb{R}}^3\times\vec{\mathbb{R}}^3 &\longrightarrow \vec{\mathbb{R}}^3\\ (\vec{x},\vec{y}) &\longmapsto \vec{x}\curlywedge\vec{y} := (x_2y_3-x_3y_2)\mathrm{i}_1\\ &\qquad\qquad +(-x_1y_3+x_3y_1)\mathrm{i}_2\\ &\qquad\qquad +(x_1y_2-x_2y_1)\mathrm{i}_3.\end{aligned}$$

O produto interno quaterniônico difere do produto interno usual em $\mathbb{R}^4$, mas os outros dois produtos são o produto interno e o produto vetorial usuais em $\mathbb{R}^3$.

$\curlywedge$	i_1	i_2	i_3
i_1	0	i_3	$-\mathrm{i}_2$
i_2	$-\mathrm{i}_3$	0	i_1
i_3	i_2	$-\mathrm{i}_1$	0

Tabela 10.2: Tabela de produto vetorial em $\vec{\mathbb{R}}^3$.

▷ **Exercício 10.7.** *1. Para todos $x,y\in\mathbb{R}^4$,*

$$\langle x,y\rangle_{1,3} = x_0y_0-\vec{x}\cdot\vec{y}.$$

2. O produto escalar $\cdot : \vec{\mathbb{R}}^3\times\vec{\mathbb{R}}^3\to\mathbb{R}$ é um produto interno no subespaço vetorial $\vec{\mathbb{R}}^3$;

3. O produto vetorial $\curlywedge : \vec{\mathbb{R}}^3\times\vec{\mathbb{R}}^3\to\vec{\mathbb{R}}^3$ é uma função bilinear alternada tal que, para todos $\vec{x},\vec{y}\in\vec{\mathbb{R}}^3$,

3.1. $(\vec{x}\curlywedge\vec{y})\cdot\vec{x} = (\vec{x}\curlywedge\vec{y})\cdot\vec{y} = 0$;

3.2. $(\vec{x}\curlywedge\vec{y})\cdot(\vec{x}\curlywedge\vec{y}) = \det\begin{bmatrix}\vec{x}\cdot\vec{x} & \vec{x}\cdot\vec{y}\\ \vec{y}\cdot\vec{x} & \vec{y}\cdot\vec{y}\end{bmatrix}$.

4. *Para todos* $\vec{x}, \vec{y}, \vec{z} \in \vec{\mathbb{R}}^3$,

$$(\vec{x} \curlywedge \vec{y}) \curlywedge \vec{z} = (\vec{x} \cdot \vec{z})\vec{y} - (\vec{y} \cdot \vec{z})\vec{x};$$

5. *A função*

$$\begin{aligned} \mathrm{vol}_{\curlywedge} \colon \vec{\mathbb{R}}^3 \times \vec{\mathbb{R}}^3 \times \vec{\mathbb{R}}^3 &\longrightarrow \vec{\mathbb{R}}^3 \\ (\vec{x}, \vec{y}, \vec{z}) &\longmapsto (\vec{x} \curlywedge \vec{y}) \cdot \vec{z} \end{aligned}$$

é uma função trilinear alternada.

⊢ **Proposição 10.16.** 1. *Para todos* $x, y \in \vec{\mathbb{R}}^3$ *(*$x_0 = y_0 = 0$*),*

$$xy = -\vec{x} \cdot \vec{y} + \vec{x} \curlywedge \vec{y};$$

2. *Para todos* $x, y \in \mathbb{R}^4$,

$$xy = x_0 y_0 - \vec{x} \cdot \vec{y} + x_0 \vec{y} + y_0 \vec{x} + \vec{x} \curlywedge \vec{y},$$

sendo

$$\begin{aligned} (xy)_0 &= x_0 y_0 - \vec{x} \cdot \vec{y}, \\ \vec{(xy)} &= x_0 \vec{y} + y_0 \vec{x} + \vec{x} \curlywedge \vec{y}. \end{aligned}$$

3. *Para todos* $x, y \in \mathbb{R}^4$, $xy = yx$ *se, e somente se,* $\vec{x} \curlywedge \vec{y} = 0$*;*
4. *Para todo* $x \in \mathbb{R}^4 \smallsetminus \{0\}$,

$$x^{-1} = \frac{x^*}{\|x\|^2};$$

5. $\vec{\mathbb{S}}^2 = \{x \in \mathbb{R}^4 \mid x^2 = -1\}$*;*
6. *Para todos* $x \in \mathbb{R}^4 \smallsetminus \{0\}$ *e* $y \in \mathbb{R}^4$,

$$xyx^{-1} = y_0 + 2\frac{\vec{x} \cdot \vec{y}}{\|x\|^2}\vec{x} + \frac{\langle x, x \rangle_{1,3}}{\|x\|^2}\vec{y} + 2\frac{x_0}{\|x\|^2}(\vec{x} \curlywedge \vec{y}).$$

7. *Para todos* $x, y \in \mathbb{R}^4 \smallsetminus \{0\}$ *tais que* $xy \in \mathbb{R}_0$, $\vec{x} \parallel \vec{y}$.

□ *Demonstração.* 1. Como $x_0 = y_0 = 0$,

$$\begin{aligned} xy = &- (x_1 y_1 + x_2 y_2 + x_3 y_3) \\ &+ (x_2 y_3 - x_3 y_2)\mathrm{i}_1 \\ &+ (-x_1 y_3 + x_3 y_1)\mathrm{i}_2 \\ &+ (x_1 y_2 - x_2 y_1)\mathrm{i}_3 \\ = &- \vec{x} \cdot \vec{y} + \vec{x} \curlywedge \vec{y}. \end{aligned}$$

2. Segue da bilinearidade do produto e de $\vec{x}$ e $\vec{y}$ serem puramente vetoriais que

$$\begin{aligned} xy &= (x_0 + \vec{x})(y_0 + \vec{y}) \\ &= x_0 y_0 + x_0 \vec{y} + y_0 \vec{x} + \vec{x}\vec{y} \\ &= x_0 y_0 + x_0 \vec{y} + y_0 \vec{x} + (-\vec{x} \cdot \vec{y} + \vec{x} \curlywedge \vec{y}) \\ &= (x_0 y_0 - \vec{x} \cdot \vec{y}) + (x_0 \vec{y} + y_0 \vec{x} + \vec{x} \curlywedge \vec{y}). \end{aligned}$$

3. Para todos $x, y \in \mathbb{R}^4$,

$$\begin{aligned} xy - yx &= x_0 y_0 - \vec{x} \cdot \vec{y} + x_0 \vec{y} + y_0 \vec{x} + \vec{x} \curlywedge \vec{y} \\ &\quad - y_0 x_0 + \vec{y} \cdot \vec{x} - y_0 \vec{x} - x_0 \vec{y} - \vec{y} \curlywedge \vec{x} \\ &= \vec{x} \curlywedge \vec{y} - \vec{y} \curlywedge \vec{x} \\ &= 2\vec{x} \curlywedge \vec{y}. \end{aligned}$$

   Isso implica que $xy - yx = 0$ se, e somente se, $\vec{x} \curlywedge \vec{y} - \vec{y} \curlywedge \vec{x} = 0$, o que ocorre se, e somente se, $\vec{x} \curlywedge \vec{y} = 0$.

4. Para todo $x \in \mathbb{R}^4 \setminus \{0\}$,

$$\begin{aligned} xx^{-1} &= (x_0 + \vec{x}) \left\|x\right\|^{-2} (x_0 - \vec{x}) \\ &= \left\|x\right\|^{-2} (x_0^2 - \vec{x} \cdot (-\vec{x}) + x_0(-\vec{x}) + x_0 \vec{x} + \vec{x} \curlywedge (-\vec{x})) \\ &= \left\|x\right\|^{-2} (x_0^2 + \vec{x} \cdot \vec{x} - x_0 \vec{x} + x_0 \vec{x} - \vec{x} \curlywedge \vec{x}) \\ &= \left\|x\right\|^{-2} \left\|x\right\|^2 \\ &= 1 \end{aligned}$$

   e, como $\vec{x} \curlywedge (-\vec{x}) = 0$, $x^{-1}x = xx^{-1} = 1$.

5. Para todo $x \in \mathbb{R}^4$,

$$\begin{aligned} x^2 &= x_0 x_0 - \vec{x} \cdot \vec{x} + x_0 \vec{x} + x_0 \vec{x} + \vec{x} \curlywedge \vec{x} \\ &= x_0^2 - \left\|\vec{x}\right\|^2 + 2x_0 \vec{x}. \end{aligned}$$

   Segue que $x_0 = 0$ e $\left\|\vec{x}\right\| = 1$ se, e somente se, $x^2 = -1$.

6. Para todos $x, y \in \mathbb{R}^4$,

$$\begin{aligned} (xyx^{-1})_0 &= (xy)_0 (x^{-1})_0 - (\vec{xy}) \cdot (\vec{x^{-1}}) \\ &= (x_0 y_0 - \vec{x} \cdot \vec{y}) \frac{x_0}{\left\|x\right\|^2} - (x_0 \vec{y} + y_0 \vec{x} + \vec{x} \curlywedge \vec{y}) \cdot \frac{(-\vec{x})}{\left\|x\right\|^2} \\ &= \frac{x_0^2 y_0 - x_0(\vec{x} \cdot \vec{y}) + x_0 \vec{y} \cdot \vec{x} + y_0 \vec{x} \cdot \vec{x} + (\vec{x} \curlywedge \vec{y}) \cdot \vec{x}}{\left\|x\right\|^2} \\ &= \frac{x_0^2 y_0 + \left\|\vec{x}\right\|^2 y_0}{\left\|x\right\|^2} \\ &= y_0 \end{aligned}$$

e

$$
\begin{aligned}
(xy\vec{x^{-1}}) &= (xy)_0(\vec{x^{-1}}) + (x^{-1})_0(\vec{xy}) + (\vec{xy}) \curlywedge (\vec{x^{-1}}) \\
&= (x_0y_0 - \vec{x}\cdot\vec{y})\frac{(-\vec{x})}{\|x\|^2} + \frac{x_0}{\|x\|^2}(x_0\vec{y} + y_0\vec{x} + \vec{x}\curlywedge\vec{y}) \\
&\quad + (x_0\vec{y} + y_0\vec{x} + \vec{x}\curlywedge\vec{y}) \curlywedge \frac{(-\vec{x})}{\|x\|^2} \\
&= \frac{(\vec{x}\cdot\vec{y})\vec{x} + x_0^2\vec{y} + x_0(\vec{x}\curlywedge\vec{y})}{\|x\|^2} \\
&\quad - \frac{x_0(\vec{y}\curlywedge\vec{x}) + (\vec{x}\curlywedge\vec{y})\curlywedge\vec{x}}{\|x\|^2} \\
&= \frac{(\vec{x}\cdot\vec{y})\vec{x} + x_0^2\vec{y} + 2x_0(\vec{x}\curlywedge\vec{y}) - (\vec{x}\cdot\vec{x})\vec{y} + (\vec{y}\cdot\vec{x})\vec{x}}{\|x\|^2} \\
&= \frac{2(\vec{x}\cdot\vec{y})\vec{x} + (x_0^2 - \|\vec{x}\|^2)\vec{y} + 2x_0(\vec{x}\curlywedge\vec{y})}{\|x\|^2} \\
&= 2\frac{\vec{x}\cdot\vec{y}}{\|x\|^2}\vec{x} + \frac{\langle x, x\rangle_{1,3}}{\|x\|^2}\vec{y} + 2\frac{x_0}{\|x\|^2}(\vec{x}\curlywedge\vec{y}).
\end{aligned}
$$

Portanto

$$
xyx^{-1} = y_0 + 2\frac{\vec{x}\cdot\vec{y}}{\|x\|^2}\vec{x} + \frac{\langle x, x\rangle_{1,3}}{\|x\|^2}\vec{y} + 2\frac{x_0}{\|x\|^2}(\vec{x}\curlywedge\vec{y}).
$$

7. Sejam $x, y \in \mathbb{R}^4 \setminus \{0\}$ tais que $xy \in \mathbb{R}_0$. Então

$$
0 = x_0\vec{y} + y_0\vec{x} + \vec{x}\curlywedge\vec{y}.
$$

Isso implica que $(\vec{x}, \vec{y}, \vec{x}\curlywedge\vec{y})$ não é uma base de $\vec{\mathbb{R}}^3$, pois caso contrário

$$
x_0\vec{y} + y_0\vec{x} + \vec{x}\times\vec{y} \neq 0,
$$

já que $(x_0, y_0, 1) \neq (0, 0, 0)$. Mas a tripla não é base se, e somente se, $\vec{x} = 0$ ou $\vec{y} = 0$ ou $\vec{x}\curlywedge\vec{y} = 0$, porque sempre vale $\vec{x}\cdot(\vec{x}\curlywedge\vec{y}) = \vec{y}\cdot(\vec{x}\curlywedge\vec{y}) = 0$. Consideramos cada caso.

- $(\vec{x} = 0)$ Nesse caso, segue que $x_0 \neq 0$, pois $x \neq 0$, e que $\vec{x}\times\vec{y} = 0$, logo

$$
0 = x_0\vec{y}.
$$

  o que implica que $\vec{y} = 0$, portanto $\vec{x} \parallel \vec{y}$ e $x, y \in \mathbb{R}_0$.

- ($\vec{y} = 0$) Nesse caso, segue que $y_0 \neq 0$, pois $y \neq 0$, e que $\vec{x} \times \vec{y} = 0$, logo

$$0 = y_0\vec{x},$$

  o que implica que $\vec{x} = 0$, portanto $\vec{x} \parallel \vec{y}$ e $x, y \in \mathbb{R}_0$.
- ($\vec{x} \neq 0$, $\vec{y} \neq 0$ e $\vec{x} \curlywedge \vec{y} = 0$) Nesse caso, segue que $\vec{x} \parallel \vec{y}$.

■

10.4.2.3 Rotações em $\mathbb{R}^3$

⊢ **Proposição 10.17.** *A função*

$$\begin{aligned} R\colon \mathbb{S}^3 &\longrightarrow \mathrm{SO}(3) \\ u &\longmapsto R_u\colon \vec{\mathbb{R}}^3 \longrightarrow \vec{\mathbb{R}}^3 \\ &\qquad\qquad x \longmapsto uxu^{-1}. \end{aligned}$$

é um homomorfismo sobrejetivo de grupos diferenciável e $\mathrm{nuc}(R) = \mathbb{S}^0$.

□ *Demonstração.* A função R é diferenciável porque $R_u(x)$ é um polinômio em cada entrada. Primeiro, vamos mostrar que R de fato está bem definida, ou seja, que, para todo $u \in \mathbb{S}^3$, $R_u \in \mathrm{SO}(3)$. Seja $u \in \mathbb{S}^3$. Pela fórmula de conjugação por quatérnio, temos que, para todo $x \in \vec{\mathbb{R}}^3$,

$$(R_u(x))_0 = (uxu^{-1})_0 = x_0 = 0,$$

portanto $R_u(x) \in \vec{\mathbb{R}}^3$ e então $R_u\colon \vec{\mathbb{R}}^3 \longrightarrow \vec{\mathbb{R}}^3$. A função R_u é linear, pois, para todos $x, x' \in \vec{\mathbb{R}}^3$ e $c \in \mathbb{R}$, segue da bilinearidade do produto quaterniônico que

$$R_u(cx+x') = u(cx+x')u^{-1} = (cux+ux')u^{-1} = cuxu^{-1}+ux'u^{-1} = cR_u(x)+R_u(x').$$

Isso mostra que $R_u \in \mathrm{GL}(3, \mathbb{R})$. Além disso, R_u preserva norma, pois

$$\|R_u(x)\| = \left\|uxu^{-1}\right\| = \|u\|\,\|x\|\,\|u\|^{-1} = \|x\|\,.$$

Isso mostra que $R_u \in \mathrm{O}(3)$. Agora, como R é diferenciável, em particular é contínua, logo, como $\mathbb{S}^3$ é conexo, $R(\mathbb{S}^3)$ é conexo. Isso implica que $R(\mathbb{S}^3) \subseteq \mathrm{SO}(3)$.

Seja $u \in \mathbb{S}^3$ tal que $R_u = \mathrm{I}$. Então, para todo $x \in \vec{\mathbb{R}}^3$, $uxu^{-1} = R_u(x) = x$, logo $ux = xu$. Isso ocorre se, e somente se, $\vec{u} \curlywedge x = 0$. Como isso vale para todo $x \in \vec{\mathbb{R}}^3$, segue que $\vec{u} = 0$, portanto $u = u_0 \in \mathbb{R}_0$; como $u \in \mathbb{S}^3$, concluímos que $u \in \mathbb{S}^0 = \{1, -1\}$, ou $\mathrm{nuc}(R) = \mathbb{S}^0$.

Para a sobrejetividade, primeiro notamos que, como $R(\mathbb{S}^3) \subseteq \mathrm{SO}(3)$ é compacto, pois $\mathbb{S}^3$ é compacto, e $\mathrm{SO}(3)$ é separado[5], segue que $R(\mathbb{S}^3)$ é fechado. Além disso,

[5] Hausdorff

R é difeomorfismo local, pois o núcleo $\mathbb{S}^0$ é um subgrupo discreto (portanto tem álgebra de Lie trivial), logo R é difeomorfismo local na identidade, e portanto em todo $\mathbb{S}^3$, já que um morfismo de grupos diferenciáveis[6] tem posto constante. Assim, segue que $R(\mathbb{S}^3) \subseteq \mathrm{SO}(3)$ é aberto. Como $\mathrm{SO}(3)$ é conexo, segue que $R(\mathbb{S}^3) = \mathrm{SO}(3)$. ■

Em particular, obtemos desse recobrimento duplo que

$$\mathbb{S}^3 \big/ \mathbb{S}^0 \simeq \mathrm{SO}(3).$$

Para calcular exatamente qual é a rotação que R_u representa em $\mathrm{SO}(3)$, lembremos que, pela decomposição polar, todo $u \in \mathbb{S}^3$ pode ser escrito como

$$u = \cos(\phi) + \operatorname{sen}(\phi)\hat{u} = \mathrm{e}^{\phi\hat{u}},$$

com $\phi = \cos(u_0) \in [0, \tau \diagup 2]$ e $\hat{u} = \frac{\vec{u}}{\|\hat{u}\|} \in \vec{\mathbb{S}}^2$.

A rotação de $x \in \mathbb{R}^3$ por um ângulo $\theta \in [0, \tau]$ em torno de um vetor unitário $\hat{u} \in \mathbb{S}^2$ é dada por

$$R^\theta_{\hat{u}}(x) = \mathrm{p}_{\|\hat{u}}(x) + \cos(\theta)\mathrm{p}_{\perp\hat{u}}(x) + \operatorname{sen}(\theta)\hat{u} \curlywedge x.$$

Isso ocorre pois, se rotacionamos um vetor x em torno de um eixo $\hat{u}$, a componente de x que é paralela a $\hat{u}$, dada pela projeção $\mathrm{p}_{\|\hat{u}}(x)$ de x em $\hat{u}$, fica constante, não rotaciona pois é paralela ao eixo de rotação, enquanto que a componente de x que é perpendicular a $\hat{u}$, dada pela projeção $\mathbb{P}_{\perp\hat{u}}(x)$, rotaciona no plano perpendicular a $\hat{u}$, dado pela base $\mathbb{P}_{\perp\hat{u}}(x)$ e $\hat{u} \curlywedge x$, e essa rotação é por um ângulo θ, portanto resultará em $\cos(\theta)\mathrm{p}_{\perp\hat{u}}(x) + \operatorname{sen}(\theta)\hat{u} \curlywedge x$. Um detalhe é que o plano perpendicular é dado por essa base se, e somente se, $\hat{u}$ e x não são paralelos, mas quando eles são paralelos vale $\mathrm{p}_{\|\hat{u}}(x) = x$ e $\mathrm{p}_{\perp\hat{u}}(x) = \hat{u} \curlywedge x = 0$, portanto a fórmula acima ainda vale.

Isso significa que

$$R^\theta_{\hat{u}} = \mathrm{p}_{\|\hat{u}} + \cos(\theta)\mathrm{p}_{\perp\hat{u}} + \operatorname{sen}(\theta)\hat{u} \curlywedge .$$

Como

$$\begin{aligned} 2\operatorname{sen}(\phi)^2 &= 1 - \cos(2\phi) \\ \cos(\phi)^2 - \operatorname{sen}(\phi)^2 &= \cos(2\phi) \\ 2\cos(\phi)\operatorname{sen}(\phi) &= \operatorname{sen}(2\phi) \end{aligned}$$

[6] Grupos de lie

segue que da fórmula de conjugação por quatérnio que

$$\begin{aligned}
\mathrm{e}^{\phi\hat{u}}x\mathrm{e}^{-\phi\hat{u}} &= 2(\vec{u}\cdot x)\vec{u} + \langle u,u\rangle_{1,3}\, x + 2u_0(\vec{u}\curlywedge x)\\
&= 2\operatorname{sen}(\phi)(\hat{u}\cdot x)(\operatorname{sen}(\phi)\hat{u})\\
&\qquad + (\cos(\phi)^2 - \operatorname{sen}(\phi)^2)x + 2\cos(\phi)\operatorname{sen}(\phi)(\hat{u}\curlywedge x)\\
&= (1-\cos(2\phi))(x\cdot\hat{u})\hat{u} + \cos(2\phi)x + \operatorname{sen}(2\phi)(\hat{u}\curlywedge x)\\
&= (x\cdot\hat{u})\hat{u} + \cos(2\phi)(x-(x\cdot\hat{u})\hat{u}) + \operatorname{sen}(2\phi)(\hat{u}\curlywedge x)\\
&= \mathrm{p}_{\parallel\hat{u}}(x) + \cos(2\phi)\mathrm{p}_{\perp\hat{u}}(x) + \operatorname{sen}(2\phi)(x)\\
&= R^{2\phi}_{\hat{u}}(x).
\end{aligned}$$

Isso mostra que

$$\begin{aligned}
R\colon \mathbb{S}^3 &\longrightarrow \mathrm{SO}(3)\\
\mathrm{e}^{\phi\hat{u}} &\longmapsto R^{2\phi}_{\hat{u}}\colon \vec{\mathbb{R}}^3 \longrightarrow \vec{\mathbb{R}}^3\\
&\qquad\qquad x \longmapsto \mathrm{e}^{\phi\hat{u}}x\mathrm{e}^{-\phi\hat{u}}.
\end{aligned}$$

Como $\operatorname{nuc}(R) = \mathbb{S}^0$, sabemos que $R^{-1}(R^{\theta}_{\hat{u}}) = \{\mathrm{e}^{\frac{\theta}{2}\hat{u}}, -\mathrm{e}^{\frac{\theta}{2}\hat{u}}\}$. Podemos achar $-\mathrm{e}^{\frac{\theta}{2}\hat{u}}$ explicitamente em função de seu ângulo em $[0,\tau\diagup 2]$ e eixo de rotação em $\vec{\mathbb{S}}^2$. Como para todo $\phi\in[0,\tau\diagup 2]$ vale

$$\begin{aligned}
-\cos(\phi) &= \cos((\tau\diagup 2)-\phi)\\
\operatorname{sen}(\phi) &= \operatorname{sen}((\tau\diagup 2)-\phi),
\end{aligned}$$

segue que

$$\begin{aligned}
-\mathrm{e}^{\frac{\theta}{2}\hat{u}} &= -(\cos(\theta\diagup 2) + \operatorname{sen}(\theta\diagup 2)\hat{u})\\
&= -\cos(\theta\diagup 2) - \operatorname{sen}(\theta\diagup 2)\hat{u}\\
&= \cos((\tau-\theta)\diagup 2) + \operatorname{sen}((\tau-\theta)\diagup 2)(-\hat{u})\\
&= \mathrm{e}^{\frac{\tau-\theta}{2}(-\hat{u})}.
\end{aligned}$$

Notemos que como $(\tau\diagup 2)-\phi\in[0,\tau\diagup 2]$ e $-\hat{u}\in\vec{\mathbb{S}}^2$, $\mathrm{e}^{\frac{\tau-\theta}{2}(-\hat{u})}$ é a decomposição polar de $-\mathrm{e}^{\frac{\theta}{2}\hat{u}}$. Obtemos assim

$$R^{-1}(R^{\theta}_{\hat{u}}) = \{\mathrm{e}^{\frac{\theta}{2}\hat{u}}, -\mathrm{e}^{\frac{\tau-\theta}{2}(-\hat{u})}\}.$$

As funções $\mathrm{p}_{\parallel u}$, $\mathrm{p}_{\perp u}$ e $u\curlywedge$ são funções lineares e, na base canônica de $\mathbb{R}^3$, são dadas pelas matrizes

$$[\mathrm{p}_{\parallel u}] = \begin{bmatrix} {u_1}^2 & u_1u_2 & u_1u_3\\ u_1u_2 & {u_2}^2 & u_2u_3\\ u_1u_3 & u_2u_3 & {u_3}^2\end{bmatrix},$$

$$[\mathrm{p}_{\perp u}] = \begin{bmatrix} 1-{u_1}^2 & -u_1u_2 & -u_1u_3 \\ -u_1u_2 & 1-{u_2}^2 & -u_2u_3 \\ -u_1u_3 & -u_2u_3 & 1-{u_3}^2 \end{bmatrix}$$

e

$$[u\curlywedge] = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}.$$

Matricialmente, $[R_u]$ é dada por

$$\begin{aligned} [R_u] &= [\mathrm{p}_{\|u}] + \cos(\theta)[\mathrm{p}_{\perp u}] + \operatorname{sen}(\theta)[u\curlywedge] \\ &= (1-\cos(\theta)) \begin{bmatrix} {u_1}^2 & u_1u_2 & u_1u_3 \\ u_1u_2 & {u_2}^2 & u_2u_3 \\ u_1u_3 & u_2u_3 & {u_3}^2 \end{bmatrix} + \cos(\theta)\mathrm{I} + \operatorname{sen}(\theta) \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \end{aligned}$$

10.4.2.4 Rotações em $\mathbb{R}^4$

:⊢ **Definição 10.17.** A função

$$\begin{aligned} R\colon \mathbb{S}^3 \times \mathbb{S}^3 &\longrightarrow \mathrm{SO}(4) \\ (u,v) &\longmapsto Ru,v\colon \mathbb{R}^4 \longrightarrow \mathbb{R}^4 \\ &\qquad\qquad\quad x \longmapsto uxv^{-1} \end{aligned}$$

é um homomorfismo sobrejetivo de grupos tal que $\operatorname{nuc}(R) = (1,1)\mathbb{S}^0$.

□ *Demonstração.* A função R é diferenciável porque $R_u(x)$ é um polinômio em cada entrada. Primeiro, vamos mostrar que R de fato está bem definida, ou seja, que, para todos $u, v \in \mathbb{S}^3$, $R_{u,v} \in \mathrm{SO}(4)$. Sejam $u, v \in \mathbb{S}^3$. A função $R_{u,v}$ é linear, pois, para todos $x, x' \in \vec{\mathbb{R}}^3$ e $c \in \mathbb{R}$, segue da bilinearidade do produto quaterniônico que

$$\begin{aligned} R_{u,v}(cx+x') &= u(cx+x')v^{-1} \\ &= (cux+ux')v^{-1} \\ &= cuxv^{-1} + ux'v^{-1} \\ &= cR_{u,v}(x) + R_{u,v}(x'). \end{aligned}$$

Isso mostra que $R_{u,v} \in \mathrm{GL}(4,\mathbb{R})$. Além disso, $R_{u,v}$ preserva norma, pois

$$\|R_{u,v}(x)\| = \left\|uxv^{-1}\right\| = \|u\|\,\|x\|\,\|v\|^{-1} = \|x\|\,.$$

Isso mostra que $R_{u,v} \in \mathrm{O}(4)$. Agora, como R é diferenciável, em particular é contínua, logo, como $\mathbb{S}^3 \times \mathbb{S}^3$ é conexo, $R(\mathbb{S}^3 \times \mathbb{S}^3)$ é conexo. Isso implica que $R(\mathbb{S}^3 \times \mathbb{S}^3) \subseteq \mathrm{SO}(4)$.

Sejam $u, v \in \mathbb{S}^3$ tais que $R_{u,v} = \mathrm{I}$. Então, para todo $x \in \mathbb{R}^4$, $uxv^{-1} = R_{u,v}(x) = x$. Em particular, $u1v^{-1} = 1$, portanto $u = v$ e $R_{u,v}(x) = uxu^{-1}$. Suponhamos por absurdo que $\vec{u} \neq 0$. Então existiria $\vec{x} \in \vec{\mathbb{R}}^3$ tal que $\vec{u} \cdot \vec{x} = 0$, de modo que $\{\vec{u}, \vec{x}, \vec{u} \curlywedge \vec{x}\}$ seria uma base de $\vec{\mathbb{R}}^3$. Pela fórmula de conjugação por quatérnio, seguiria que

$$\vec{u} = u\vec{u}u^{-1} = 2(\vec{u} \cdot \vec{u})\vec{u} + \langle u, u\rangle_{1,3}\,\vec{u} + 2u_0(\vec{u} \curlywedge \vec{u}) = (2 + \langle u, u\rangle_{1,3})\vec{u},$$

portanto $\langle u, u\rangle_{1,3} = -1$, e que

$$\vec{x} = u\vec{x}u^{-1} = 2(\vec{u} \cdot \vec{x})\vec{u} + \langle u, u\rangle_{1,3}\,\vec{x} + 2u_0(\vec{u} \curlywedge \vec{x}) = \langle u, u\rangle_{1,3}\,\vec{x} + 2u_0(\vec{u} \curlywedge \vec{x}).$$

portanto $\langle u, u\rangle_{1,3} = 1$, o que é uma contradição. Isso mostra que $\vec{u} = 0$, portanto que $u = u_0 \in \mathbb{R}_0$; como $u \in \mathbb{S}^3$, concluímos que $u \in \mathbb{S}^0$, ou seja, $\operatorname{nuc} R = \{(1,1), (-1,-1)\} = (1,1)\mathbb{S}^0$.

Para a sobrejetividade, primeiro notamos que, como $R(\mathbb{S}^3 \times \mathbb{S}^3) \subseteq \mathrm{SO}(4)$ é compacto, pois $\mathbb{S}^3 \times \mathbb{S}^3$ é compacto, e $\mathrm{SO}(4)$ é separado[7], segue que $R(\mathbb{S}^3 \times \mathbb{S}^3)$ é fechado. Além disso, R é difeomorfismo local, pois o núcleo $(1,1)\mathbb{S}^0$ é um subgrupo discreto (portanto tem álgebra de Lie trivial), logo R é difeomorfismo local na identidade, e portanto em todo $\mathbb{S}^3 \times \mathbb{S}^3$, já que um morfismo de grupos diferenciáveis[8] tem posto constante. Assim, segue que $R(\mathbb{S}^3 \times \mathbb{S}^3) \subseteq \mathrm{SO}(4)$ é aberto. Como $\mathrm{SO}(4)$ é conexo, segue que $R(\mathbb{S}^3 \times \mathbb{S}^3) = \mathrm{SO}(4)$. ■

10.4.3 Classificação das álgebras de divisão reais

Uma álgebra de divisão associativa real é uma álgebra associativa, unitária e invertível sobre $\mathbb{R}$. Nessa seção usaremos o teorema fundamental da álgebra, o teorema de Cayley-Hamilton e propriedades de produto interno, ou seja, 2-formas simétricas positivo-definidas, como a existência de uma base ortonormal.

⊢ **Proposição 10.18.** *Seja $\boldsymbol{A}$ uma álgebra de divisão associativa sobre $\mathbb{R}$. Então $\boldsymbol{A}$ é isomorfa a $\mathbb{R}$, $\mathbb{R}^2$ ou $\mathbb{R}^4$.*

□ *Demonstração.* Primeiro identificamos $\mathbb{R}$ com $\mathbb{R}1 = \{c1 \mid c \in \mathbb{R}\}$, em que $1 \in A$ é a unidade de A, e, conforme essa identificação, para todos $a, a' \in \mathbb{R}$, consideramos a ordem induzida $a \leq a'$. Mostraremos que $V := \{a \in A \mid a^2 \in \mathbb{R} \wedge a^2 \leq 0\}$ é um subespaço vetorial de $\boldsymbol{A}$ e $\boldsymbol{A} = \mathbb{R} \oplus \boldsymbol{V}$.

Defina $d := \dim(A)$. Para todo $a \in A$, a função multiplicação à esquerda

$$\begin{aligned} a\times\colon A &\longrightarrow A \\ x &\longmapsto ax \end{aligned}$$

[7] Hausdorff
[8] Grupos de lie

é uma função linear e, da associatividade de $\times$, segue que, para todos $a, a' \in A$, $(a\times) \circ (a'\times) = (aa')\times$. Seja p_a o polinômio característico de $a\times$. pelo teorema fundamental da álgebra, existem números reais $(c_k)_{k\in[n]}$ e números complexos $(z_k)_{k\in[n']}$ tais que

$$p_a(x) = (x - c_0)\cdots(x - c_{n-1})(x - z_0)(x - (z_0)^*)\cdots(x - z_{n-1})(x - (z_{n-1})^*),$$

e, para todo $k \in [n']$, o polinômio $(x - z_k)(x - (z_k)^*)$ é irredutível em $\mathbb{R}$. Pelo teorema de Cayley-Hamilton, segue que $p_a(a\times) = 0$. Da linearidade de $a\times$ e de $(a\times) \circ (a'\times) = (aa')\times$ segue que

$$p_a(a\times) = p_a(a) \times .$$

Como $\boldsymbol{A}$ é álgebra de divisão, segue que $p_a(a) = 0$ e que, para algum $k \in [n]$, $a - c_k = 0$ ou, para algum $k \in [n']$, $(a - z_k)(a - (z_k)^*) = 0$. No primeiro caso, segue que $a = c_k \in \mathbb{R}$ e, no segunda caso,

$$\mu_a(x) := (x - z_k)(x - (z_k)^*) = x^2 - 2\Re(z_k)x + |z_k|^2$$

é o polinômio minimal de $a\times$. Como p é real e tem as mesmas raízes complexas que μ_a, segue que, para algum $p \in \mathbb{Z}_>$,

$$p_a(x) = (\mu_a(x))^p = (x^2 - 2\Re(z_k)x + |z_k|^2)^p.$$

Como p_a é polinômio característico de $a\times$, o coeficiente de x^{2p-1} em p_a é $-\operatorname{tr} a\times$ (inverso aditivo do traço de $a\times$). Mas esse coeficiente é 0 se, e somente se, $\Re(z_k) = 0$, o que por sua vez é equivalente a $a^2 = -|z_k|^2$. Com $-|z_k|^2 < 0$, concluímos que

$$V = \{a \in A \mid \operatorname{tr} a\times = 0\},$$

portanto é um subespaço linear de $\boldsymbol{A}$ e tem codimensão 1, pois é o núcleo do traço $\operatorname{tr}\cdot$, um funcional linear não nulo, e conclui-se que $\boldsymbol{A} = \mathbb{R} \oplus \boldsymbol{V}$.

Isso implica que, para todo $v \in V$, $v^2 \in \mathbb{R}$e segue que, para todos $v, v' \in A$, $vv' + v'v = (v + v')^2 - v^2 - (v')^2 \in \mathbb{R}$. Assim, definimos

$$\begin{aligned} \langle\cdot,\cdot\rangle \colon V \times V &\longrightarrow \mathbb{R} \\ (v, v') &\longmapsto -\frac{vv' + v'v}{2}. \end{aligned}$$

Essa função é uma 2-forma simétrica positiva definida em V:

- (Bilinearidade) Segue da bilinearidade da multiplicação;
- (Simetria) Para todo $v \in V$,

$$\langle v, v'\rangle = -\frac{vv' + v'v}{2} = -\frac{v'v + vv'}{2} = \langle v', v\rangle;$$

- (Positiva-definida) Para todo $v \in V$, $v^2 \leq 0$, logo
$$\langle v, v\rangle = -\frac{vv + vv}{2} = -v^2 \geq 0.$$

Seja V' o subespaço de V que gera A como álgebra e seja $(b_k)_{k\in[n]}$ uma base ortonormal de V' com respeito a $\langle \cdot, \cdot \rangle$. A ortonormalidade implica que, para todo $k \in [n]$,
$${b_k}^2 = -\langle b_k, b_k\rangle = -1$$
e, para todos $k, k' \in [n]$ distintos,
$$\begin{aligned} b_k b_{k'} &= b_k b_{k'} + 0 \\ &= b_k b_{k'} + 2\langle b_k, b_{k'}\rangle \\ &= b_k b_{k'} - (b_k b_{k'} + b_{k'} b_k) \\ &= -b_{k'} b_k. \end{aligned}$$
Disso segue que, para todos $k, k' \in [n]$ distintos,
$$(b_k b_{k'})^2 = b_k b_{k'} b_k b_{k'} = -b_{k'} b_k b_k b_{k'} = -b_{k'}(-1) b_{k'} = b_{k'} b_{k'} = -1$$
e, para todos $k, k', k'' \in [n]$ distintos dois a dois,
$$\begin{aligned} (b_k b_{k'} b_{k''})^2 &= b_k b_{k'} b_{k''} b_k b_{k'} b_{k''} \\ &= -b_k b_{k'} b_k b_{k''} b_{k'} b_{k''} \\ &= (b_k b_k)(b_{k'} b_{k''})(b_{k'} b_{k''}) \\ &= -(b_{k'} b_{k''})^2 \\ &= 1. \end{aligned}$$
Isso implica que $n \in \{0, 1, 2\}$. Suponhamos, por absurdo, que $n \geq 3$, e seja $u := b_0 b_1 b_2$. Pelas relações anteriormente calculadas, seguiria que $u^2 = 1$, portanto $0 = u^2 - 1 = (u+1)(u-1)$. Como A é invertível, valeria $u + 1 = 0$ ou $u - 1 = 0$, o que implicaria que $b_2 = b_0 b_1$ ou $b_2 = -b_0 b_1$, o que seria uma contradição com a minimalidade de V'. Consideramos por fim cada um dos 3 casos.

- Se $n = 0$, segue que $V = \{0\}$ e portanto que $A \simeq \mathbb{R}$.
- Se $n = 1$, $V = \{c_0 b_0 \mid c_0 \in \mathbb{R}\}$ e segue que $A \simeq \mathbb{R}^2$ pelo isomorfismo
$$\begin{aligned} h\colon A &\longrightarrow \mathbb{R}^2 \\ c + c_0 b_0 &\longmapsto (c, c_0). \end{aligned}$$
- Se $n = 2$, $V = \{c_0 b_0 + c_1 b_1 + c_2(b_0 b_1) \mid c_0, c_1, c_2 \in \mathbb{R}\}$ e $A \simeq \mathbb{R}^4$ pelo isomorfismo
$$\begin{aligned} h\colon A &\longrightarrow \mathbb{R}^4 \\ c + c_0 b_0 + c_1 b_1 + c_2(b_0 b_1) &\longmapsto (c, c_0, c_1, c_2). \end{aligned}$$

■

10.4.4 Octônios $\mathbb{R}^8$

10.4.4.1 Multiplicação octoniônica

Consideremos o espaço linear $\mathbb{R}^8$ com a norma 2

$$\begin{aligned}\|\cdot\|\colon \mathbb{R}^8 &\longrightarrow \mathbb{R}\\ x &\longmapsto ({x_0}^2+{x_1}^2+{x_2}^2+{x_3}^2+{x_4}^2+{x_5}^2+{x_6}^2+{x_7}^2)^{\frac{1}{2}}.\end{aligned}$$

A base canônica de $\mathbb{R}^8$ será denotada por

$$\begin{aligned}
\mathrm{i}_0 &:= (1,0,0,0,0,0,0,0)\\
\mathrm{i}_1 &:= (0,1,0,0,0,0,0,0)\\
\mathrm{i}_2 &:= (0,0,1,0,0,0,0,0)\\
\mathrm{i}_3 &:= (0,0,0,1,0,0,0,0)\\
\mathrm{i}_4 &:= (0,0,0,0,1,0,0,0)\\
\mathrm{i}_5 &:= (0,0,0,0,0,1,0,0)\\
\mathrm{i}_6 &:= (0,0,0,0,0,0,1,0)\\
\mathrm{i}_7 &:= (0,0,0,0,0,0,0,1).
\end{aligned}$$

:⊢ **Definição 10.18.** A *multiplicação octoniônica* em $\mathbb{R}^8$ é a operação

$$\begin{aligned}\times\colon \mathbb{R}^8\times\mathbb{R}^8 &\longrightarrow \mathbb{R}^8\\ (x,y) &\longmapsto xy := (x_0y_0-x_1y_1-x_2y_2-x_3y_3-x_4y_4-x_5y_5-x_6y_6-x_7y_7,\\ &\qquad x_0y_1+x_1y_0+x_2y_3-x_3y_2+x_4y_5+x_5y_4-x_6y_7+x_7y_6,\\ &\qquad x_0y_2-x_1y_3+x_2y_0+x_3y_1+x_4y_6-x_5y_7-x_6y_4-x_7y_5,\\ &\qquad x_0y_3+x_1y_2-x_2y_1+x_3y_0+x_4y_7-x_5y_6+x_6y_5-x_7y_4,\\ &\qquad x_0y_4-x_1y_5-x_2y_6-x_3y_7+x_4y_0+x_5y_1+x_6y_2+x_7y_3,\\ &\qquad x_0y_5+x_1y_4-x_2y_7+x_3y_6-x_4y_1+x_5y_0-x_6y_3+x_7y_2,\\ &\qquad x_0y_6+x_1y_7+x_2y_4-x_3y_5-x_4y_2+x_5y_3+x_6y_0-x_7y_1,\\ &\qquad x_0y_7-x_1y_6+x_2y_5+x_3y_4-x_4y_3+x_5y_2+x_6y_1+x_7y_0).\end{aligned}$$

A *unidade* de $(\mathbb{R}^8,\times)$ é

$$1_{\mathbb{R}^8} := \mathrm{i}_0 = (1,0,0,0,0,0,0,0).$$

A *inversão* de $(\mathbb{R}^8,\times,1)$ é a operação

$$\begin{aligned}\diagup\colon \mathbb{R}^8\smallsetminus\{0\} &\longrightarrow \mathbb{R}^8\smallsetminus\{0\}\\ x &\longmapsto x^{-1} := \|x\|^{-2}\,(x_0,-x_1,-x_2,-x_3,-x_4,-x_5,-x_6,-x_7).\end{aligned}$$

Em geral, por não haver ambiguidade, denotaremos $1_{\mathbb{R}^8}$ simplesmente por 1. Assim, todo $x \in \mathbb{R}^8$ pode ser escrito

$$x = x_0 1 + x_1 \mathrm{i}_1 + x_2 \mathrm{i}_2 + x_3 \mathrm{i}_3 + x_4 \mathrm{i}_4 + x_5 \mathrm{i}_5 + x_6 \mathrm{i}_6 + x_7 \mathrm{i}_7.$$

Como a notação acima sugere, omitiremos o símbolo de multiplicação $\times$ sempre que possível pois isso não gerará ambiguidade.

A multiplicação octoniônica, quando restrita à base canônica, é dada pela tabela 10.3 e estendida linearmente.

$\times$	1	i_1	i_2	i_3	i_4	i_5	i_6	i_7
1	1	i_1	i_2	i_3	i_4	i_5	i_6	i_7
i_1	i_1	-1	i_3	$-\mathrm{i}_2$	i_5	$-\mathrm{i}_4$	$-\mathrm{i}_7$	i_6
i_2	i_2	$-\mathrm{i}_3$	-1	i_1	i_6	i_7	$-\mathrm{i}_4$	$-\mathrm{i}_5$
i_3	i_3	i_2	$-\mathrm{i}_1$	-1	i_7	$-\mathrm{i}_6$	i_5	$-\mathrm{i}_4$
i_4	i_4	$-\mathrm{i}_5$	$-\mathrm{i}_6$	$-\mathrm{i}_7$	-1	i_1	i_2	i_3
i_5	i_5	i_4	i_7	i_6	i_1	-1	$-\mathrm{i}_3$	$-\mathrm{i}_2$
i_6	i_6	i_7	i_4	$-\mathrm{i}_5$	$-\mathrm{i}_2$	i_3	-1	$-\mathrm{i}_1$
i_7	i_7	$-\mathrm{i}_6$	i_5	i_4	$-\mathrm{i}_3$	$-\mathrm{i}_2$	i_1	-1

TABELA 10.3: Tabela de multiplicação dos octônios.

Pela definição da multiplicação, vale que

$$\begin{aligned} x &= x_0 + x_1\mathrm{i}_1 + x_2\mathrm{i}_2 + x_3\mathrm{i}_3 + x_4\mathrm{i}_4 + x_5\mathrm{i}_5 + x_6\mathrm{i}_6 + x_7\mathrm{i}_7 \\ &= (x_0 + x_1\mathrm{i}_1 + x_2\mathrm{i}_2 + x_3\mathrm{i}_3) + (x_4 + x_5\mathrm{i}_1 + x_6\mathrm{i}_2 + x_7\mathrm{i}_3)\mathrm{i}_4 \\ &= (x_0 + x_1\mathrm{i}_1) + (x_2 + x_3\mathrm{i}_1)\mathrm{i}_2 + (x_4 + x_5\mathrm{i}_1)\mathrm{i}_4 + (x_6 - x_7\mathrm{i}_1)\mathrm{i}_6, \end{aligned}$$

FIGURA 10.2: Grafo para memorizar multiplicação octoniônica.

▷ **Exercício 10.8.** *A lista*

$$((\mathbb{R}^8, +, -, 0, \cdot, \times, /, 1), \|\cdot\|)$$

é uma álgebra[9] *alternativa invertível normada sobre* $\mathbb{R}$.

1. *(Bilinearidade) A multiplicação quaterniônica* $\times\colon \mathbb{R}^8 \times \mathbb{R}^8 \longrightarrow \mathbb{R}^8$ *é uma função bilinear;*
2. *(Alternatividade) Para todos* $x, y \in \mathbb{R}^8$,

$$\begin{aligned}(xx)y &= x(xy)\\ (xy)y &= x(yy);\end{aligned}$$

3. *(Identidade) Para todo* $x \in \mathbb{R}^8$,

$$x1 = 1x = x;$$

4. *(Inversa) Para todo* $x \in \mathbb{R}^8$,

$$x^{-1}x = 1 = xx^{-1};$$

5. *(Submultiplicatividade) Para todos* $x, y \in \mathbb{R}^8$,

$$\|xx'\| \leq \|x\| \, \|x'\| \, .$$

6. *(Unitariedade)* $\|1\| = 1$;
7. *(Invertibilidade) Para todo* $x \in \mathbb{R}^8$,

$$\left\|x^{-1}\right\| = \|x\|^{-1} \, .$$

$\curlywedge$	i_1	i_2	i_3	i_4	i_5	i_6	i_7
i_1	0	i_3	$-\mathrm{i}_2$	i_5	$-\mathrm{i}_4$	$-\mathrm{i}_7$	i_6
i_2	$-\mathrm{i}_3$	0	i_1	i_6	i_7	$-\mathrm{i}_4$	$-\mathrm{i}_5$
i_3	i_2	$-\mathrm{i}_1$	0	i_7	$-\mathrm{i}_6$	i_5	$-\mathrm{i}_4$
i_4	$-\mathrm{i}_5$	$-\mathrm{i}_6$	$-\mathrm{i}_7$	0	i_1	i_2	i_3
i_5	i_4	i_7	i_6	i_1	0	$-\mathrm{i}_3$	$-\mathrm{i}_2$
i_6	i_7	i_4	$-\mathrm{i}_5$	$-\mathrm{i}_2$	i_3	0	$-\mathrm{i}_1$
i_7	$-\mathrm{i}_6$	i_5	i_4	$-\mathrm{i}_3$	$-\mathrm{i}_2$	i_1	0

Tabela 10.4: Tabela de produto vetorial em $\vec{\mathbb{R}}^7$.

[9] A álgebra de $\mathbb{R}^8$ com essa multiplicação costuma ser denotada por $\mathbb{O}$, mas manteremos a notação $\mathbb{R}^8$ por simplicidade de notação, pois nenhuma ambiguidade resultará disso.

10.4.5 2-Hiperbólicos $\mathbb{R}^{1,1}$

Definiremos o 'produto hiperbólico' no espaço linear $\mathbb{R}^2$, cuja base canônica denotaremos por

$$\begin{aligned} \mathbf{1} &:= (1,0) \\ \mathrm{j} &:= (0,1). \end{aligned}$$

Como de costume, denotaremos o vetor $\mathbf{1}$ simplesmente por 1, pois isso não gerará ambiguidade, e sempre que possível ele será sempre omitido da notação, de modo que todo $x \in \mathbb{R}^2$ será denotado

$$x = x_0 + x_1\mathrm{j}.$$

Além disso, denotaremos o espaço $\mathbb{R}^2$ por $\mathbb{R}^{1,1}$ para enfatizar que estamos usando um produto diferente do complexo e que a geometria desse espaço será diferente.

:⊢ **Definição 10.19.** O *produto hiperbólico* em $\mathbb{R}^{1,1}$ é a operação

$$\begin{aligned} \times \colon \mathbb{R}^{1,1} \times \mathbb{R}^{1,1} &\longrightarrow \mathbb{R}^{1,1} \\ (x,y) &\longmapsto xy := (x_0y_0 + x_1y_1, x_0y_1 + x_1y_0) \end{aligned}$$

A *unidade* de $(\mathbb{R}^{1,1}, \times)$ é $1 = (1,0)$.

Escrevendo esse produto explicitamente na base de $\mathbb{R}^{1,1}$, temos

$$xy = (x_0y_0 + x_1y_1) + (x_0y_1 + x_1y_0)\mathrm{j}.$$

Notemos que

$$\mathrm{j}^2 = (0,1)(0,1) = (0 \times 0 + 1 \times 1) + (0 \times 1 + 1 \times 0)\mathrm{j} = 1.$$

⊢ **Proposição 10.19.** *A lista*

$$(\mathbb{R}^{1,1}, +, -, 0, \times, 1, \cdot)$$

é uma álgebra associativa, comutativa e unitária sobre $\mathbb{R}$.

□ *Demonstração.* • (Comutatividade) Para todos $x, y \in \mathbb{R}^{1,1}$,

$$xy = (x_0y_0 + x_1y_1, x_0y_1 + x_1y_0) = (y_0x_0 + y_1x_1, y_0x_1 + y_1x_0) = yx.$$

- (Associatividade) Faremos as contas por coordenadas. Para todos $x, y, z \in \mathbb{R}^{1,1}$, segue da definição do produto que as coordenadas de xy são

$$\begin{aligned}(xy)_0 &= x_0y_0 + x_1y_1\\ (xy)_1 &= x_0y_1 + x_1y_0\end{aligned}$$

e que as coordenadas de yz são

$$\begin{aligned}(yz)_0 &= y_0z_0 + y_1z_1\\ (yz)_1 &= y_0z_1 + y_1z_0.\end{aligned}$$

Ainda da definição do produto, as coordenadas de $(xy)z$ são

$$\begin{aligned}((xy)z)_0 &= (x_0y_0 + x_1y_1)z_0 + (x_0y_1 + x_1y_0)z_1\\ ((xy)z)_1 &= (x_0y_0 + x_1y_1)z_1 + (x_0y_1 + x_1y_0)z_0)\end{aligned}$$

e as coordenadas de $x(yz)$ são

$$\begin{aligned}(x(yz))_0 &= x_0(y_0z_0 + y_1z_1) + x_1(y_0z_1 + y_1z_0)\\ (x(yz))_1 &= x_0(y_0z_1 + y_1z_0) + x_1(y_0z_0 + y_1z_1).\end{aligned}$$

Para a coordenada 0, segue que

$$\begin{aligned}((xy)z)_0 &= (x_0y_0 + x_1y_1)z_0 + (x_0y_1 + x_1y_0)z_1\\ &= x_0y_0z_0 + x_1y_1z_0 + x_0y_1z_1 + x_1y_0z_1\\ &= x_0(y_0z_0 + y_1z_1) + x_1(y_0z_1 + y_1z_0)\\ &= (x(yz))_0\end{aligned}$$

e, para a coordenada 1, segue que

$$\begin{aligned}((xy)z)_1 &= (x_0y_0 + x_1y_1)z_1 + (x_0y_1 + x_1y_0)z_0)\\ &= x_0y_0z_1 + x_1y_1z_1 + x_0y_1z_0 + x_1y_0z_0\\ &= x_0(y_0z_1 + y_1z_0) + x_1(y_0z_0 + y_1z_1)\\ &= (x(yz))_1.\end{aligned}$$

Portanto

$$(xy)z = ((xy)z)_0 1 + ((xy)z)_1 \mathrm{j} = (x(yz))_0 1 + (x(yz))_1 \mathrm{j} = x(yz).$$

- (Bilinearidade) Como o produto é comutativo, basta mostrar a linearidade na primeira entrada. Para todos $x, x', y \in \mathbb{R}^{1,1}$ e $c \in \mathbb{R}$,

$$\begin{aligned}(cx + x')y &= ((cx_0 + x'_0)y_0 + (cx_1 + x'_1)y_1, (cx_0 + x'_0)y_1 + (cx_1 + x'_1)y_0)\\ &= (cx_0y_0 + x'_0y_0 + cx_1y_1 + x'_1y_1, cx_0y_1 + x'_0y_1 + cx_1y_0 + x'_1y_0)\\ &= (cx_0y_0 + cx_1y_1, cx_0y_1 + cx_1y_0) + (x'_0y_0 + x'_1y_1, x'_0y_1 + x'_1y_0)\\ &= c(x_0y_0 + x_1y_1, x_0y_1 + x_1y_0) + (x'_0y_0 + x'_1y_1, x'_0y_1 + x'_1y_0)\\ &= c(xy) + x'y.\end{aligned}$$

- (Identidade) Como o produto é comutativo, basta mostrar a identidade à esquerda. Para todo $x \in \mathbb{R}^{1,1}$,

$$1x = (1,0)(x_0, x_1) = (1x_0 + 0x_1, 1x_1 + 0x_0) = (x_0, x_1). \qquad \blacksquare$$

:⊢ **Definição 10.20.** Seja $x \in \mathbb{R}^{1,1}$. O *conjugado* de x é $x^* := x_0 - x_1\mathrm{j}$. A

▷ **Exercício 10.9.** *Seja $x \in \mathbb{R}^{1,1}$. Então $xx^* = {x_0}^2 - {x_1}^2 = \langle x, x\rangle_{1,1}$.*

10.4.6 4-Hiperbólicos $\mathbb{R}^{1,3}$

Definiremos o 'produto hiperbólico' no espaço linear $\mathbb{R}^4$, cuja base canônica denotaremos por

$$\begin{aligned}
\mathbf{1} &:= (1,0,0,0)\\
\mathrm{j}_1 &:= (0,1,0,0)\\
\mathrm{j}_2 &:= (0,0,1,0)\\
\mathrm{j}_3 &:= (0,0,0,1).
\end{aligned}$$

Como de costume, denotaremos o vetor $\mathbf{1}$ simplesmente por 1, pois isso não gerará ambiguidade, e sempre que possível ele será sempre omitido da notação, de modo que todo $x \in \mathbb{R}^4$ será denotado

$$x = x_0 + x_1\mathrm{j}_1 + x_2\mathrm{j}_2 + x_3\mathrm{j}_3.$$

Além disso, denotaremos o espaço $\mathbb{R}^4$ por $\mathbb{R}^{1,3}$ para enfatizar que estamos usando um produto diferente do quaterniônico e que a geometria desse espaço será diferente.

:⊢ **Definição 10.21.** O *produto hiperbólico* em $\mathbb{R}^{1,3}$ é a operação

$$\begin{aligned}
\times\colon \mathbb{R}^{1,3} \times \mathbb{R}^{1,3} &\longrightarrow \mathbb{R}^{1,3}\\
(x,y) &\longmapsto xy := (x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3)1\\
&\qquad\quad + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathrm{j}_1\\
&\qquad\quad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathrm{j}_2\\
&\qquad\quad + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathrm{j}_3.
\end{aligned}$$

A *unidade* de $(\mathbb{R}^{1,3}, \times)$

$$1 := (1,0,0,0).$$

Notemos que

$$\begin{aligned}
(\mathrm{j}_1)^2 &= 1\\
(\mathrm{j}_2)^2 &= 1\\
(\mathrm{j}_3)^2 &= 1\\
\mathrm{j}_1\mathrm{j}_2 &= \mathrm{j}_3 = -\mathrm{j}_2\mathrm{j}_1\\
\mathrm{j}_2\mathrm{j}_3 &= \mathrm{j}_1 = -\mathrm{j}_3\mathrm{j}_2\\
\mathrm{j}_3\mathrm{j}_1 &= \mathrm{j}_2 = -\mathrm{j}_1\mathrm{j}_3.
\end{aligned}$$

▷ **Exercício 10.10.** *A lista*

$$(\mathbb{R}^{1,3}, +, -, 0, \times, 1, \cdot)$$

é uma álgebra unitária sobre $\mathbb{R}$.

Capítulo 11

Ordem em estruturas algébricas

11.1 Corpos ordenados

Nesta seção, pretendemos formalizar a ideia de elementos positivos e negativos em um corpo, e para isso usaremos o conceito de ordem. Uma ordem, também chamada de ordem total, ordem linear ou cadeia, é uma relação reflexiva, antissimétrica, transitiva e total em um conjunto X. As três primeiras propriedades definem uma ordem parcial, e a quarta, também chamada de conexidade, significa existe relação entre qualquer par de elementos. De fato, as três últimas implicam a reflexividade, mas definimos assim para podermos dizer que uma ordem total é uma ordem parcial que satisfaz a totalidade.

Um corpo é um conjunto em que existe adição, subtração, multiplicação e divisão. Para unir os dois conceitos de modo que as estruturas de corpo e de ordem se relacionem, devemos exigir que elas satisfaçam algumas propriedades. Para isso, devemos relembrar dos conceitos de função monótona, de translação e de expansão. Uma função monótona crescente entre conjuntos ordenados é uma função f tal que, se dois elementos do domínio x, x' têm uma certa relação de ordem $x \leq x'$ (um é menor ou igual ao outro), então suas imagens têm a mesma relação de ordem, $f(x) \leq f(x')$. Isso quer dizer que a função preserva a ordem dos conjuntos. Uma função monótona decrescente inverte a relação de ordem, de modo que se dois elementos têm uma relação de ordem $x \leq x'$, suas imagens têm a relação de ordem dual, ou invertida, $f(x') \leq f(x)$. Lembremos ainda que, em uma ordem total, definimos para cada $e \in X$ os seguintes intervalos

$$\begin{aligned}
[e, \infty[&= \{e \in X \mid e \leq x\} \\
]e, \infty[&= \{e \in X \mid e < x\} \\
]-\infty, e] &= \{e \in X \mid x \leq e\} \\
]-\infty, e[&= \{e \in X \mid x < e\}\,.
\end{aligned}$$

Claramente, supomos que ∞ não é um símbolo usado para representar um elemento de X, de modo a não gerar confusão.

Para entrelaçar as estruturas de corpo e de ordem, usamos as translações e expansões em corpos e exigimos que elas sejam monótonas. Seja $\boldsymbol{C}$ um corpo e $c \in C$. A translação em C por c é a função

$$\begin{aligned} T_c \colon C &\longrightarrow C \\ c' &\longmapsto c + c' \end{aligned}$$

e a expansão de C por c é a função

$$\begin{aligned} E_c \colon C &\longrightarrow C \\ c' &\longmapsto cc'. \end{aligned}$$

Aqui assumimos que os corpos são comutativos, por isso não é necessário considerar translações e expansões à esquerda e à direita.

Pedir que as funções sejam monótonas é bem natural para preservar a estrutura de ordem do corpo. No entanto, se escolhemos ter translação monótona, não podemos ter qualquer expansão monótona, já que se E_{-1} fosse monótona, seguiria que, para todos $c, c' \in C$,

$$c \leq c' \Rightarrow -c \leq -c'$$

e da monotonicidade da adição seguiria que

$$c' = (c' + c) + (-c) \leq (c' + c) + (-c') = c.$$

Como $c \leq c'$ e $c' \leq c$, seguiria então da antissimetria da ordem que $c = c'$. Como a ordem é total, isso implicaria ainda que todos elementos são iguais, já que quaisquer dois elementos podem ser comparados. Isso nos sugere, então, a imitar a relação de ordem nos números racionais e reais, e exigir que a expansão por elementos positivos sejam crescente e por elementos negativos seja decrescente. Claro, isso nos obriga a definir o que é ser positivo ou negativo. Positivo significa simplesmente ser maior ou igual a 0, e negativo significa ser menor ou igual a 0. A definição de 0 como positivo ou negativo é, na maioria dos casos, irrelevante, mas em geral é mais conveniente considerá-lo ambos, em vez de nenhum dos dois.

:⊢ **Definição 11.1.** Um *corpo ordenado* é um par $(\boldsymbol{C}, \leq)$ em que $\boldsymbol{C}$ é um corpo e $\leq$ é uma ordem em C que satisfaz:

1. (Monotonicidade da translação) Para todo $c \in C$, a translação $T_c \colon C \to C$ é uma função monótona crescente: para todos $c', c'' \in C$

$$c' \leq c'' \Rightarrow c + c' \leq c + c'';$$

2. (Monotonicidade da expansão) Para todo $c \in [0, \infty[$ ($c \in C$ tal que $c \geq 0$), a expansão $E_c \colon C \to C$ é uma função monótona crescente: para todos $c', c'' \in C$
$$c' \leq c'' \Rightarrow cc' \leq cc''.$$

Os elementos *positivos*, *negativos*, *estritamente positivos* e *estritamente negativos* de C são os elementos $c \in C$ tais que $c \geq 0$, $c \leq 0$, $c > 0$ e $c < 0$, respectivamente. O conjuntos desses elementos são denotados, respectivamente, $C_{\geq 0} := [0, \infty[$, $C_{\leq 0} :=]-\infty, 0]$, $C_{>0} :=]0, \infty[$ e $C_{<0} :=]-\infty, 0[$.

Poderia-se perguntar por que devemos ter uma ordem total, por que não só uma ordem parcial. Se queremos definir o conceito de positividade e negatividade, devemos garantir que todos os elementos sejam relacionados com o 0. Suponhamos então que temos uma ordem parcial tal que todos elementos são relacionados com 0 e mostremos que essa ordem é total. Para todos $c, c' \in C$, consideremos então $c' - c \in C$. Como todos elementos podem ser comparados com 0, segue que $c' - c \geq 0$ ou $c' - c \leq 0$. Da monotonicidade crescente da translação por c, segue que $c' \geq c$ ou $c' \leq c$, portanto a ordem é total. O conceito de monotonicidade da expansão ser crescente ou decrescente de acordo com a positividade do elemento pelo qual se multiplica exige, em si, que todos elementos sejam relacionados com 0. Nas seções seguintes veremos como o conceito de cone positivo se relaciona com o de ordem em corpos.

$\vdash$ **Proposição 11.1.** *Seja* $(\boldsymbol{C}, \leq)$ *um corpo ordenado.*

1. *Para todo* $c \in C$, $c \in C_{\geq 0}$ *se, e somente se,* $-c \in C_{\leq 0}$*;*
2. *Para todo* $c \in C_{\leq 0}$*, a expansão* $E_c \colon C \to C$ *é uma função monótona decrescente: para todos* $c', c'' \in C$
$$c' \leq c'' \Rightarrow cc'' \leq cc'.$$
3. *Para todos* $c, c' \in C$*,*
   3.1. *se* $c, c' \in C_{\geq 0}$*, então* $c + c' \in C_{\geq 0}$*;*
   3.2. *se* $c, c \in C_{\leq 0}$*, então* $c + c' \in C_{\leq 0}$*;*
4. *Para todos* $c, c' \in C$*,* $cc' \in C_{\geq 0}$ *se, e somente se,* $c, c' \in C_{\geq 0}$ *ou* $c, c' \in C_{\leq 0}$*;*
5. *A inversa da adição* $- \colon C \to C$ *é decrescente: para todos* $c, c' \in C$ *tais que* $c \leq c'$*,* $-c' \leq -c$*;*
6. *A inversa da multiplicação* ${}^{-1} \colon C \smallsetminus \{0\} \to C \smallsetminus \{0\}$ *é decrescente: para todos* $c, c' \in C \smallsetminus \{0\}$ *tais que* $c \leq c'$*,* $c'^{-1} \leq c^{-1}$*;*
7. *Para todo* $c \in C$*,* $c^2 \geq 0$*;*
8. $-1 \leq 0 \leq 1$*;*
9. *Para todos* $c, c', d, d' \in C$ *tais que* $c \leq c'$ *e* $d \leq d'$*,* $c + d \leq c' + d'$*;*

□ *Demonstração.* 1. Seja $c \in C_{\geq 0}$. Isso significa que $c \geq 0$. Pela monotonicidade crescente da translação por $-c$, segue que

$$-c = -c + 0 \leq -c + c = 0,$$

portanto $-c \in C_{\leq 0}$. A recíproca segue direto de $c = -(-c)$.

2. Sejam $c \in C_{\leq 0}$ e $c', c'' \in C$ tais que $c' \leq c''$. Então temos que $-c \in C_{\leq 0}$, portanto

$$-cc' \leq -cc''$$

e somando $cc' + cc''$, segue que

$$cc'' = (cc' + cc'') - cc' \leq (cc' + cc'') - cc'' = cc'.$$

3.
4. Se $c, c' \in C_{\geq 0}$, então $c \geq 0$ e $c' \geq 0$, logo da monotonicidade crescente de E_c segue que $cc' \geq c0 = 0$. Se $c, c' \in C_{\leq 0}$, então $c \leq 0$ e $c' \leq 0$, logo da monotonicidade decrescente de E_c segue que $cc' \geq c0 = 0$.
Reciprocamente, se $c \in C_{\geq 0}$ e $c' \in C_{\leq 0}$, então $c \geq 0$ e $c' \leq 0$, logo da monotonicidade crescente de E_c segue que $cc' \leq c0 = 0$. O caso em que $c \in C_{\leq 0}$ e $c' \in C_{\geq 0}$ é o mesmo, trocando c e c'.

■

⊢ **Proposição 11.2** (Relação de ordem e ciclicidade)**.** *Seja* $(\boldsymbol{C}, \leq)$ *um corpo ordenado. Se* $\operatorname{car}(\boldsymbol{C}) \neq 0$*, então* $\boldsymbol{C}$ *é o corpo trivial (*$C = \{0\}$*).*

□ *Demonstração.* Seja $\operatorname{car}(C) = n \in \mathbb{N} \setminus \{0\}$. Consideremos $0, 1, n_C \in C$. Pela monotonicidade da adição, segue que,

$$0 \leq 1 \leq 1 + 1 \leq \ldots \leq n_C = 0,$$

o que implica, pela antissimetria da ordem, que $1 = 0$. ■

11.1.1 Imersão dos inteiros e racionais

Lembremos que a função

$$\begin{aligned} h \colon \mathbb{Z} &\longrightarrow C \\ n &\longmapsto \begin{cases} +_{i \in [n]} 1_C, & n > 0 \\ 0_C, & n = 0 \\ +_{i \in [-n]} (-1_C), & n < 0 \end{cases} \end{aligned}$$

é um homomorfismo de anel. Como $(\boldsymbol{C}, \leq)$ é um corpo ordenado não trivial, ele não tem torção, logo h é injetivo e sua imagem é $\mathbb{Z}_C$. A função

$$\begin{aligned} \bar{h} \colon \mathbb{Q} &\longrightarrow C \\ \frac{n}{d} &\longmapsto h(n)h(d)^{-1} \end{aligned}$$

é um homomorfismo de corpo injetivo cuja imagem é $\mathbb{Q}_C$. O homomorfismo $\bar{h}$ é o único homomorfismo injetivo que estende h, de acordo com a propriedade universal dos corpos de frações 7.44. Assim, todo corpo sem torção, em particular um corpo ordenado não trivial, tem um subcorpo isomorfo aos racionais. Note que $\bar{h}$ é crescente, pois $\mathbb{Q}_C$ tem a mesma relação de ordem que $\mathbb{Q}$.

⊢ **Proposição 11.3.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado não trivial. Para todo $\frac{n}{d}, \frac{n'}{d'} \in \mathbb{Q}_C$,*

$$\frac{p}{q} \leq \frac{p'}{q'} \Leftrightarrow pq' \leq qp'$$

□ *Demonstração.* Tomemos $q, q' \geq 0$ sem perda de generalidade. Suponhamos que $\frac{p}{q} \leq \frac{p'}{q'}$. Da monotonicidade crescente de $E_{qq'}$, segue que

$$pq' = qq'\frac{p}{q} \leq qq'\frac{p'}{q'} = qp'.$$

A recíproca é evidente. ■

11.1.2 Cones positivos

Apresentamos brevemente uma definição alternativa de ordem em corpos.

:⊢ **Definição 11.2.** Seja $\boldsymbol{C}$ um corpo. Um *cone pré-positivo* em C é um conjunto $P \subseteq C$ tal que

1. (Fechamento da adição) Para todos $p, p' \in P$, $p + p' \in P$;
2. (Fechamento da multiplicação) Para todos $p, p' \in P$, $pp' \in P$;
3. (Positividade do quadrado) Para todo $c \in C$, $c^2 \in P$;
4. (Não-positividade do oposto da unidade) $-1 \notin P$.

Um *cone positivo* é um cone pré-positivo $P \subseteq C$ tal que $C = P \cup -P$.

⊢ **Proposição 11.4.** *Sejam $\boldsymbol{C}$ um corpo e $P \subseteq C$ um cone pré-positivo.*

1. *$0, 1 \in P$;*
2. *Para todo $p \in P \smallsetminus \{0\}$, $p^{-1} \in P$;*

3. $P \cap -P = \{0\}$*;*
4. $P \cup -P$ *é um subcorpo de* C*.*

□ *Demonstração.* 1. Pela positividade do quadrado, segue que $0 = 0^2 \in P$ e $1 = 1^2 \in P$.
2. Exercício.
3. Seja $p \in P \cap -P$. Então $p, -p \in P$. Se $p \neq 0$, então $p^{-1} \in P$ e segue que $-1 = -pp^{-1} \in P$, o que é uma contradição. Logo $p = 0$.
4. Exercício.

■

⊢ **Proposição 11.5.** *Seja* $(\boldsymbol{C}, \leq)$ *um corpo ordenado. O conjunto* $C_{\geq 0}$ *de elementos positivos de* C *é um cone positivo.*

□ *Demonstração.* Segue direto das proposições anteriores. ■

:⊢ **Definição 11.3.** Sejam $\boldsymbol{C}$ um corpo e $P \subseteq C$ um cone pré-positivo. A *ordem* induzida por P é a relação $\leq_P$ definida por: para todos $c, c' \in C$, $c \leq c'$ se, e somente se, $c' - c \geq 0$.

⊢ **Proposição 11.6.** *Sejam* $\boldsymbol{C}$ *um corpo e* $P \subseteq C$ *um cone pré-positivo.*

1. A ordem $\leq_P$ *induzida por* P *é uma ordem parcial em* C*;*
2. Para todo $c \in C$*, a translação* $T_c \colon C \to C$ *é monótona crescente;*
3. Para todo $c \in C$*,*

3.1. se $c \in P$*, a expansão* $E_c \colon C \to C$ *é monótona crescente;*
3.2. se $c \in -P$*, a expansão* $E_c \colon C \to C$ *é monótona decrescente.*

□ *Demonstração.* 1. (Reflexividade) Seja $c \in C$. Isso significa que $c - c = 0 \in P$, logo $c \leq_P c$.
(Antissimetria) Sejam $c, c' \in C$ tais que $c \leq_P c'$ e $c' \leq_P c$. Isso significa que $c' - c \in P$ e $c - c' = -(c' - c) \in P$, o que implica que $c' - c = 0$, logo $c = c'$.
(Transitividade) Sejam $c, c', c'' \in C$ tais que $c \leq_P c'$ e $c' \leq_P c''$. Isso significa que $c' - c \in P$ e $c'' - c' \in P$, logo $(c'' - c') + (c' - c) = c'' - c \in P$, portanto $c \leq_P c''$.
2. Sejam $c \in C$ e $c', c'' \in C$ tais que $c' \leq_P c''$. Isso significa que $c'' - c \in P$, portanto
$$(c + c'') - (c + c') = c'' - c' \in P,$$
o que significa que $c + c' \leq_P c + c''$. Isso mostra que T_c é monótona crescente.
3. Sejam $c \in C$ e $c', c'' \in C$ tais que $c' \leq_P c''$. Isso significa que $c'' - c' \in P$.

3.1. Se $c \in P$, segue do fechamento por multiplicação que
$$cc'' - cc' = c(c'' - c') \in P,$$
o que significa que $cc' \leq_P cc''$. Isso mostra que E_c é monótona crescente.

3.2. Se $c \in -P$, então $-c \in P$, e segue do fechamento por multiplicação que

$$cc' - cc'' = -c(c'' - c') \in P,$$

o que significa que $cc'' \leq_P cc'$. Isso mostra que E_c é monótona decrescente.

■

No entanto, para elementos $c \in C$ que não estão em $P \cup -P$, não podemos garantir a monotonicidade da expansão por ele. Note que $C \neq P \cup -P$ é equivalente a termos elementos que não são relacionados com 0.

⊢ **Proposição 11.7.** *Sejam $\boldsymbol{C}$ um corpo e $P \subseteq C$ um cone positivo. O par $(\boldsymbol{C}, \leq_P)$ é um corpo ordenado.*

□ *Demonstração.* Pela proposição anterior, $\leq_P$ é uma ordem parcial. Basta mostrar a totalidade. Sejam $c, c' \in C$. Como $C = P \cup -P$, então $c' - c \in P \cup -P$. Isso implica que $c' - c \in P$ ou $c' - c \in -P$, logo $c \leq_P c'$ ou $c' \leq_P c$. Assim concluímos que $\leq_P$ é uma ordem.

Pela proposição anterior, a translação é monótona. Como $C = P \cup -P$, então $c \geq_P 0$ se, e somente se, $c \in P$ e $c \leq_P 0$ se, e somente se, $c \in -P$, e segue da proposição anterior que, se $c \geq_P 0$, a expansão E_c é monótona crescente e, se $c \leq_P 0$, a expansão E_c é monótona decrescente. ■

⊢ **Proposição 11.8.** *Seja $\boldsymbol{C}$ um corpo. Existe uma bijeção entre as ordens em $\boldsymbol{C}$ tais que $(\boldsymbol{C}, \leq)$ é um corpo ordenado e os cones positivos de $\boldsymbol{C}$.*

□ *Demonstração.* Usando os resultados anteriores, essa demonstração fica simples. A bijeção é dada por $\leq \mapsto C_{\geq 0}$ e sua inversa é $P \mapsto \leq_P$. ■

11.1.3 Corpos ordenados completos

:⊢ **Definição 11.4.** Um corpo ordenado *completo* é um corpo ordenado $(\boldsymbol{C}, \leq)$ tal que todo conjunto não-vazio limitado superiormente admite supremo: para todo $A \subseteq C$ não-vazio imitado superiormente, existe $s \in C$ tal que

$$s = \sup A.$$

⊢ **Proposição 11.9.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado completo (não-trivial). Existe isomorfismo de corpos crescente $h \colon \mathbb{R} \longrightarrow C$.*

□ *Demonstração.* Definimos h como

$$\begin{aligned} h\colon \mathbb{R} &\longrightarrow C \\ x &\longmapsto \sup\left\{ \frac{i_C}{j_C} \;\middle|\; (i,j) \in \mathbb{Z}\times\mathbb{Z}\smallsetminus\{0\}, \frac{i}{j} \leq x \right\}. \end{aligned}$$

Definamos $L(x) := \{q \in \mathbb{Q} \mid q \leq x\}$. Mostremos que h é uma bijeção. (Injetividade) Sejam $x, x' \in \mathbb{R}$ tais que $x \neq x'$. Da totalidade da ordem, $x < x'$ ou $x' < x$, logo $\sup L(x) < \sup L(x')$ ou $\sup L(x') < \sup L(x)$. No primeiro caso, segue que $h(x) < h(x')$, e no segundo $h(x') < h(x)$, portanto $h(x) \neq h(x')$. (Sobrejetividade) Seja $c \in C$. Então $c = \sup\{q \in \mathbb{Q}_C \mid q \leq c\}$, portanto definindo $x := \sup\left\{\frac{i_C}{j_C} \;\middle|\; (i,j) \in \mathbb{Z}\times\mathbb{Z}\smallsetminus\{0\}, \frac{i_C}{j_C} \leq c\right\}$, temos que $h(x) = y$.

Agora, mostremos que h é isomorfismo de corpos. (Homomorfismo de grupo aditivo) Sejam $x, x' \in \mathbb{R}$. Então $x + x' = \sup L(x + x')$, logo

$$h(x + x') = h(\sup L(x + x')) = \sup h(L(x + x')) = h(x) + h(x').$$

(Homomorfismo de grupo multiplicativo) Sejam $x, x' \in \mathbb{R}$. Então $xx' = \sup L(xx')$, logo

$$h(xx') = h(\sup L(xx')) = \sup h(L(xx')) = h(x)h(x').$$

Os detalhes podem ser completados pelo leitor.

O isomorfismo h é crescente pois, se $x \leq x'$, então $L(x) \subseteq L(x')$, logo $h(L(x)) \subseteq h(L(x'))$, o que implica $h(x) \leq h(x')$. ■

Essa proposição mostra que corpos ordenados completos são, basicamente, cópias de $\mathbb{R}$. Isso nos permite usar $\mathbb{R}$ sempre que quisermos usar algum corpo que seja ordenado e completo, e tira um pouco a arbitrariedade algébrica de usar o corpo dos números reais.

11.1.4 Corpos ordenados infinitesimais

:⊢ **Definição 11.5.** Um *corpo não-infinitesimal* (ou *arquimediano*) é um corpo ordenado $(\boldsymbol{C}, \leq)$ tal que $\mathbb{N}_C$ é ilimitado superiormente: para todo $c \in C$, existe $n \in \mathbb{N}_C$ tal que $c < n$.

⊢ **Proposição 11.10.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. São equivalentes:*

1. *$(\boldsymbol{C}, \leq)$ é arquimediano;*
2. *Para todos $c \in C_{>0}$ e $c' \in C$, existe $n \in \mathbb{N}_C$ tal que $nc > c'$;*
3. *Para todo $c \in C_{>0}$, existe $n \in \mathbb{N}_C$ tal que $0 < \frac{1}{n} < c$;*
4. *Para todo $c \in C_{>0}$, existe $n \in \mathbb{N}_C$ tal que $0 < \frac{1}{2^n} < c$.*
5. *Para todos $c, c' \in C$ tais que $c < c'$, existe $q \in \mathbb{Q}_C$ tal que $c < q < c'$.*

11.1.5 Valor absoluto e distância

:⊢ **Definição 11.6.** Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. O *valor absoluto* de C é a função

$$\begin{aligned} |\cdot| \colon C &\longrightarrow C_{\geq 0} \\ c &\longmapsto \begin{cases} c, & c \in C_{\geq 0} \\ -c, & c \in C_{\leq 0}. \end{cases} \end{aligned}$$

⊢ **Proposição 11.11.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. O valor absoluto satisfaz*

1. *(Separação) Para todo $c \in C$, $|c| = 0$ se, e somente se, $c = 0$;*
2. *(Multiplicatividade) Para todos $c, c' \in C$, $|cc'| = |c|\,|c'|$;*
3. *(Subaditividade) Para todos $c, c' \in C$, $|c + c'| \leq |c| + |c'|$;*
4. *Para todo $c \in C$, $|-c| = |c|$;*
5. *Para todo $c \in C \smallsetminus \{0\}$, $|c^{-1}| = |c|^{-1}$.*

□ *Demonstração.*

1. Como $0 \in C_{\geq 0}$, $|0| = 0$. Reciprocamente, se $|c| = 0$, então $c = 0$ ou $-c = 0$. Ambos os casos implicam $c = 0$.
2. Sejam $c, c' \in C$. Consideramos quatro casos. Se $c, c' \in C_{\geq 0}$, então $cc' \in C_{\geq 0}$, logo
$$|cc'| = cc' = |c|\,|c'|\,.$$
Se $c, c' \in C_{\leq 0}$, então $cc' \in C_{\geq 0}$, logo
$$|cc'| = cc' = (-c)(-c') = |c|\,|c'|\,.$$
Se $c \in C_{\geq 0}$ e $c \in C_{\leq 0}$, então $cc' \in C_{\leq 0}$, logo
$$|cc'| = -cc' = |c|\,|c'|\,.$$
O caso em que $c \in C_{\leq 0}$ e $c \in C_{\geq 0}$ é o mesmo, trocando c e c'.
3. Sejam $c, c' \in C$. Consideramos quatro casos. Se $c, c' \in C_{\geq 0}$, então $c + c' \in C_{\geq 0}$, logo
$$|c + c'| = c + c' = |c| + |c'|\,.$$
Se $c, c' \in C_{\leq 0}$, então $c + c' \in C_{\leq 0}$, logo
$$|c + c'| = -(c + c') = (-c) + (-c') = |c| + |c'|\,.$$
Se $c \in C_{\geq 0}$ e $c' \in C_{\leq 0}$, então $-c \leq c$ e $c' \leq -c'$. Consideramos dois casos. Se $c + c' \in C_{\geq 0}$, então
$$|c + c'| = c + c' \leq c + (-c') = |c| + |c'|$$
Se $c + c' \in C_{\leq 0}$, então
$$|c + c'| = -(c + c') = (-c) + (-c') \leq c + (-c') = |c| + |c'|$$

4. Seja $c \in C$. Se $c \in C_{\geq 0}$, então $|c| = c$ e $-c \in C_{\leq 0}$, logo $|-c| = c = |c|$; se $c \in C_{\leq 0}$, então $|c| = -c$ e $-c \in C_{\geq 0}$, logo $|-c| = -c = |c|$.
5. Seja $c \in C \smallsetminus \{0\}$. Então

$$\left|c^{-1}\right| = \left|c^{-1}\right| |c| \, |c|^{-1} = \left|c^{-1}c\right| |c|^{-1} = |1| \, |c|^{-1} = |c|^{-1} .$$

■

:⊢ **Definição 11.7.** Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. A *distância* de C é a função

$$\begin{aligned} |\cdot, \cdot| \colon C \times C &\longrightarrow C_{\geq 0} \\ (c, c') &\longmapsto |c' - c| . \end{aligned}$$

⊢ **Proposição 11.12.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. A distância $|\cdot, \cdot| : C \times C \to C_{\geq 0}$ satisfaz:*

1. *(Separação) Para todos $c, c' \in C$,*

$$|c, c'| = 0 \quad \Longleftrightarrow \quad c = c';$$

2. *(Simetria) Para todos $c, c' \in C$,*

$$|c, c'| = |c', c| ;$$

3. *(Desigualdade Triangular) Para todos $c, c', c'' \in C$,*

$$|c, c''| \leq |c, c'| + |c', c''| ;$$

4. *(Invariância por Translação) Para todos $c, c', c'' \in C$,*

$$|c + c', c + c''| = |c', c''| ;$$

5. *(Homogeneidade Absoluta) Para todos $c, c', c'' \in C$,*

$$|cc', cc''| = |c| \, |c', c''| .$$

6. *Para todos $c, c' \in C$,*

$$|-c, -c'| = |c, c'| ;$$

7. *Para todo $c \in C$,*

$$|c| = |0, c| .$$

11.1.6 Conicidade

:⊢ **Definição 11.8** (Cone)**.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$. Um *cone* (ou *conjunto cônico*) em V é um conjunto $X \subseteq V$ tal que, para todo $c \in C_{\geq 0}$,

$$cX \subseteq X.$$

Isso é equivalente a, para todos $c \in C_{\geq 0}$ e $x \in X$, $cx \in X$.

:⊢ **Definição 11.9** (Combinação cônica)**.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$ e $v_0, \cdots, v_{n-1} \in V$. Uma *combinação cônica* de $v_0, \cdots, v_{n-1}$ é um vetor $v \in V$ para o qual existem $c_0, \cdots, c_{n-1} \in C_{\geq 0}$ tais que

$$v = \mathop{+}_{i \in [n]} c_i v_i.$$

Um *cone agudo* é um cone X tal que $X \cap -X = \{0\}$.

⊢ **Proposição 11.13.** *Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$ e $X \subseteq V$ um cone.*

1. *Se $X \neq \emptyset$, então $0 \in X$;*
2. *$X \cap -X$ é o maior subespaço linear contido em X.*

:⊢ **Definição 11.10.** Sejam X um conjunto e $(\boldsymbol{C}, \leq)$ um corpo ordenado. Uma *função positiva* de X para C é uma função de X para $C_{\geq 0}$; ou seja, uma função $f \colon X \to C$ tal que, para todo $x \in X$, $f(x) \geq 0$. O espaço de funções positivas de X para C é $C^X_{\geq 0}$.

⊢ **Proposição 11.14.** *Sejam X um conjunto e $(\boldsymbol{C}, \leq)$ um corpo ordenado. O espaço $C^X_{\geq 0}$ de funções positivas é um cone em C^X.*

□ *Demonstração.* Sejam $c \in C_{\geq 0}$ e $f \in C^X_{\geq 0}$. Então, para todo $x \in X$, $f(x) \geq 0$, portanto $cf(x) \geq 0$, logo $cf \geq 0$. ■

11.1.7 Convexidade

:⊢ **Definição 11.11** (Conjunto convexo)**.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$. Um *conjunto convexo* em V é um conjunto $X \subseteq V$ tal que, para todo $c \in [0, 1]$,

$$(1 - c)X + cX \subseteq X.$$

Isso é equivalente a dizer que, para todos $x, x' \in X$ e $c \in [0, 1]$,

$$(1 - c)x + cx' \in X.$$

Uma combinação convexa é uma combinação cônica cuja soma dos coeficientes é 1.

:⊢ **Definição 11.12** (Combinação convexa)**.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$ e $v_0, \cdots, v_{n-1} \in V$. Uma *combinação convexa* de $v_0, \cdots, v_{n-1}$ é um vetor $v \in V$ para o qual existem $c_0, \cdots, c_{n-1} \in C_{\geq 0}$ tais que $+_{i\in[n]} c_i = 1$ e

$$v = \mathop{+}_{i\in[n]} c_i v_i.$$

Um conjunto convexo é um conjunto fechado por combinações convexas.

⊢ **Proposição 11.15.** *Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$ e X um conjunto convexo em V. Para todos $c_0, \cdots, c_{n-1} \in C_{\geq 0}$ tais que $+_{i\in[n]} c_i = 1$,*

$$\mathop{+}_{i\in[n]} c_i X \subseteq X.$$

:⊢ **Definição 11.13** (Independência convexa (afim))**.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$. Uma lista de vetores *convexamente independentes* (ou *afimente independentes*) é uma lista $v_0, \ldots, v_d \subseteq V$ tais que $v_1 - v_0, \ldots, v_d - v_0$ são linearmente independentes.

11.1.7.1 Simplexos

:⊢ **Definição 11.14.** Sejam $(\boldsymbol{C}, \leq)$ um corpo ordenado e $d \in \mathbb{N}$. O *d-simplexo unitário* (ou simplexo unitário d-dimensional) sobre $\boldsymbol{C}$ é o conjunto

$$\Delta^d_{\boldsymbol{C}} = \left\{ c \in C^{d+1}_{\geq 0} \;\middle|\; \mathop{+}_{k\in[d+1]} c_k = 1 \right\}.$$

O *d-simplexo unitário real* é o d-simplexo unitário sobre o corpo dos números reais, denotado

$$\mathbb{\Delta}^d := \Delta^d_{\mathbb{R}} = \left\{ x \in \mathbb{R}^{d+1}_{\geq 0} \;\middle|\; \mathop{+}_{k\in[d+1]} x_k = 1 \right\}.$$

O simplexo unitário é portanto o conjunto de pontos de $c = (c_0, \ldots, c_d) \in C^{d+1}$ tais que, para todo $k \in [d+1]$, $c_k \geq 0$ e satisfazendo $+_{k\in[d+1]} c_k = 1$.

(A) O 0-simplexo é um ponto.

(B) O 1-simplexo é um intervalo.

(C) O 2-simplexo é um triângulo.

(D) O 3-simplexo é um tetraedro.

FIGURA 11.1: Simplexos unitários reais $\mathbb{\Delta}^0$, $\mathbb{\Delta}^1$, $\mathbb{\Delta}^2$ e $\mathbb{\Delta}^3$. Os eixos estão representados em cinza para referência; os simplexos estão representados em preto e, nos casos de dimensão 2 e 3, suas faces estão representadas com opacidade para que os eixos sejam visíveis.

:⊢ **Definição 11.15.** Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo ordenado $(\boldsymbol{C}, \leq)$ e $v_0, \ldots, v_d \in E$ convexamente independentes. O *simplexo gerado* por $v_0, \ldots, v_d$ é o conjunto

$$[v_0, \ldots, v_d]_{\Delta^d} = \left\{ \mathop{+}_{k \in [d+1]} c_k v_k = 1 \;\middle|\; c \in \Delta^d_C \right\}$$

e o interior desse simplexo é

$$]v_0, \ldots, v_d[_{\Delta^d} = \left\{ \mathop{+}_{k \in [d+1]} c_k v_k = 1 \;\middle|\; c \in \Delta^d_C \right\}.$$

No caso real, denota-se

$$[v_0, \ldots, v_d]_{\mathbb{\Delta}^d} = \left\{ \mathop{+}_{k \in [d+1]} x_k v_k = 1 \;\middle|\; x \in \mathbb{\Delta}^d \right\}$$

e

$$]v_0, \ldots, v_d[_{\mathbb{\Delta}^d} = \left\{ \mathop{+}_{k \in [d+1]} x_k v_k = 1 \;\middle|\; x \in \mathbb{\Delta}^d \right\}.$$

Em particular, o 1-simplexo é o conjunto $\Delta^1 = \{(1-t,t) \mid t \in [0,1]\}$ e dados $v, v' \in E$,

$$[v, v']_{\Delta^1} = \{(1-t)v + tv' \mid t \in [0,1]\}$$

e

$$]v, v'[_{\Delta^1} = \{(1-t)v + tv' \mid t \in]0,1[\}\,.$$

Capítulo 12

Espaços afins

12.1 Espaço e subespaço afins

:⊢ **Definição 12.1.** Seja $\boldsymbol{C}$ um corpo. Um *espaço afim* sobre $\boldsymbol{C}$ é uma tripla $\boldsymbol{A} = (A, \vec{\boldsymbol{A}}, +)$ em que

1. A é um conjunto, o *conjunto de pontos* de $\boldsymbol{A}$;
2. $\vec{\boldsymbol{A}}$ é um espaço linear sobre $\boldsymbol{C}$, o *espaço de translações* de $\boldsymbol{A}$;
3. $+\colon \vec{\boldsymbol{A}} \curvearrowright A$ é uma ação de grupo simplesmente transitiva[1] (transitiva e livre), a *translação* de $\boldsymbol{A}$, denotada por

$$\begin{aligned} +\colon \vec{A} &\longrightarrow \overset{\leftrightarrow}{\mathcal{F}}(A) \\ v &\longmapsto +v\colon A \longrightarrow A \\ &\qquad\qquad a \longmapsto a + v \end{aligned}$$

A dimensão de $\boldsymbol{A}$ é a dimensão de $\vec{\boldsymbol{A}}$.

Não distinguiremos a notação da ação $+\colon \vec{\boldsymbol{A}} \curvearrowright A$, da operação binária de grupo $+\colon \vec{A}^2 \longrightarrow \vec{A}$ e da adição do corpo $+\colon C^2 \longrightarrow C$, pois elas são todas compatíveis entre si, mas deve-se ter em mente que são funções distintas. Sempre que possível, denotaremos, para todos $a \in A$ e $v \in \vec{A}$, $a - v := a + (-v)$.

⊢ **Proposição 12.1.** *Sejam A um conjunto, $\vec{\boldsymbol{A}}$ um espaço linear sobre um corpo $\boldsymbol{C}$ e $+\colon \vec{\boldsymbol{A}} \curvearrowright A$ ação. São equivalentes:*

1. *A ação $+$ é simplesmente transitiva;*
2. *(Afinidade) Para todo $a \in A$,*

$$\begin{aligned} a+\colon \vec{A} &\longrightarrow A \\ v &\longmapsto a + v \end{aligned}$$

[1] Também chamada de *regular*.

é uma bijeção;

3. *(Subtração) Para todos $a, a' \in A$, existe único $v \in \vec{A}$ tal que $a + v = a'$.*

Por causa dessa proposição, podemos definir uma função de subtração em A, que toma $a, a' \in A$ e dá o único vetor $v \in \vec{A}$ tal que $a + v = a'$, que será denotado $a' - a := v$, de modo a termos a igualdade

$$a + (a' - a) = a'.$$

:⊢ **Definição 12.2.** Seja $\boldsymbol{A} = (A, \vec{\boldsymbol{A}}, +)$ um espaço afim sobre um corpo $\boldsymbol{C}$. A *subtração* em A é a função

$$\begin{aligned} -\colon A^2 &\longrightarrow \vec{A} \\ (a, a') &\longmapsto a' - a. \end{aligned}$$

É importante notar que não está definida adição entre pontos de A, de modo que $a + (a' - a)$ não pode ser escrita como $(a + a') - a$. A seguir estão duas propriedades conhecidas como axiomas de Weyl.

⊢ **Proposição 12.2.** *Seja $\boldsymbol{A} = (A, \vec{\boldsymbol{A}}, +)$ um espaço afim sobre um corpo $\boldsymbol{C}$.*

1. *Para todos $a \in A$ e $v \in \vec{A}$, existe único $a' \in A$ tal que $a' - a = v$;*
2. *(Paralelogramo) Para todos $a, a', a'' \in A$,*

$$(a'' - a') + (a' - a) = a'' - a.$$

□ *Demonstração.* Ambas propriedades seguem da propriedade de subtração da proposição anterior.

1. Para a existência, tomamos $a' := a + v$. Pela propriedade de subtração da proposição anterior, $v \in \vec{A}$ é o único vetor tal que $a + v = a'$ e como, por definição, $a + (a' - a) = a'$, segue que $v = a' - a$. Para a unicidade, sejam $a', a'' \in A$ tais que $a' - a = v = a'' - a$. Então

$$a' = a + (a' - a) = a + (a'' - a) = a''.$$

2. Basta notar que

$$a + (a'' - a') + (a' - a) = a + (a' - a) + (a'' - a') = a' + (a'' - a') = a'',$$

portanto, da unicidade da propriedade de subtração,

$$(a'' - a') + (a' - a) = (a'' - a).$$

■

▷ **Exercício 12.1.** *Seja* $\boldsymbol{A} = (A, \vec{\boldsymbol{A}}, +)$ *um espaço afim sobre um corpo* $\boldsymbol{C}$*. Para todos* $a, a', a'' \in A$*,*

$$(a'' - a) - (a' - a) = (a'' - a').$$

:⊢ **Definição 12.3.** Seja $\boldsymbol{A} = (A, \vec{\boldsymbol{A}}, +)$ um espaço afim sobre um corpo $\boldsymbol{C}$. Um *subespaço afim* de $\boldsymbol{A}$ é uma tripla $\boldsymbol{S} = (S, \vec{\boldsymbol{S}}, +_S)$ tal que

1. $S \subseteq A$;
2. Existe $a \in A$ tal que $\vec{S} := \{s - a \mid s \in S\}$ é subespaço linear de $\vec{A}$;
3. $+_S$ é a ação $+\colon \vec{\boldsymbol{A}} \curvearrowright A$ restrita a $\vec{S}$.

Isso não depende da escolha de $a \in A$.

12.2 Transformação afim

:⊢ **Definição 12.4.** Sejam $\boldsymbol{A}$ e $\boldsymbol{A}'$ espaços afins sobre um corpo $\boldsymbol{C}$. Um *transformação afim* de $\boldsymbol{A}$ para $\boldsymbol{A}'$ é uma função $f\colon A \longrightarrow B$ tal que

1. Para todos $a, a', b, b' \in A$ tais que $a' - a = b' - b$,

$$f(a') - f(a) = f(b') - f(b);$$

2. A função

$$\begin{aligned} \vec{f}\colon \vec{A} &\longrightarrow \vec{A}' \\ a' - a &\longmapsto f(a') - f(a) \end{aligned}$$

é linear.

Denota-se $f\colon \boldsymbol{A} \longrightarrow \boldsymbol{A}'$. O conjunto de todas transformações afins é denotado $\vec{\boldsymbol{A}} \rtimes \mathcal{L}(\vec{\boldsymbol{A}})$.

Isso implica que, para todos $a \in A$ e $v \in \vec{A}$,

$$f(a + v) = f(a) + \vec{f}(v)$$

e, portanto, f está unicamente determinada por seu valor $f(a)$ e pela função linear $\vec{f}$, como mostra a proposição a seguir.

⊢ **Proposição 12.3.** *Sejam* $\boldsymbol{A}$ *e* $\boldsymbol{A}'$ *espaços afins sobre um corpo* $\boldsymbol{C}$ *e* $f\colon A \longrightarrow B$ *uma função. Então* f *é uma transformação afim de* $\boldsymbol{A}$ *para* $\boldsymbol{A}'$ *se, e somente se, existe transformação linear* $L\colon \vec{\boldsymbol{A}} \longrightarrow \vec{\boldsymbol{A}}'$ *tal que, para todos* $a \in A$ *e* $v \in \vec{A}$*,*

$$f(a + v) = f(a) + L(v).$$

Nesse caso, a transformação linear L *é única e* $L = \vec{f}$*.*

□ *Demonstração.* ($\Rightarrow$) Suponhamos que f é afim. Nesse caso, basta tomar $L := \vec{f}$ e segue que L, é linear e, para todos $a \in A$ e $v \in \vec{A}$,

$$\begin{aligned} f(a+v) &= f(a) + (f(a+v) - f(a)) \\ &= f(a) + f((a+v) - a) \\ &= f(a) + \vec{f}(v) \\ &= f(a) + L(v). \end{aligned}$$

($\Leftarrow$) Suponhamos que existe tal L linear. Então, para todos $a, a' \in A$,

$$f(a') - f(a) = f(a + (a' - a)) - f(a) = (f(a) + L(a' - a)) - f(a) = L(a' - a).$$

Isso implica que

0.1. Para todos $a, a', b, b' \in A$ tais que $a' - a = b' - b$,

$$f(a') - f(a) = L(a' - a) = L(b' - b) = f(b') - f(b);$$

0.2. A função $\vec{f} = L$, logo é linear.

A unicidade segue pois, se L, L' são funções lineares satisfazendo a propriedade da proposição, então, para todo $v \in \vec{A}$, tomamos $a \in A$ e então

$$\begin{aligned} L(v) &= (f(a) + L(v)) - f(a) \\ &= f(a+v) - f(a) \\ &= (f(a) + L'(v)) - f(a) \\ &= L'(v). \end{aligned}$$

■

⊢ **Proposição 12.4.** *1. (Fechamento) Sejam $\boldsymbol{A}$, $\boldsymbol{A}'$ e $\boldsymbol{A}''$ espaços afins sobre um corpo $\boldsymbol{C}$ e $f\colon \boldsymbol{A} \longrightarrow \boldsymbol{A}'$ e $f'\colon \boldsymbol{A}' \longrightarrow \boldsymbol{A}''$ transformações afins. Então $f' \circ f\colon \boldsymbol{A} \longrightarrow \boldsymbol{A}''$ é uma transformação afim e $(f' \vec{\circ} f) = \vec{f'} \circ \vec{f}$;*

2. (Identidade) Seja $\boldsymbol{A}$ um espaço afim sobre um corpo $\boldsymbol{C}$. A identidade $\mathrm{I}_A\colon \boldsymbol{A} \longrightarrow \boldsymbol{A}$ é uma transformação afim e $\vec{\mathrm{I}}_A = \mathrm{I}_{\vec{A}}$;

□ *Demonstração.* 1. Para todos $a \in A$ e $v \in \vec{A}$,

$$\begin{aligned} f' \circ f(a+v) &= f'(f(a) + \vec{f}(v)) \\ &= f'(f(a)) + \vec{f'}(\vec{f}(v)) \\ &= f' \circ f(a) + \vec{f'} \circ \vec{f}(v). \end{aligned}$$

Como $\vec{f'} \circ \vec{f}$ é linear, segue que $f' \circ f$ é afim e $(f' \vec{\circ} f) = \vec{f'} \circ \vec{f}$.

2. Para todos $a \in A$ e $v \in \vec{A}$,

$$\mathrm{I}_A(a+v) = a+v = \mathrm{I}_A(a) + \mathrm{I}_{\vec{A}}(v).$$

Como $\mathrm{I}_{\vec{A}}$ é linear, segue que I_A é linear e $\vec{\mathrm{I}}_A = \mathrm{I}_{\vec{A}}$.

■

12.3 Bases e coordenadas

12.3.1 Coordenadas baricêntricas

⊢ **Proposição 12.5.** *Sejam* $\boldsymbol{A}$ *um espaço afim sobre um corpo* $\boldsymbol{C}$, $n \in \mathbb{N}$ *e* $a_1, \ldots, a_n \in A$, $c_1, \ldots, c_n \in C$.

1. *Se* $c_1 + \cdots + c_n = 0$, *existe único* $v \in \vec{A}$ *tal que, para todo* $o \in A$,

$$v = c_1(a_1 - o) + \ldots + c_n(a_n - o);$$

2. *Se* $c_1 + \cdots + c_n = 1$, *existe único* $a \in A$ *tal que, para todo* $o \in A$,

$$a - o = c_1(a_1 - o) + \ldots + c_n(a_n - o).$$

□ *Demonstração.* Segue da propriedade do paralelogramo que, para todos $a, a', o \in A$,

$$(a' - o) - (a - o) = (a' - a).$$

1. Sejam $o, o' \in A$ e definamos

$$v := c_1(a_1 - o) + \ldots + c_n(a_n - o).$$

Mostraremos que $v = c_1(a_1 - o') + \ldots + c_n(a_n - o')$. A unicidade segue da

definição de v. Como $c_1 + \cdots + c_n = 0$, então $c_n = -(c_1 + \cdots + c_{n-1})$, logo

$$\begin{aligned}
v &= \mathop{+}_{i=1}^{n} c_i(a_i - o) \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - o) - \mathop{+}_{i=1}^{n-1} c_i(a_n - o) \\
&= \mathop{+}_{i=1}^{n-1} c_i((a_i - o) - (a_n - o)) \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - a_n) \\
&= \mathop{+}_{i=1}^{n-1} c_i((a_i - o') - (a_n - o')) \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - o') - \mathop{+}_{i=1}^{n-1} c_i(a_n - o') \\
&= \mathop{+}_{i=1}^{n} c_i(a_i - o').
\end{aligned}$$

2. Sejam $o, o' \in A$. Pela proposição 12.2, existe único $a \in A$ tal que

$$a - o = c_1(a_1 - o) + \ldots + c_n(a_n - o).$$

Mostraremos que $a - o' = c_1(a_1 - o') + \ldots + c_n(a_n - o')$. A unicidade segue da definição de a. Como $c_1 + \cdots + c_n = 1$, então $c_n = 1 - (c_1 + \cdots + c_{n-1})$, logo

$$\begin{aligned}
a - o &= \mathop{+}_{i=1}^{n} c_i(a_i - o) \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - o) + \left(1 - \mathop{+}_{i=1}^{n-1} c_i\right)(a_n - o) \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - a_n) + (a_n - o).
\end{aligned}$$

Como

$$(a - o) - (a_n - o) = (a - a_n) = (a - o') - (a_n - o'),$$

segue que

$$\begin{aligned}
a - o' &= (a - o) - (a_n - o) + (a_n - o') \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - a_n) + (a_n - o) - (a_n - o) + (a_n - o') \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - a_n) + (a_n - o') \\
&= \mathop{+}_{i=1}^{n-1} c_i(a_i - o') + \left(1 - \mathop{+}_{i=1}^{n-1} c_i\right)(a_n - o') \\
&= \mathop{+}_{i=1}^{n} c_i(a_i - o).
\end{aligned}$$

■

Isso nos permite fazer a seguinte definição, que não depende de $o \in A$.

:⊢ **Definição 12.5.** Sejam $\boldsymbol{A}$ um espaço afim sobre um corpo $\boldsymbol{C}$, $n \in \mathbb{N}$ e $a_1, \ldots, a_n \in A$, $c_1, \ldots, c_n \in C$ tais que $c_1 + \ldots + c_n = 1$. O *baricentro* (ou *centroide*) de $(a_i)_{i\in[n]}$ com *pesos* $(c_i)_{i\in[n]}$ é o ponto

$$c_1a_1 + \cdots + c_na_n := o + c_1(a_1 - o) + \ldots + c_n(a_n - o).$$

Além disso, podemos também definir, para $c_1 + \cdots + c_n = 0$, o vetor

$$c_1a_1 + \cdots + c_na_n := c_1(a_1 - o) + \cdots + c_n(a_n - o).$$

Esse vetor pode ser interpretado da seguinte maneira. Primeiro, notemos que para $n = 2$, $c_1 = 1$ e $c_2 = -1$, temos o vetor $a_0 - a_1$ da propriedade de subtração, que leva a para a'. Agora, no caso geral em que $c_1 + \cdots + c_n = 0$ e $v = c_1a_1 + \cdots + c_na_n$, separamos os coeficientes positivos, denotados c_j^+, e seus respectivos pontos p_j, dos negativos, denotados $-c_k^-$, e seus respectivos pontos n_k de modo que temos

$$\Delta := \mathop{+}_{j} c_j^+ = \mathop{+}_{j} c_j^+ - \mathop{+}_{i} c_i = \mathop{+}_{k} c_k^-.$$

Assim, podemos concluir que $\frac{v}{\Delta}$ é o vetor que desloca o baricentro dos pontos negativos n_k com pesos $\frac{c_k^-}{\Delta}$ ao baricentro dos pontos positivos p_j com pesos $\frac{c_k^+}{\Delta}$, ou seja,

$$\frac{1}{\Delta}(c_1a_1 + \cdots + c_na_n) = \mathop{+}_{j} \frac{c_j^+}{\Delta} p_j - \mathop{+}_{k} \frac{c_k^-}{\Delta} n_k.$$

Capítulo 13

Álgebra universal

13.1 Álgebra e estrutura algébrica

:⊢ **Definição 13.1.** Sejam A um conjunto e $n \in \mathbb{N}$. Uma *operação n-ária* em A é uma função

$$O\colon A^n \longrightarrow A.$$

A *aridade* de O é o número n.

Aqui a notação A^n denota o produto de conjuntos $\prod_{i\in[n]} A = A \times \cdots \times A$.

:⊢ **Definição 13.2.** Uma *álgebra* é um par $\boldsymbol{A} = (A, \mathcal{O})$ em que

1. A é um conjunto, seu *conjunto de pontos*;
2. $\mathcal{O} = (\mathcal{O}_n)_{n\in\mathbb{N}}$, sua *estrutura algébrica*, sendo $\mathcal{O}_n = (O_i)_{i\in\#(\mathcal{O}_n)}$ uma sequência finita de operações n-árias em A

$$O_i\colon A^n \longrightarrow A.$$

O *tipo* de $(A, \mathcal{O})$ é a sequência $(\#(\mathcal{O}_n))_{n\in\mathbb{N}}$.

Álgebras de mesmo tipo têm bijeção entre suas estruturas algébricas, dada por

$$O_i \in \mathcal{O}_n \mapsto O'_i \in \mathcal{O}'_n,$$

em que $n \in \mathbb{N}$ e $i \in \#(\mathcal{O}_n)$. Isso ocorre porque $\#(\mathcal{O}_n) = \#(\mathcal{O}'_n)$. É importante notar que a bijeção relaciona operações de mesma aridade n. Quando conveniente, as operações O_i e O'_i relacionadas pela bijeção serão denotadas simplesmente O_i por simplicidade mas deve-se ter em mente que são operações em conjuntos diferentes, a primeira em A e a segunda em A'.

13.2 Morfismos

:⊢ **Definição 13.3.** Sejam $\boldsymbol{A} = (A, \mathcal{O})$ e $\boldsymbol{A}' = (A', \mathcal{O}')$ álgebras de mesmo tipo. Um *morfismo* de $\boldsymbol{A}$ para $\boldsymbol{A}'$ é uma função $f\colon A \to A'$ tal que, para todo $n \in \mathbb{N}$, toda $O \in \mathcal{O}_n$ e todos $a_0, \ldots, a_{n-1} \in A$,

$$f(O(a_0, \ldots, a_{n-1})) = O'(f(a_0), \ldots, f(a_{n-1})).$$

Denota-se $f\colon \boldsymbol{A} \to \boldsymbol{A}'$. Um isomorfismo entre $\boldsymbol{A}$ e $\boldsymbol{A}'$ é um morfismo bijetivo de $\boldsymbol{A}$ para $\boldsymbol{A}'$. Denota-se $\boldsymbol{A} \simeq \boldsymbol{A}'$.

Isso é equivalente a

$$f \circ O = O' \circ (f, \ldots, f).$$

13.3 Subálgebra

:⊢ **Definição 13.4.** Seja $\boldsymbol{A} = (A, \mathcal{O})$ uma álgebra. Uma *subálgebra* de $\boldsymbol{A}$ é uma álgebra $\boldsymbol{S} = (S, \mathcal{O}^S)$ de mesmo tipo de $\boldsymbol{A}$ tal que

1. $S \subseteq A$;
2. Para todo $n \in \mathbb{N}$ e toda $O_i \in \mathcal{O}_n$, $O_i^S = O_i|_{S^n}$.

Denota-se $\boldsymbol{S} \leq \boldsymbol{A}$.

13.4 Produto

:⊢ **Definição 13.5.** Seja $(\boldsymbol{A}_i)_{i \in I} = ((A_i, \mathcal{O}_i))_{i \in I}$ uma família de álgebras de mesmo tipo. O *produto* da família $(\boldsymbol{A}_i)_{i \in I}$ é o par

$$\prod_{\boldsymbol{i} \in \boldsymbol{I}} \boldsymbol{A_i} := (A, \mathcal{O}),$$

em que

1. $A := \prod_{i \in I} A_i$;
2. $\mathcal{O} := (\mathcal{O}_n)_{n \in \mathbb{N}}$, e, para todo $n \in \mathbb{N}$ e todo $k \in \#(\mathcal{O}_n)$,

$$\begin{aligned} O_k\colon A^2 &\longrightarrow A \\ a_0, \cdots, a_{n-1} &\longmapsto ((O_0)_k(a_{0,i})_{i \in I}, \ldots, (O_{n-1})_k(a_{n-1,i})_{i \in I}). \end{aligned}$$

13.5 Congruências e quociente

Nesta subseção, trataremos de equivalências especiais. De modo geral, se $\equiv$ é uma equivalência em A, $n \in \mathbb{N}$ e $a = (a_0, \ldots, a_{n-1}) \in A^n$ e $a' = (a'_0, \ldots, a'_{n-1}) \in A^n$ são tais que, para todo $k \in [n]$, $a_k \equiv a'_k$, denotaremos isso por

$$a \equiv a'.$$

Isso é uma equivalência em A^n.

:⊢ **Definição 13.6.** Seja $\boldsymbol{A} = (A, \mathcal{O})$ uma álgebra. Uma *congruência* em $\boldsymbol{A}$ é uma equivalência $\equiv$ em A tal que

1. (Compatibilidade) Para todo $n \in \mathbb{N}$, toda $O \in \mathcal{O}_n$ e todos $a, a' \in A^n$ tais que $a \equiv a'$,
$$O(a) \equiv O(a').$$

Isso basicamente significa que uma congruência é uma equivalência em A compatível com as operações n-árias da sua estrutura algébrica. Essa compatibilidade pode ser enunciada de outra forma, como mostra a proposição a seguir.

⊢ **Proposição 13.1.** *Seja $\boldsymbol{A} = (A, \mathcal{O})$ uma álgebra. Uma equivalência $\equiv$ em A é uma congruência em $\boldsymbol{A}$ se, e somente se, é uma subálgebra de $\boldsymbol{A}^2$ (com as operações).*

□ *Demonstração.* Basta notar que, para todo $n \in \mathbb{N}$, toda $O \in \mathcal{O}_n$ e todos $a, a' \in A^n$ tais que $a \equiv a'$, como por definição das operações de $\boldsymbol{A}$ no produto $\boldsymbol{A}^2$ temos $O((a_i, a'_i)_{i\in[n]}) = (O(a), O(a'))$, então $O(a) \equiv O(a')$ é equivalente a $O((a_i, a'_i)_{i\in[n]}) \in \equiv$, ou seja, a compatibilidade de $\equiv$ é equivalente a O pode ser restrita ao subconjunto $\equiv \subseteq A^2$. ■

Como toda equivalência, podemos usar uma congruência para quocientar o conjunto A. A compatibilidade garante que o quociente tenha uma estrutura algébrica de mesmo tipo induzida pela estrutura algébrica de $\boldsymbol{A}$.

:⊢ **Definição 13.7.** Sejam $\boldsymbol{A} = (A, \mathcal{O})$ uma álgebra e $\equiv$ uma congruência em $\boldsymbol{A}$. O *quociente* de $\boldsymbol{A}$ por $\equiv$ é o par $\boldsymbol{A}/\!\equiv\; := (A/\!\equiv, \mathcal{O}^{\equiv})$, em que

1. O conjunto
$$A\big/_{\equiv} = \{[\![a]\!] = \{a' \in A \mid a' \equiv a\} \mid a \in A\}$$
é o conjunto quociente;
2. $\mathcal{O}^{\equiv} := (\mathcal{O}^{\equiv}_n)$, em que $\mathcal{O}^{\equiv}_n := (O^{\equiv}_i)$, sendo $O^{\equiv}_i$ definida, para todos a_0, …, $a_{n-1} \in A$, por
$$O^{\equiv}_i([\![a_0]\!], \ldots, [\![a_{n-1}]\!]) := [\![O(a_0, \ldots, a_{n-1})]\!].$$

⊢ **Proposição 13.2.** *Sejam $\boldsymbol{A} = (A, \mathcal{O})$ uma álgebra e $\equiv$ uma congruência em $\boldsymbol{A}$. O quociente $\boldsymbol{A}/\!\equiv\, := (A/\!\equiv, \mathcal{O}^{\equiv})$ é uma estrutura algébrica do mesmo tipo que $\boldsymbol{A}$ e a projeção quociente $[\![\cdot]\!] : A \to A/\!\equiv$ é um morfismo de álgebras.*

□ *Demonstração.* Devemos mostrar que, para todo $n \in \mathbb{N}$ e toda $O \in \mathcal{O}_n$, O é uma operação n-ária em $A/\!\equiv$. Sejam $a_0, \ldots, a_{n-1} \in A$ e $a'_0, \ldots, a'_{n-1} \in A$ tais que, para todo $i \in [n]$, $a_i \equiv a'_i$. Então

$$\begin{aligned} O_i^{\equiv}([\![a_0]\!], \ldots, [\![a_{n-1}]\!]) &= [\![O(a_0, \ldots, a_{n-1})]\!] \\ &= [\![O(a'_0, \ldots, a'_{n-1})]\!] \\ &= O_i^{\equiv}([\![a'_0]\!], \ldots, [\![a'_{n-1}]\!]), \end{aligned}$$

o que mostra que $O_i^{\equiv}$ está bem definida.

Por fim, notemos que a projeção é morfismo por definição da estrutura quociente, pois, para todo $n \in \mathbb{N}$, toda $O \in \mathcal{O}_n$ e todos $a_0, \ldots, a_{n-1} \in A^n$,

$$[\![O(a_0, \ldots, a_{n-1})]\!] = O^{\equiv}([\![a_0]\!], \ldots, [\![a_{n-1}]\!]). \qquad \blacksquare$$

▷ **Exercício 13.1.** *1. Existe uma bijeção entre os subgrupos normais de um grupo $\boldsymbol{G}$ e as congruências em $\boldsymbol{G}$.*

2. Existe uma bijeção entre os ideais de um anel $\boldsymbol{A}$ e as congruências em $\boldsymbol{A}$.

13.6 Núcleo e imagem

:⊢ **Definição 13.8.** Sejam $\boldsymbol{A}$ e $\boldsymbol{A}'$ álgebras de mesmo tipo e $f\colon \boldsymbol{A} \to \boldsymbol{A}'$ um morfismo de álgebras. O *núcleo* de f é o conjunto

$$\operatorname{nuc}(f) := \left\{(a, a') \in A^2 \;\middle|\; f(a) = f(a')\right\} \subseteq A^2.$$

⊢ **Proposição 13.3** (Isomorfismo de imagem-coimagem)**.** *Sejam $\boldsymbol{A}$ e $\boldsymbol{A}'$ álgebras de mesmo tipo e $f\colon \boldsymbol{A} \to \boldsymbol{A}'$ um morfismo de álgebras.*

1. *O núcleo* $\operatorname{nuc}(f)$ *é uma congruência em $\boldsymbol{A}$;*
2. *A imagem* $\operatorname{im}(f)$ *é uma subálgebra de $\boldsymbol{A}'$;*
3. *O quociente de A pelo núcleo de f é isomorfo à imagem de f:*

$$\boldsymbol{A}\big/_{\mathbf{nuc}(\boldsymbol{f})} \simeq \mathbf{im}\,(\boldsymbol{f}).$$

$$\boldsymbol{A}\big/_{\equiv_{\boldsymbol{f}}} \simeq \boldsymbol{f}(\boldsymbol{A}).$$

□ *Demonstração.* 1. A equivalência definida pelo núcleo é $a \equiv_f a'$ se, e somente se, $f(a) = f(a')$. Para ver que é equivalência, notemos que

1.1. Para todo $a \in A$, $f(a) = f(a)$, logo $a \equiv_f a$;
1.2. Para todos $a, a \in A$ tais que $a \equiv_f a'$, vale $f(a) = f(a')$, logo $f(a') = f(a)$, portanto $a' \equiv_f a$;
1.3. Para todos $a, a', a'' \in A$ tais que $a \equiv_f a'$ e $a' \equiv_f a''$, vale $f(a) = f(a')$ e $f(a') = f(a'')$, logo $f(a) = f(a'')$, portanto $a \equiv_f a''$.

Para mostrarmos a compatibilidade, sejam $O_i \in \mathcal{O}_n$ e $a, a' \in A^n$ tais que $a \equiv_f a'$. Isso significa que, para todo $k \in [n]$, vale $a_k \equiv_f a'_k$, ou seja, $f(a_k) = f(a'_k)$. Como f é morfismo, segue que

$$\begin{aligned} f(O_i(a)) &= O'_i(f(a_0), \ldots, f(a_{n-1})) \\ &= O'_i(f(a'_0), \ldots, f(a'_{n-1})) \\ & f(O_i(a')), \end{aligned}$$

portanto $O_i(a) \equiv_f O_i(a')$, o que mostra que $\equiv_f$ é uma congruência.

2. Exercício.
3. Exercício.

■

Parte 3

Geometria

Capítulo 14

Topologia

14.1 Espaços topológicos

14.1.1 Topologia, abertos e fechados

A noção de topologia que será abordada nesta parte do livro é a topologia baseada em teoria dos conjuntos.

:⊢ **Definição 14.1.** Seja X um conjunto. Uma *topologia* de X é um conjunto $\mathcal{T} \subseteq \wp(X)$ que satisfaz

1. Vazio e conjunto todo são abertos.

$$\emptyset, X \in \mathcal{T}$$

2. União de abertos é aberto.

$$(A_i)_{i \in I} \subseteq \mathcal{T} \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{T}$$

3. Interseção finita de abertos é aberto.

$$(A_i)_{i \in [n]} \subseteq \mathcal{T} \Rightarrow \bigcap_{i=0}^{n-1} A_i \in \mathcal{T}$$

⊢ **Proposição 14.1.** *Seja X um conjunto.*

1. *(Topologia trivial) $\{\emptyset, X\}$ é uma topologia de X;*
2. *(Topologia discreta) $\wp(X)$ é uma topologia de X.*

□ *Demonstração.* 1. Primeiramente, $\emptyset$ e X pertencem a $\{\emptyset, X\}$. Agora, consideremos uma família $(A_i)_{i \in I} \subseteq \{\emptyset, X\}$ de abertos. Caso $A_i = \emptyset$ para todo

$i \in I$, então $\bigcup_{i\in I} A_i = \emptyset \in \mathcal{T}$. Caso contrário, existe $j \in I$ tal que $A_j = X$, o que implica $\bigcup_{i\in I} A_i = X \in \mathcal{T}$. Assim, concluímos que vale a segunda propriedade. Para mostrar a terceira propriedade, seja $(A_i)_{i\in[n]} \subseteq \mathcal{T}$ uma família finita de abertos. Caso $A_i = X$ para todo $i \in I$, o que implica $\bigcap_{i=0}^{n-1} A_i = X \in \mathcal{T}$, Caso contrário, existe $j \in I$ tal que $A_j = \emptyset$, o que implica $\bigcap_{i=0}^{n-1} A_i = \emptyset \in \mathcal{T}$.

2. Todas propriedades valem pois $\mathcal{T} = \wp(X)$. ■

:⊢ **Definição 14.2.** Um *espaço topológico* é um par $(X, \mathcal{T})$ em que X é um conjunto não vazio e $\mathcal{T}$ é uma topologia de X. Um *aberto* de $(X, \mathcal{T})$ é um conjunto de $\mathcal{T}$. Um *fechado* de $(X, \mathcal{T})$ é um conjunto cujo complementar em X é um aberto de $(X, \mathcal{T})$. O conjunto dos fechados de $(X, \mathcal{T})$ é denotado $\mathcal{T}_\bullet$.

⊢ **Proposição 14.2** (Dualidade de abertos e fechados)**.** *Seja X um conjunto. Então*

1. *Vazio e conjunto todo são fechados.*

$$\emptyset, X \in \mathcal{T}_\bullet$$

2. *Interseção de fechados é fechado.*

$$(F_i)_{i\in I} \subseteq \mathcal{T}_\bullet \Rightarrow \bigcap_{i\in I} F_i \in \mathcal{T}_\bullet$$

3. *União finita de fechados é fechado.*

$$(F_i)_{i\in[n]} \subseteq \mathcal{T}_\bullet \Rightarrow \bigcup_{i=0}^{n-1} F_i \in \mathcal{T}_\bullet$$

□ *Demonstração.* Todas as demonstrações dependem de propriedades básicas de teoria de conjuntos.

1. Como $\emptyset, X \in \mathcal{T}$, $\emptyset^\complement = X$ e $X^\complement = \emptyset$, segue que $\emptyset, X \in \mathcal{T}_\bullet$.
2. Seja $(F_i)_{i\in I} \subseteq \mathcal{T}_\bullet$. Então $((F_i)^\complement)_{i\in I} \subseteq \mathcal{T}$, o que implica que $\bigcup_{i\in I}(F_i)^\complement \in \mathcal{T}$. Para concluir a demonstração, basta notar que

$$\left(\bigcup_{i\in I}(F_i)^\complement\right)^\complement = \bigcap_{i\in I} F_i.$$

3. Seja $(F_i)_{i\in[n]} \subseteq \mathcal{T}_\bullet$. Então $((F_i)^\complement)_{i\in[n]} \subseteq \mathcal{T}$, o que implica que $\bigcap_{i=0}^{n-1}(F_i)^\complement \in \mathcal{T}$. Para concluir a demonstração, basta notar que

$$\left(\bigcap_{i=0}^{n-1}(F_i)^\complement\right)^\complement = \bigcup_{i=0}^{n-1} F_i.$$

■

14.1.1.1 Interior e fecho

Intuitivamente, sabemos dizer quais pontos de um subconjunto da reta, do plano ou do espaço estão dentro do conjunto, quais estão fora e quais formam uma espécie de fronteira entre a parte de dentro e a de fora. Para uma bola aberta de raio unitário e centro na origem, é claro que os pontos de norma menor que 1 são os pontos do interior da bola, os pontos de norma igual a 1 são os pontos de fronteira e os pontos com norma maior que 1 são os pontos do exterior. Para conjuntos mais complicados, no entanto, parece ser mais difícil dizer o que está dentro e o que está fora. Se analisamos o conjunto dos racionais na reta, não é óbvio o que está dentro, o que está fora, o que é fronteira.

Além disso, a nossa intuição usa a ideia de norma ou distância no caso da bola e em espaços topológicos gerais isso não é possível. Para continuar a generalização dos conceitos topológicos existentes na reta, plano e espaço, devemos tentar formular a ideia de ponto interior usando os conceitos topológicos gerais, e fazemos isso notando que, no caso da bola e de conjuntos simples dos espaços tradicionais, um ponto está dentro do conjunto se existe um aberto em torno do ponto que está inteiramente contido no conjunto. Esse conceito não é o conceito que usaremos para definir o interior de um conjunto, mas é equivalente.

No começo dessa seção, o foco será o conceito de interior de um conjunto. A definição de interior que usaremos é a de que o interior de um conjunto é o maior conjunto aberto contido no conjunto. Maior, nesse caso, significa maior em relação à ordem parcial de contenção de conjuntos abertos. Como união de abertos é aberto, sabemos que a união de todos abertos contidos em um conjunto é aberto e, portanto, é o maior aberto contido no conjunto. Podemos perceber, ainda, que essa definição é equivalente à comentada acima pois, se um ponto está em um aberto do conjunto, ele está na união de todos eles e, se ele está na união, está, em particular, em algum aberto. Seguem abaixo a definição formal de interior de um conjunto e, em seguida, algumas propriedades básicas.

:⊢ **Definição 14.3.** Sejam $\boldsymbol{X}$ um espaço topológico e $C \subseteq X$. O *interior de* C é o conjunto aberto

$$C^\circ := \bigcup \{A \in \mathcal{T} \mid A \subseteq C\}.$$

De acordo com essa definição, o interior de um conjunto é o maior aberto contido no conjunto, no sentido de que qualquer aberto contido no conjunto está também contido no seu interior. Ainda, podemos ver o interior como um operador topológico — uma função $\mathcal{I} : \wp(X) \to \wp(X)$ que leva $C \mapsto C^\circ$. Nesse sentido, podemos pensar em propriedades que esse operador satisfaz. Algumas dessas propriedades estão na proposição abaixo. Além disso, se temos um operador qualquer em $\wp(X)$ que satisfaça algumas das propriedades abaixo, então esse operador é o operador

interior de alguma topologia de X. Isso está demonstrado na proposição que segue a proposição abaixo.

⊢ **Proposição 14.3.** *Sejam* $\boldsymbol{X}$ *um espaço topológico e* $A, B \subseteq X$. *Então*

1. $A^\circ \subseteq A$
2. $A \in \mathcal{T} \Leftrightarrow A^\circ = A$
3. $(A \cap B)^\circ = A^\circ \cap B^\circ$
4. $(A^\circ)^\circ = A^\circ$
5. $A \subseteq B \Rightarrow A^\circ \subseteq B^\circ$

□ *Demonstração.* Sejam $(A_i)_{i\in I}$, $(B_j)_{j\in J}$ e $(C_k)_{k\in K}$ indexações dos subconjuntos abertos de A, B e $A \cap B$, respectivamente.

1. Seja $a \in A^\circ$. Então existe $j \in I$ tal que $a \in A_j$. Como $A_i \subseteq A$ para todo $i \in I$, então $a \in A$, o que mostra que $A^\circ \subseteq A$.
2. Suponha que A um aberto. Como $A^\circ \subseteq A$ para todo $A \subseteq X$, basta mostrar que $A \subseteq A^\circ$. Seja $a \in A$. Se A é aberto, então existe $i \in I$ tal que $A_i = A$, o que implica $a \in A_i$ e, portanto, que $a \in A^\circ$, o que mostra que $A \subseteq A^\circ$. Assim, concluímos que $A^\circ = A$. Reciprocamente, suponha que $A^\circ = A$. Como A° é união de abertos, é um conjunto aberto e isso implica que A é aberto.
3. Vamos mostrar a inclusão $(A \cap B)^\circ \subseteq A^\circ \cap B^\circ$. Seja $a \in (A \cap B)^\circ$. Então existe $k \in K$ tal que $a \in C_k$. Como C_k é subconjunto aberto de $A \cap B$, então C_k é subconjunto aberto de A e de B, o que implica que $a \in A^\circ$ e $a \in B^\circ$. Assim, concluímos que $a \in A^\circ \cap B^\circ$. Reciprocamente, vamos mostrar a inclusão $A^\circ \cap B^\circ \subseteq (A \cap B)^\circ$. Seja $a \in A^\circ \cap B^\circ$. Então $a \in A^\circ$ e $a \in B^\circ$, o que implica que existem $i \in I$ e $j \in J$ tais que $a \in A_i$ e $a \in B_j$; ou seja, $a \in A_i \cap B_j$. Como A_i e B_j são abertos, sua interseção é aberto e, como $A_i \subseteq A$ e $B_j \subseteq B$, segue que $A_i \cap B_j \subseteq A \cap B$, e isso implica que $a \in (A \cap B)^\circ$.
4. Como A° é um conjunto aberto, pelo item 2 segue que $(A^\circ)^\circ = A^\circ$. ■

⊢ **Proposição 14.4.** *Sejam* X *um conjunto e* $\mathcal{I} : \wp(X) \to \wp(X)$ *uma função que satisfaz, para todos* $A, B \subseteq X$,

1. $\mathcal{I}(X) = X$;
2. $\mathcal{I}(A) \subseteq A$;
3. $\mathcal{I}(A \cap B) = \mathcal{I}(A) \cap \mathcal{I}(B)$;
4. $\mathcal{I}(\mathcal{I}(A)) = \mathcal{I}(A)$.

Então o conjunto $\mathcal{T} := \{C \subseteq X \mid C = \mathcal{I}(C)\}$ *é uma topologia de* X *e* $\mathcal{I}(C) = C^\circ$ *para cada* $C \subseteq X$.

□ *Demonstração.* Primeiro, notemos que, pela primeira propriedade, $X \in \mathcal{T}$ e, pela segunda propriedade, como $\mathcal{I}(\emptyset) \subseteq \emptyset$, concluímos que $\mathcal{I}(\emptyset) = \emptyset$, o que significa que $\emptyset \in \mathcal{T}$.

Vamos mostrar a seguinte propriedade: para todos $A, B \subseteq X$,

$$A \subseteq B \Rightarrow \mathcal{I}(A) \subseteq \mathcal{I}(B)$$

Supondo que $A \subseteq B$, temos que $A = A \cap B$. Então, pela terceira propriedade, $\mathcal{I}(A) = \mathcal{I}(A) \cap \mathcal{I}(B)$. Assim, segue que $\mathcal{I}(A) \subseteq \mathcal{I}(B)$.

Agora, seja $(A_i)_{i \in I}$ uma família de conjuntos em $\mathcal{T}$. Então, para cada $j \in I$, $A_j \subseteq \bigcup_{i \in I} A_i$ e, portanto, $\mathcal{I}(A_j) \subseteq \mathcal{I}(\bigcup_{i \in I} A_i)$. Mas a família satisfaz $\mathcal{I}(A_i) = A_i$. Assim, concluímos que

$$\bigcup_{i \in I} \mathcal{I}(A_i) = \bigcup_{i \in I} A_i \subseteq \mathcal{I}\left(\bigcup_{i \in I} A_i\right)$$

Pela segunda propriedade, $\mathcal{I}(\bigcup_{i \in I} A_i) \subseteq \bigcup_{i \in I} A_i$. Portanto $\bigcup_{i \in I} A_i = \mathcal{I}(\bigcup_{i \in I} A_i)$, e concluímos que $\bigcup_{i \in I} A_i \in \mathcal{T}$.

Então, seja $(A_i)_{i \in [n]}$ uma família finita de conjuntos em $\mathcal{T}$. Usando a terceira propriedade e indução, provaremos que $\bigcap_{i=0}^{n-1} A_i \in \mathcal{T}$. Para 0, a propriedade é trivialmente verdade. Suponhamos que vale para k. Então, pela terceira propriedade,

$$\begin{aligned}
\mathcal{I}\left(\bigcap_{i=0}^{k} A_i\right) &= \mathcal{I}\left(\left(\bigcap_{i=0}^{k-1} A_i\right) \cap A_k\right) \\
&= \mathcal{I}\left(\bigcap_{i=0}^{k-1} A_i\right) \cap \mathcal{I}\left(A_k\right) \\
&= \bigcap_{i=0}^{k-1} A_i \cap A_k \\
&= \bigcap_{i=0}^{k} A_i,
\end{aligned}$$

o que mostra que $\bigcap_{i=0}^{k} A_i \in \mathcal{T}$ e que, para todo $n \in \mathbb{N}$, $\bigcap_{i=0}^{n-1} A_i \in \mathcal{T}$. Logo concluímos que $\mathcal{T}$ é uma topologia de X.

Devemos, por fim, mostrar que $\mathcal{I}(C) = C^\circ$ para todo $C \subseteq X$. Seja $(C_I)_{i \in I}$ uma indexação dos subconjuntos abertos de C. Vamos mostrar primeiro que $C^\circ \subseteq \mathcal{I}(C)$. Para todo $i \in I$, temos que $C_i \subseteq C$ implica $\mathcal{I}(C_i) \subseteq \mathcal{I}(C)$. Como $C_i = \mathcal{I}(C_i)$, segue que $C_i \subseteq \mathcal{I}(C)$ e, portanto,

$$C^\circ = \bigcup_{i \in I} C_i \subseteq \mathcal{I}(C).$$

Por outro lado, notemos que $\mathcal{I}(C)$ é um aberto, pois, pela quarta propriedade, $\mathcal{I}(\mathcal{I}(C)) = \mathcal{I}(C)$. Pela segunda propriedade, $\mathcal{I}(C) \subseteq C$, e segue que $\mathcal{I}(C)$ é um dos subconjuntos abertos C_i de C. Portanto

$$\mathcal{I}(C) \subseteq \bigcup_{i \in I} C_i = C^\circ.$$

■

⊢ **Proposição 14.5.** *Sejam $(X, \mathcal{T})$ um espaço topológico e $C \subseteq X$. Então*

$$C^\circ = \{x \in X \mid \exists A \in \mathcal{T} \quad x \in A \subseteq C\}$$

□ *Demonstração.* Se o único aberto contido em C é $\emptyset$, então $C^\circ = \emptyset$ e segue que, para todo $x \in X$, todo aberto $A \in \mathcal{T}$ tal que $x \in A$ não está contido em C. Então os conjuntos são iguais. Se $C^\circ = \emptyset$, então o único aberto contido em

Se $x \in C^\circ$, então existe ■

Dualmente, consideraremos agora o conceito de fecho.

:⊢ **Definição 14.4.** Sejam $(X, \mathcal{T})$ um espaço topológico e $C \subseteq X$. O *fecho de C* é o conjunto fechado

$$\overline{C} := \bigcap \{F \in \mathcal{T}_\bullet \mid C \subseteq F\}.$$

⊢ **Proposição 14.6.** *Seja $\boldsymbol{X}$ um espaço topológico. Então*

1. *Para todo $A \subseteq X$, $A \subseteq \overline{A}$;*
2. *Um conjunto A é fechado se, e somente se, $\overline{A} = A$;*
3. *$\overline{(A \cup B)} = \overline{A} \cup \overline{B}$;*
4. *$\overline{(\overline{A})} = \overline{A}$;*

□ *Demonstração.* Demonstração análoga à de interior.

■

⊢ **Proposição 14.7.** *Sejam X um conjunto e $\mathcal{F} : \wp(X) \to \wp(X)$ uma função que satisfaz, para todos $A, B \subseteq X$,*

1. *$\mathcal{F}(\emptyset) = \emptyset$;*
2. *$A \subseteq \mathcal{F}(A)$;*
3. *$\mathcal{F}(A \cup B) = \mathcal{F}(A) \cup \mathcal{F}(B)$.*
4. *$\mathcal{F}(\mathcal{F}(A)) = \mathcal{F}(A)$.*

Então o conjunto $\mathcal{T}_\bullet := \{C \subseteq X \mid C = \mathcal{F}(C)\}$ é o conjunto de fechados de uma topologia $\mathcal{T}$ de X e $\mathcal{F}(C) = \overline{C}$ para cada $C \subseteq X$.

□ *Demonstração.* Demonstração análoga à de interior. ■

▷ **Exercício 14.1.** *Sejam $\boldsymbol{X}$ um espaço topológico e $A \subseteq X$. Então*

1. *$(\overline{A})^{\complement} = (A^{\complement})^\circ$;*
2. *$(A^\circ)^{\complement} = \overline{(A^{\complement})}$.*

14.1.1.2 Fronteira

:⊢ **Definição 14.5.** Sejam $\boldsymbol{X}$ um espaço topológico e $C \subseteq X$. A *fronteira* de C é o conjunto

$$\partial C := \overline{C} \setminus C^\circ.$$

▷ **Exercício 14.2.** *Sejam $\boldsymbol{X}$ um espaço topológico e $C \subseteq X$. Então*

1. *O fecho é a união disjunta de interior e fronteira.*

$$\overline{C} = C^\circ \cup \partial C \quad e \quad C^\circ \cap \partial C = \emptyset$$

2. *A fronteira é a interseção dos fechos do conjunto e de seu complementar.*

$$\partial C = \overline{C} \cap \overline{(C^{\complement})}$$

3. *Um conjunto é aberto se, e somente se, não contém pontos da fronteira.*

$$C \in \mathcal{T} \Leftrightarrow \partial C \cap C = \emptyset$$

4. *Um conjunto é fechado se, e somente se, contém todos pontos da fronteira.*

$$C \in \mathcal{T}_\bullet \Leftrightarrow \partial C \subseteq C$$

5. *A fronteira de um conjunto é fechado.*

$$\partial C \in \mathcal{T}_\bullet$$

6. $\partial(\partial(\partial C)) = \partial(\partial C)$.

14.1.2 Topologias geradas, bases e sub-bases

14.1.2.1 Topologias finas e grossas e topologias geradas

:⊢ **Definição 14.6.** Seja X um conjunto e $\mathcal{T}$ uma topologia de X. Uma *topologia mais grossa* (ou *fraca*) que a topologia $\mathcal{T}$ é uma topologia $\mathcal{T}'$ de X tal que $\mathcal{T}' \subseteq \mathcal{T}$. Uma *topologia mais fina* (ou *forte*) que a topologia $\mathcal{T}$ é uma topologia $\mathcal{T}'$ de X tal que $\mathcal{T} \subseteq \mathcal{T}'$.

A topologia mais fraca é a topologia trivial e a topologia mais forte é a topologia discreta.

⊢ **Proposição 14.8.** *Sejam X um conjunto e $(\mathcal{T}_i)_{i \in I}$ uma família de topologias de X. Então*

$$\mathcal{T} := \bigcap_{i \in I} \mathcal{T}_i$$

é uma topologia de X.

□ *Demonstração.* Primeiro, notemos que, para todo $i \in I$, $\emptyset, X \in \mathcal{T}_i$, e segue que $\emptyset, X \in \mathcal{T}$. Agora, seja $(A_j)_{j\in J} \subseteq \mathcal{T}$. Então, para todo $i \in I$, $(A_j)_{j\in J} \in \mathcal{T}_i$, e segue que $\bigcup_{j\in J} A_j \in \mathcal{T}_i$, o que implica que $\bigcup_{j\in J} A_j \in \mathcal{T}$. Por fim, seja $(A_j)_{j\in[n]} \subseteq \mathcal{T}$. Então, para todo $i \in I$, $(A_j)_{j\in[n]} \in \mathcal{T}_i$, e segue que $\bigcap_{j=0}^{n-1} A_j \in \mathcal{T}_i$, o que implica que $\bigcap_{j=0}^{n-1} A_j \subseteq \mathcal{T}$. ■

:⊢ **Definição 14.7.** Sejam X um conjunto, $G \subseteq \wp(X)$ e $(\mathcal{T}_i)_{i\in I}$ uma indexação do conjunto de todas as topologias de X das quais G é subconjunto. A *topologia gerada por* G é a topologia

$$\langle G \rangle := \bigcap_{i\in I} \mathcal{T}_i.$$

Nesse caso, dizemos que G é o *conjunto gerador* de $\langle G \rangle$ ou que G gera $\langle G \rangle$.

14.1.2.2 Bases e sub-bases

14.1.3 Funções contínuas

:⊢ **Definição 14.8.** Sejam $\boldsymbol{X} = (X, \mathcal{T}_X)$ e $\boldsymbol{Y} = (Y, \mathcal{T}_Y)$ espaços topológicos. Uma *função contínua* de $\boldsymbol{X}$ para $\boldsymbol{Y}$ é uma função $f : X \to Y$ tal que, para todo $A \in \mathcal{T}_Y$,

$$f^{-1}(A) \in \mathcal{T}_X.$$

Denota-se $f : \boldsymbol{X} \to \boldsymbol{Y}$. O conjunto dessas funções é denotado por $\mathscr{C}(\boldsymbol{X}, \boldsymbol{Y})$.

⊢ **Proposição 14.9.** (Propriedades categóricas).

1. *(Identidade) Seja* $\boldsymbol{X}$ *um espaço topológico. A função* $\mathrm{I}_X : X \to X$ *é uma função contínua.*
2. *(Associatividade) Sejam* $\boldsymbol{X_0}$, $\boldsymbol{X_1}$ *e* $\boldsymbol{X_2}$ *espaços topológicos e* $f_0 : \boldsymbol{X_0} \to \boldsymbol{X_1}$ *e* $f_1 : \boldsymbol{X_1} \to \boldsymbol{X_2}$ *funções contínuas. Então* $f_1 \circ f_0 : \boldsymbol{X_0} \to \boldsymbol{X_2}$ *é uma função contínua.*

$$\begin{tikzcd} \boldsymbol{X_0} \arrow[r, "f_0"] \arrow[rr, bend right, "f_1 \circ f_0"'] & \boldsymbol{X_1} \arrow[r, "f_1"] & \boldsymbol{X_2} \end{tikzcd}$$

□ *Demonstração.* 1. Seja $A \in \mathcal{T}$. Então $\mathrm{I}_X^{-1}(A) = A \in \mathcal{T}$, logo I_x é contínua.

2. Seja $A \in \mathcal{T}_2$. Como f_1 é contínua, $f_1^{-1}(A) \in \mathcal{T}_1$. Como f_0 é contínua, $f_0^{-1}(f_1^{-1}(A)) \in \mathcal{T}_0$. Portanto

$$(f_1 \circ f_0)^{-1}(A) = f_1^{-1} \circ f_0^{-1}(A) = f_0^{-1}(f_1^{-1}(A)) \in \mathcal{T}_0.$$

Logo $f_1 \circ f_0$ é contínua. ■

:⊢ **Definição 14.9.** Sejam $\boldsymbol{X_0}$ e $\boldsymbol{X_1}$ espaços topológicos. Um *homeomorfismo* de $\boldsymbol{X_0}$ para $\boldsymbol{X_1}$ é uma função contínua $f\colon \boldsymbol{X_0} \longrightarrow \boldsymbol{X_1}$ invertível cuja inversa é contínua. O conjunto de todos esses homeomorfismos é denotado por $\overleftrightarrow{\mathscr{C}}(\boldsymbol{X_0}, \boldsymbol{X_1})$. Esses espaços topológicos são *homeomorfos* e denota-se $\boldsymbol{X_0} \simeq \boldsymbol{X_1}$.

▷ **Exercício 14.3.** *Sejam $\boldsymbol{X_0}$, $\boldsymbol{X_1}$ e $\boldsymbol{X_2}$ espaços topológicos.*

1. *(Reflexividade)* $\boldsymbol{X_0} \simeq \boldsymbol{X_0}$;
2. *(Antissimetria)* $\boldsymbol{X_0} \simeq \boldsymbol{X_1} \Rightarrow \boldsymbol{X_1} \simeq \boldsymbol{X_0}$;
3. *(Transitividade)* $\boldsymbol{X_0} \simeq \boldsymbol{X_1}$ *e* $\boldsymbol{X_1} \simeq \boldsymbol{X_2} \Rightarrow \boldsymbol{X_0} \simeq \boldsymbol{X_2}$.

14.1.3.1 Suporte de funções vetoriais

:⊢ **Definição 14.10.** Sejam $\boldsymbol{X}$ um espaço topológico, $\boldsymbol{L}$ um espaço linear topológico e $f\colon X \to L$ uma função contínua. O *suporte fechado* de f é o conjunto

$$\operatorname{supp}^{\bullet}(f) := \overline{\operatorname{supp}(f)} = \overline{(f^{-1}(L \smallsetminus \{0\}))}.$$

14.1.4 Topologias induzidas

Nesta seção estudaremos como induzir topologias em conjuntos a partir de topologias que já temos. Isso será feito, em geral, de modo que uma ou mais funções sejam contínuas e a topologia induzida seja a menor ou a maior possível, dependendo do caso. Quando temos uma função de um conjunto em um espaço topológico, podemos induzir uma topologia nesse conjunto de modo que a topologia faça com que a função seja contínua. Nesse caso, a maior topologia que faz a função ser contínua é a topologia discreta, e o nosso interesse será achar a menor topologia tal que a função é discreta. Quando temos uma função de um espaço topológico em um conjunto, o caso se inverte. A menor topologia tal que a função é contínua sempre é a topologia trivial, e o nosso interesse está na maior topologia que garante a continuidade. De maneira parecida, podemos induzir topologias garantindo a continuidade de várias funções e também considerando funções injetivas e sobrejetivas quado necessário. O estudo desta seção envolve a definição dessas noções e a investigação inicial como sobre esses objetos se comportam e por que são os menores ou maiores com determinadas características.

14.1.4.1 Topologias puxada e inicial

:⊢ **Definição 14.11.** Sejam X um conjunto, $\boldsymbol{Y} = (Y, \mathcal{T}_Y)$ um espaço topológico e $f : X \to Y$ uma função. A *topologia puxada* por f de $\boldsymbol{Y}$ para X é

$$f^{\star}(\mathcal{T}_Y) := \left\{ f^{-1}(A) \;\middle|\; A \in \mathcal{T}_Y \right\}.$$

⊢ **Proposição 14.10.** *Sejam X um conjunto, $\boldsymbol{Y} = (Y, \mathcal{T}_Y)$ um espaço topológico e $f : X \to Y$ uma função. Então $f^\star(\mathcal{T}_Y)$, a topologia puxada por f de $\boldsymbol{Y}$ para X, é uma topologia de X.*

□ *Demonstração.* (1) Notemos que, como $\emptyset \in \mathcal{T}_Y$ e $\emptyset = f^{-1}(\emptyset)$, temos que $\emptyset \in f^\star(\mathcal{T}_Y)$. (2) Seja $(A_i)_{i\in I}$ uma família de conjuntos em $f^\star(\mathcal{T}_Y)$. Então, para cada $i \in I$, existe aberto $U_i \in \mathcal{T}_Y$ tal que $A_i = f^{-1}(U_i)$. Como $\mathcal{T}_Y$ é topologia, a união de abertos é aberto $\bigcup_{i\in I} U_i \in \mathcal{T}_Y$. Portanto

$$\bigcup_{i\in I} A_i = \bigcup_{i\in I} f^{-1}(U_i) = f^{-1}\left(\bigcup_{i\in I} U_i\right),$$

logo $\bigcup_{i\in I} A_i \in f^\star(\mathcal{T}_Y)$. (3) Seja $(A_i)_{i=1}^n$ uma família de conjuntos em $f^\star(\mathcal{T}_Y)$. Então, para cada $1 \leq i \leq n$, existe aberto $U_i \in \mathcal{T}_Y$ tal que $A_i = f^{-1}(U_i)$. Como $\mathcal{T}_Y$ é topologia, a interseção finita de abertos é aberto $\bigcap_{i=1}^n U_i \in \mathcal{T}_Y$. Portanto

$$\bigcap_{i=1}^n A_i = \bigcap_{i=1}^n f^{-1}(U_i) = f^{-1}\left(\bigcap_{i=1}^n U_i\right),$$

logo $\bigcap_{i=1}^n A_i \in f^\star(\mathcal{T}_Y)$. ■

⊢ **Proposição 14.11.** *Sejam $\boldsymbol{X} = (X, \mathcal{T}_X)$ e $\boldsymbol{Y} = (Y, \mathcal{T}_Y)$ espaços topológicos. Uma função $f : X \to Y$ é função contínua de $\boldsymbol{X}$ para $\boldsymbol{Y}$ se, e somente se, a topologia $f^\star(\mathcal{T}_Y)$ puxada por f de $\boldsymbol{Y}$ para X é uma subtopologia de $\mathcal{T}_X$.*

$$f \in \mathscr{C}(\boldsymbol{X}, \boldsymbol{Y}) \quad \Longleftrightarrow \quad f^\star(\mathcal{T}_Y) \subseteq \mathcal{T}_X.$$

□ *Demonstração.* Suponha que $f \in \mathscr{C}(\boldsymbol{X}, \boldsymbol{Y})$ e seja $B \in f^\star(\mathcal{T}_Y)$. Então existe $A \in \mathcal{T}_Y$ tal que $B = f^{-1}(A)$. Como f é contínua, segue que $f^{-1}(A) \in \mathcal{T}_X$, portanto, $f^\star(\mathcal{T}_Y) \subseteq \mathcal{T}_X$. Reciprocamente, suponha que $f^\star(\mathcal{T}_Y) \subseteq \mathcal{T}_X$. Então, para todo $A \in \mathcal{T}_Y$, $f^{-1}(A) \in f^\star(\mathcal{T}_Y)$, portanto $f^{-1}(A) \in \mathcal{T}_X$, o que mostra que $f \in \mathscr{C}(\boldsymbol{X}, \boldsymbol{Y})$. ■

:⊢ **Definição 14.12.** Sejam $\boldsymbol{X}$ um conjunto, $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos e, para todo $i \in I$, $f_i : X \to X_i$ uma função. A *topologia inicial* de X com respeito à família $(f_i)_{i\in I}$ é a menor topologia de X tal que, para todo $i \in I$, f_i é contínua.

⊢ **Proposição 14.12.** *Sejam $\boldsymbol{X}$ um conjunto, $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos e, para todo $i \in I$, $f_i : X \to X_i$ uma função. A topologia inicial de X com respeito à família $(f_i)_{i\in I}$ é a topologia*

$$\left\langle \bigcup_{i\in I} f_i^\star(\mathcal{T}_i) \right\rangle.$$

□ *Demonstração.* Seja $\mathcal{T}$ uma topologia de X tal que, para todo $i \in I$, $f_i : X \to X_i$ é contínua. Então, para todo $i \in I$, $f_i^\star(\mathcal{T}_i) \subseteq \mathcal{T}$ (14.11), o que implica que $\bigcup_{i\in I} f_i^\star(\mathcal{T}_i) \subseteq \mathcal{T}$ e, portanto, $\langle \bigcup_{i\in I} f_i^\star(\mathcal{T}_i) \rangle \subseteq \mathcal{T}$. ■

14.1.4.2 Topologias empurrada e final

:⊢ **Definição 14.13.** Sejam $\boldsymbol{X} = (X, \mathcal{T}_X)$ um espaço topológico, Y um conjunto e $f : X \to Y$ uma função. A *topologia empurrada* por f de $\boldsymbol{X}$ para Y é

$$f_\star(\mathcal{T}_X) := \left\{ A \subseteq Y \;\middle|\; f^{-1}(A) \in \mathcal{T}_X \right\}.$$

⊢ **Proposição 14.13.** *Sejam $\boldsymbol{X} = (X, \mathcal{T}_X)$ um espaço topológico, Y um conjunto e $f : X \to Y$ uma função. Então $\mathcal{T}_Y := f_\star(\mathcal{T}_X)$, a topologia empurrada por f de $\boldsymbol{X}$ para Y, é uma topologia de Y.*

□ *Demonstração.* (1) Como $f^{-1}(\emptyset) = \emptyset$ e $f^{-1}(Y) = X$, segue que $\emptyset, Y \in f_\star(\mathcal{T}_X)$. (2) Seja $(A_i)_{i\in I}$ uma família de conjuntos de $f_\star(\mathcal{T}_X)$. Então, para cada $i \in I$, $f^{-1}(A_i) \in \mathcal{T}_X$, o que implica que

$$f^{-1}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f^{-1}(A_i) \in \mathcal{T}_X,$$

portanto $\bigcup_{i\in I} A_i \in f_\star(\mathcal{T}_X)$. (3) Seja $(A_i)_{i=1}^n$ uma família de conjuntos de $f_\star(\mathcal{T}_X)$. Então, para cada $1 \leq i \leq n$, $f^{-1}(A_i) \in \mathcal{T}_X$, o que implica que

$$f^{-1}\left(\bigcap_{i=1}^n A_i\right) = \bigcap_{i=1}^n f^{-1}(A_i) \in \mathcal{T}_X,$$

portanto $\bigcap_{i=1}^n A_i \in f_\star(\mathcal{T}_X)$. ■

:⊢ **Definição 14.14.** Sejam $\boldsymbol{X}$ um conjunto, $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos e, para todo $i \in I$, $f_i : X_i \to X$ uma função. A *topologia final* de X com respeito à família $(f_i)_{i\in I}$ é a maior topologia de X tal que, para todo $i \in I$, f_i é contínua.

⊢ **Proposição 14.14.** *Sejam $\boldsymbol{X}$ um conjunto, $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos e, para todo $i \in I$, $f_i : X_i \to X$ uma função. A topologia final de X com respeito à família $(f_i)_{i\in I}$ é a topologia*

$$\bigcap_{i\in I} f_{i\star}(\mathcal{T}_i).$$

□ *Demonstração.* Seja $\mathcal{T}$ uma topologia de X tal que, para todo $i \in I$, $f_i : X_i \to X$ é contínua, e seja $A \in \mathcal{T}$. Então, para todo $i \in I$, $f_i^{-1}(A) \in \mathcal{T}_i$, o que implica que $A \in f_{i\star}(\mathcal{T}_i)$, portanto $A \in \bigcap_{i\in I} f_{i\star}(\mathcal{T}_i)$. Isso mostra que $\mathcal{T} \subseteq \bigcap_{i\in I} f_{i\star}(\mathcal{T}_i)$, e como interseção de topologias é topologia, segue o procurado. ■

14.1.4.3 Produto de espaços topológicos

:⊢ **Definição 14.15.** Seja $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos. O *produto* da família $(\boldsymbol{X_i})_{i\in I}$ é o par

$$\prod_{i\in I} \boldsymbol{X_i} := \left(\prod_{i\in I} X_i, \left\langle \bigcup_{i\in I} \pi_i{}^\star(\mathcal{T}_i) \right\rangle\right).$$

A topologia $\langle \bigcup_{i\in I} \pi_i{}^\star(\mathcal{T}_i)\rangle$ é a *topologia produto* de $\prod_{i\in I} X_i$.

⊢ **Proposição 14.15.** *Sejam $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos e $X = \prod_{i\in I} X_i$ o produto de conjuntos. A topologia produto de X é a topologia gerada pela base cujos elementos são $\prod_{i\in I} A_i$, tal que $A_i \in \mathcal{T}_i$ e existe $J \subseteq I$ finito com $A_i = X_i$ para $i \in I \smallsetminus J$.*

□ *Demonstração.* Como $\pi_i = \pi_i \circ \mathrm{I}_X$, segue de uma propriedades básica de imagem inversa de produto (2.8) que

$$\prod_{i\in I} A_i = \mathrm{I}_X^{-1}\left(\prod_{i\in I} A_i\right) = \bigcap_{i\in I} \pi_i^{-1}(A_i)$$

Para todo $i \in I \smallsetminus J$, $A_i = X_i$, então $\pi_i^{-1}(A_i) = X$ e, portanto,

$$\bigcap_{i\in I\smallsetminus J} \pi_i^{-1}(A_i) = \bigcap_{i\in I\smallsetminus J} X = X.$$

Isso implica que

$$\bigcap_{i\in I} \pi_i^{-1}(A_i) = \left(\bigcap_{i\in I\smallsetminus J} \pi_i^{-1}(A_i)\right) \cap \left(\bigcap_{j\in J} \pi_j^{-1}(A_j)\right) = \bigcap_{j\in J} \pi_j^{-1}(A_j)$$

Logo $\prod_{i\in I} A_i = \bigcap_{j\in J} \pi_j^{-1}(A_j)$. Seja A aberto da topologia descrita na proposição. Então

$$A = \bigcup_{k\in K} A_k = \bigcup_{k\in K} \prod_{i\in I} A_{ki} = \bigcup_{k\in K} \bigcap_{j\in J} \pi_j^{-1}(A_{kj}),$$

o que mostra que A é aberto da topologia produto. O resto da demonstração é simples. ■

⊢ **Proposição 14.16** (Propriedade Universal)**.** *Sejam $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos, $\boldsymbol{T} = (T, \mathcal{T}_T)$ um espaço topológico e, para todo $i \in I$, $f_i : \boldsymbol{T} \to \boldsymbol{X_i}$ uma função contínua. Então existe uma única função contínua $f : \boldsymbol{T} \to \prod_{i\in I} \boldsymbol{X_i}$ tal que, para todo $i \in I$, $\pi_i \circ f = f_i$ (o diagrama comuta).*

$$\begin{tikzcd}
 & \prod_{i\in I} \boldsymbol{X_i} \arrow[d, "\pi_i"] \\
\boldsymbol{T} \arrow[ur, dashed, "f"] \arrow[r, "f_i"'] & \boldsymbol{X_i}
\end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} f\colon T &\longrightarrow \prod_{i\in I} X_i \\ x &\longmapsto (f_i(x))_{i\in I}. \end{aligned}$$

Pela propriedade universal do produto de conjuntos, f é única e $\pi_i \circ f = f_i$. Resta mostrar que f é contínua. Seja $A \in \mathcal{T}_X$. Então $A = \bigcup_{k\in K} A_k$ é uma união de abertos básicos $A_k \in \mathcal{T}$. Isso significa que, para todo $k \in K$, $A_k = \prod_{i\in I} A_{ki}$, com $A_{ki} \in \mathcal{T}_i$ para todo $i \in I$ e existe $J_k \subseteq I$ finito tal que, para todo $i \in I \smallsetminus J_k$, $A_{ki} = X_i$. Assim, por propriedades básicas de imagem inversa de união e produto (2.8),

$$f^{-1}(A) = f^{-1}\left(\bigcup_{k\in K}\prod_{i\in I} A_{ki}\right) = \bigcup_{k\in K} f^{-1}\left(\prod_{i\in I} A_{ki}\right) = \bigcup_{k\in K}\bigcap_{i\in I} f_i^{-1}(A_{ki}).$$

Seja $k \in K$. Como, para todo $i \in I \smallsetminus J_k$, $A_{ki} = X_i$, então $f_i^{-1}(A_{ki}) = f_i^{-1}(X_i) = T$. Disso segue que

$$\bigcap_{i\in I} f_i^{-1}(A_{ki}) = \bigcap_{j\in J_k} f_j^{-1}(A_{kj})$$

e, portanto,

$$f^{-1}(A) = \bigcup_{k\in K}\bigcap_{j\in J_k} f_j^{-1}(A_{kj}).$$

Seja $k \in K$. Para todo $j \in J_k$, f_j é contínua, o que implica que $f_j^{-1}(A_{kj})$ é aberto e, por J_k ser finito, a interseção $\bigcap_{j\in J_k} f_j^{-1}(A_{kj})$ é aberta. Isso significa que a união $\bigcup_{k\in K}\bigcap_{j\in J_k} f_j^{-1}(A_{kj})$ é aberta e, portanto, $f^{-1}(A) \in \mathcal{T}_T$. Logo f é contínua. ■

14.1.4.4 Coproduto de espaços topológicos

:⊢ **Definição 14.16.** Seja $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ uma família de espaços topológicos. O *coproduto* da família $(\boldsymbol{X_i})_{i\in I}$ é o par

$$\bigsqcup_{i\in I} \boldsymbol{X_i} := \left(\bigsqcup_{i\in I} X_i, \bigcap_{i\in I} \iota_{i\star}(\mathcal{T}_i)\right).$$

A topologia $\bigcap_{i\in I} \pi_{i\star}(\mathcal{T}_i)$ é a *topologia coproduto* de $\bigsqcup_{i\in I} X_i$.

⊢ **Proposição 14.17** (Propriedade Universal)**.** *Sejam* $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{T}_i)_{i\in I}$ *uma família de espaços topológicos,* $\boldsymbol{T} = (T, \mathcal{T}_T)$ *um espaço topológico e, para todo* $i \in I$*,* $f_i : \boldsymbol{X_i} \to \boldsymbol{T}$ *uma função contínua. Então existe uma única função contínua* $f : \bigsqcup_{i\in I} \boldsymbol{X_i} \to \boldsymbol{T}$ *tal que, para todo* $i \in I$*,* $f \circ \iota_i = f_i$ *(o diagrama comuta).*

$$\begin{tikzcd}
\bigsqcup_{i\in I} \boldsymbol{X_i} \arrow[dr, dashed, "f"] & \\
\boldsymbol{X_i} \arrow[u, "\iota_i"] \arrow[r, "f_i"'] & \boldsymbol{T}
\end{tikzcd}$$

14.1.4.5 Subespaços topológicos

:⊢ **Definição 14.17.** Sejam $(X, \mathcal{T})$ um espaço topológico e $S \subseteq X$ um subconjunto. A *topologia induzida* por $\mathcal{T}$ em S é o conjunto

$$\mathcal{T}|_S := \{A \cap S \mid A \in \mathcal{T}\}.$$

⊢ **Proposição 14.18.** *Sejam* $(X, \mathcal{T})$ *um espaço topológico e* $S \subseteq X$ *um subconjunto. A topologia induzida por* $\mathcal{T}$ *em* S *é o conjunto*

$$\mathcal{T}|_S = \iota_\star(\mathcal{T}),$$

a maior topologia tal que a inclusão $\iota : S \to X$ *é contínua.*

⊢ **Proposição 14.19.** *Sejam* $(X, \mathcal{T})$ *um espaço topológico e* $S \subseteq X$ *um subconjunto. A topologia induzida por* $\mathcal{T}$ *em* S *é uma topologia de* S*.*

□ *Demonstração.* (1) Notemos que $\emptyset, S \in \mathcal{T}|_S$, pois $\emptyset \cap S = \emptyset$ e $X \cap S = S$. (2) Seja $(A_i)_{i\in I}$ uma família de abertos de $\mathcal{T}|_S$. Então, para cada $i \in I$, existe um aberto $B_i \in \mathcal{T}$ tal que $A_i = B_i \cap S$. Assim, temos que

$$\bigcup_{i\in I} A_i = \bigcup_{i\in I} (B_i \cap S) = \left(\bigcup_{i\in I} B_i\right) \cap S,$$

que pertence à topologia induzida pois $\bigcup_{i\in I} B_i \in \mathcal{T}$. (3) Seja $(A_i)_{i=1}^n$ uma família de abertos de $\mathcal{T}|_S$. Então, para cada $1 leq i leq n$, existe um aberto $B_i \in \mathcal{T}$ tal que $A_i = B_i \cap S$. Assim, temos que

$$\bigcap_{i=1}^n A_i = \bigcap_{i=1}^n (B_i \cap S) = \left(\bigcap_{i=1}^n B_i\right) \cap S,$$

que pertence à topologia induzida pois $\bigcap_{i=1}^n B_i \in \mathcal{T}$. ■

⊢ **Proposição 14.20** (Propriedade característica)**.** *Sejam* $\boldsymbol{X} = (X, \mathcal{T}_X)$ *e* $(Y, \mathcal{T}_Y)$ *espaços topológicos e* $\boldsymbol{S} \subseteq \boldsymbol{X}$ *um subespaço. Uma função* $f\colon Y \to S$ *é contínua se, e somente se,* $\iota \circ f : Y \to X$ *é contínua (o diagrama comuta).*

$$\begin{tikzcd} & \boldsymbol{X} \\ \boldsymbol{T} \arrow[ur, "\iota \circ f"] \arrow[r, "f"'] & \boldsymbol{S} \arrow[u, "\iota"'] \end{tikzcd}$$

Proposição Restrição de função contínua é contínua na topologia induzida.

⊢ **Proposição 14.21** (Colagem por abertos)**.** *Sejam* $\boldsymbol{X}$ *e* $\boldsymbol{Y}$ *espaços topológicos,* $f : X \to Y$ *uma função e* $(X_i)_{i \in I}$ *uma cobertura de* X *por conjuntos abertos tal que, para todo* $i \in I$*,* $f|_{X_i} : X_i \to Y$ *é contínua. Então* $f : X \to Y$ *é contínua.*

⊢ **Proposição 14.22** (Colagem por fechados)**.** *Sejam* $\boldsymbol{X}$ *e* $\boldsymbol{Y}$ *espaços topológicos,* $f : X \to Y$ *uma função e* $(X_i)_{i=1}^n$ *uma cobertura de* X *por conjuntos fechados tal que, para todo* $1 \leq i \leq n$*,* $f|_{X_i} : X_i \to Y$ *é contínua. Então* $f : X \to Y$ *é contínua.*

14.1.4.6 Quociente de espaços topológicos

:⊢ **Definição 14.18.** Sejam $\boldsymbol{X} = (X, \mathcal{T})$ um espaço topológico e $\sim$ uma relação de equivalência em X. O *espaço quociente* com respeito a $\sim$ é o espaço topológico

$$\boldsymbol{X}\big/\!\sim \;:= \left(X\big/\!\sim, \pi^\star(\mathcal{T})\right),$$

em que $\pi : X \to X/\!\sim$ é a projeção canônica de equivalências.

É fácil notar que os abertos de $\boldsymbol{X}/\!\sim$ são conjuntos de classes de equivalência cuja união é um aberto de $\boldsymbol{X}$. Notemos, ainda, que se $f : X \to Y$ é sobrejetivo, existe uma relação de equivalência em X induzida por f, definida como dois elementos são equivalentes se suas imagens são iguais, e essa relação de equivalência faz com que possamos identificar Y com $X/\!\sim$ como conjuntos. Definimos a topologia em Y de modo que Y e $X/\!\sim$ sejam homeomorfos.

⊢ **Proposição 14.23** (Propriedade característica)**.** *Sejam* $\boldsymbol{X} = (X, \mathcal{T}_X)$ *e* $(Y, \mathcal{T}_Y)$ *espaços topológicos e* $\boldsymbol{Q}$ *um espaço quociente de* $\boldsymbol{X}$*. Uma função* $f : Q \to Y$ *é contínua se, e somente se,* $f \circ \pi : X \to Y$ *é contínua (o diagrama comuta).*

$$\begin{tikzcd}
\boldsymbol{X} \arrow[d, "\pi"'] \arrow[dr, "f \circ \pi"] & \\
\boldsymbol{Q} \arrow[r, "f"'] & \boldsymbol{Y}
\end{tikzcd}$$

Soma pontual

:⊢ **Definição 14.19.** Sejam $(\boldsymbol{X}_0, x_0)$ e $(\boldsymbol{X}_1, x_1)$ espaços topológicos pontuados. A *soma pontual* de $\boldsymbol{X}_0$ e $\boldsymbol{X}_1$ é o espaço quociente

$$\boldsymbol{X}_0 \dot{+} \boldsymbol{X}_1 := \boldsymbol{X}_0 \sqcup \boldsymbol{X}_1 \big/ \sim,$$

em que $\sim$ é a menor equivalência tal que $(0, x_0) \sim (1, x_1)$.

:⊢ **Definição 14.20.** Sejam $(\boldsymbol{X}_i, x_i)_{i \in I}$ espaços topológicos pontuados. A *soma pontual* de $(\boldsymbol{X}_i, x_i)_{i \in I}$ é o espaço quociente

$$\dot{\sum_{i \in I}} X_i := \sqcup X_i \big/ \sim,$$

em que $\sim$ é a menor equivalência tal que, para todos $i, i' \in I$, $(i, x_i) \sim (i', x_{i'})$.

Exemplo de topologia quociente para soma pontual de esferas Consideremos uma esfera d-dimensional $\mathbb{S}^d$ e o espaço $\mathbb{S}^d \dot{+} \mathbb{S}^d$, dado aqui por

$$\begin{aligned}
\mathbb{S}^d \dot{+} \mathbb{S}^d &= \left(\frac{1}{2}\mathbb{S}^d + \left(\frac{1}{2}, 0, \ldots, 0\right)\right) \cup \left(\frac{1}{2}\mathbb{S}^d - \left(\frac{1}{2}, 0, \ldots, 0\right)\right) \\
&= \left\{x \in \mathbb{R}^{d+1} \;\middle|\; \left\|x - \left(\frac{1}{2}, 0, \ldots, 0\right)\right\| = \frac{1}{2}\right\} \\
&\qquad \cup \left\{x \in \mathbb{R}^{d+1} \;\middle|\; \left\|x + \left(\frac{1}{2}, 0, \ldots, 0\right)\right\| = \frac{1}{2}\right\}.
\end{aligned}$$

Definimos assim a função

$$\begin{aligned}
\mathrm{q} \colon \mathbb{S}^d &\longrightarrow \mathbb{S}^d \dot{+} \mathbb{S}^d \\
(x_0, \ldots, x_d) &\longmapsto \left(x_0, \left(\frac{|x_0| - {x_0}^2}{1 - {x_0}^2}\right)^{\frac{1}{2}} x_1, \ldots, \left(\frac{|x_0| - {x_0}^2}{1 - {x_0}^2}\right)^{\frac{1}{2}} x_d\right)
\end{aligned}$$

Essa função é uma função quociente.

14.1.5 Topologia de ordem

Lembremos que, se $(X, \leq)$ é um conjunto (totalmente) ordenado e $e, e' \in X$, o intervalo aberto de extremos inferior e e superior e' é o conjunto

$$]e, e'[= \{x \in X \mid e < x < e'\}.$$

As semirretas abertas são os intervalos

$$]e, \infty[:= \{x \in X \mid e < x\}$$

e

$$]-\infty, e[:= \{x \in X \mid x < e\}.$$

:⊢ **Definição 14.21.** Seja $(X, \leq)$ um conjunto ordenado. A *topologia de ordem* sobre X é a topologia gerada pelas semirretas abertas de $(X, \leq)$. Denotamos essa topologia por $\mathcal{T}_{\leq}$ (ou simplesmente $\mathcal{T}$ se for claro pelo contexto).

⊢ **Proposição 14.24.** *Seja $(X, \leq)$ um conjunto ordenado. O par $(X, \mathcal{T}_{\leq})$ é um espaço topológico hereditariamente normal com base de intervalos abertos.*

14.1.5.1 Topologia de um corpo ordenado

Lembremos que, se $(X, \leq)$ é um espaço ordenado, a topologia $\mathcal{T}_{\leq}$ induzida por $\leq$ é a topologia gerada pelas semirretas abertas $]x, \infty[$ e $]\infty, x[$, e o conjunto de intervalos abertos $]x, x'[$ é uma base de $\mathcal{T}_{\leq}$.

Note que

$$\bigcup_{i \in I}]e_i, e_i'[= \left]\inf_{i \in I} e_i, \sup_{i \in I} e_i'\right[$$
$$\bigcap_{i \in I} [e_i, e_i'] = \left[\sup_{i \in I} e_i, \inf_{i \in I} e_i'\right].$$

⊢ **Proposição 14.25.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. O par $(\boldsymbol{C}, \mathcal{T}_{\leq})$ é um corpo topológico.*

□ *Demonstração.* Precisamos mostrar que $+$, $-$, $\times$ e $^{-1}$ são contínuas. Para isso, basta mostrar que elas puxam abertos geradores para abertos. Seja $e \in C$ e consideremos as semirretas $]e, \infty[$ e $]-\infty, e[$. Para mostrar que $+$ é contínua, notamos que, se $c + c' > e$, então $c' > e - c$, logo

$$+^{-1}\big(]e, \infty[\big) = \bigcup_{c \in C} \big(]c, \infty[\times]e - c, \infty[\big)$$

e, analogamente,

$$+^{-1}\Big(\,]-\infty, e[\,\Big) = \bigcup_{c\in C}\Big(\,]-\infty, c[\,\times\,]-\infty, e-c[\,\Big),$$

que são abertos pois são uniões de abertos, o que implica que $+$ é contínua. Para mostrar que $-$ é contínua, notamos que

$$-^{-1}\Big(\,]e, \infty[\,\Big) = \,]-\infty, -e[$$

e

$$-^{-1}(]-\infty, e[) = \,]-e, \infty[\,,$$

que são ambos abertos, o que mostra que $-$ é contínua.

Para mostrar que $\times$ é contínua, consideramos dois casos.

(1) Se $e = 0$ e $c \times c' > e = 0$, então $c \neq 0$, logo se $c > 0$ então $c' > 0$ e, se $c < 0$, então $c' < 0$, logo

$$\times^{-1}\Big(\,]0, \infty[\,\Big) = \Big(\,]0, \infty[\,\times\,]0, \infty[\,\Big) \cup \Big(\,]-\infty, 0[\,\times\,]-\infty, 0[\,\Big)$$

e, analogamente,

$$\times^{-1}\Big(\,]-\infty, 0[\,\Big) = \Big(\,]0, \infty[\,\times\,]-\infty, 0[\,\Big) \cup \Big(\,]-\infty, 0[\,\times\,]0, \infty[\,\Big),$$

que são ambos abertos; o caso (2) em que $e \neq 0$ é análogo, embora um pouco mais trabalhoso. Isso mostra que $\times$ é contínua. Mostrar que $^{-1}$ é contínua também é trabalhoso, mas segue direto de modo análogo. ■

⊢ **Proposição 14.26.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. O corpo topológico $(\boldsymbol{C}, \mathcal{T}_{\leq})$ é separado.*

□ *Demonstração.* Sejam $c, c' \in C$ tais que $c \neq c'$. Então $c < c'$ ou $c' < c$. Sem perda de generalidade, considere que $c < c'$. Então as vizinhanças

$$\left]-\infty, \frac{c+c'}{2}\right[\qquad \text{e} \qquad \left]\frac{c+c'}{2}, \infty\right[$$

de c e c', respectivamente, separam c e c'. ■

Consideremos em $C_{\geq 0}$ a topologia induzida.

⊢ **Proposição 14.27.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. As funções valor absoluto $|\cdot| : C \to C_{\geq 0}$ e distância $|\cdot, \cdot| : C \times C \to C_{\geq 0}$ são contínuas.*

□ *Demonstração.* (Valor absoluto) Seja $c \in C_{\geq 0}$ e consideremos os abertos $]c, \infty[$ e $[0, c[$. Então

$$|\cdot|^{-1}\Big(]c, \infty[\Big) =]-\infty, -c[\cup]c, \infty[$$

e

$$|\cdot|^{-1}\Big([0, c[\Big) =]-c, 0] \cup [0, c[=]-c, c[\,,$$

que são ambos abertos, o que mostra que $|\cdot|$ é contínua.

(Distância) Como $|\cdot, \cdot|$ é composição de $|\cdot|$, $+$, $-$, p_0 e p_1, que são todas contínuas, segue que ela é contínua. ■

⊢ **Proposição 14.28.** *Seja $(\boldsymbol{C}, \leq)$ um corpo ordenado. O corpo $\boldsymbol{C}$ é não-infinitesimal se, e somente se, $\mathbb{Q}_C$ é denso em C.*

14.2 Separação

14.2.1 Noções de separação de conjuntos

Nesta seção são apresentadas algumas noções de como dois conjuntos de um espaço topológico podem ser separados. As duas noções mas simples de separação são noções conjuntistas. A primeira é a de conjuntos distintos, ou diferentes, $A \neq B$, e a outra é a de conjuntos disjuntos, $A \cap B = \emptyset$. A seguir, mostramos noções que envolvem construções topológicas e não meramente conjuntistas.

:⊢ **Definição 14.22.** Sejam $\boldsymbol{X}$ um espaço topológico e $A, B \subseteq X$. Definimos as seguintes relações entre A e B:

1. (*Separação*) Cada conjunto é disjunto do fecho do outro.

$$A \cap \overline{B} = \overline{A} \cap B = \emptyset.$$

2. (*Separação por vizinhanças*) Existem vizinhanças $V_A \in \mathcal{V}_A$ e $V_B \in \mathcal{V}_B$ que são disjuntas.

$$V_A \cap V_B = \emptyset.$$

3. (*Separação por função contínua*) Existe $f \in \mathscr{C}(X, [0, 1])$ tal que

$$f(A) = \{0\} \;\text{ e }\; f(B) = \{1\}.$$

4. (*Separação precisa por função contínua*) Existe $f \in \mathscr{C}(X, [0, 1])$ tal que

$$f^{-1}(\{0\}) = A \;\text{ e }\; f^{-1}(\{1\}) = B.$$

Cada uma das relações vale entre pontos $x, y \in X$, ou entre um conjunto A e um ponto x, ao considerarmos no lugar no ponto o conjunto unitário que o contém: valem entre os conjuntos $\{x\}$ e $\{y\}$, ou entre A e $\{x\}$, respectivamente.

As primeiras duas relações binárias são claramente simétricas, mas isso não é necessariamente claro no caso da terceira e quarta. No entanto, isso pode ser concluído ao considerar, dada uma função f que faz A e B separados por função contínua, a função $1 - f$ que faz o mesmo entre B e A.

$\vdash$ **Proposição 14.29.** *Sejam $\boldsymbol{X}$ um espaço topológico e $A, B \subseteq X$. Então*

1. *Se A e B são precisamente separados por função contínua, então são separados por função contínua.*
2. *Se A e B são separados por função contínua, então são separados por vizinhanças.*
3. *Se A e B são separados por vizinhanças, então são separados.*
4. *Se A e B são separados, então são disjuntos.*

14.2.2 Espaços distinguíveis

$:\vdash$ **Definição 14.23.** Seja $\boldsymbol{X}$ um espaço topológico. Pontos *topologicamente indistinguíveis* em $\boldsymbol{X}$ são pontos $x, y \in X$ tais que $\mathcal{V}_x = \mathcal{V}_y$. Pontos *topologicamente disntinguíveis em* $\boldsymbol{X}$ são pontos que não são topologicamente indistinguíveis.

$\vdash$ **Proposição 14.30.** *Seja $\boldsymbol{X}$ um espaço topológico. A relação binária de indistinguibilidade topológica é uma relação de equivalência em X.*

$\square$ *Demonstração.* Denotemos por $\sim$ a relação binária de indistinguibilidade topológica. Sejam $x, y, z \in X$. Para mostrar a reflexividade, notemos que, como $\mathcal{V}_x = \mathcal{V}_y$, então $x \sim x$. Para mostrar a simetria, se $x \sim y$, então $\mathcal{V}_x = \mathcal{V}_y$, o que é equivalente a $\mathcal{V}_y = \mathcal{V}_x$ e, portanto, $y \sim x$. Por fim, para mostrar a transitividade, se $x \sim y$ e $y \sim z$, então $\mathcal{V}_x = \mathcal{V}_y$ e $\mathcal{V}_y = \mathcal{V}_z$, o que implica $\mathcal{V}_x = \mathcal{V}_z$ e, portanto, $x \sim z$. $\blacksquare$

O fato de que essa é uma relação de equivalência mostra que podemos obter a partir de qualquer espaço topológico um espaço topológico em que nenhum ponto é topologicamente indisntinguível. Para isso, basta considerar o espaço quociente definido pela relação. de indistinguibilidade topológica. Espaços com essa propriedade são definidos a seguir.

$:\vdash$ **Definição 14.24** (T_0)**.** Um espaço topológico *distinguível* é um espaço topológico $\boldsymbol{X}$ em que todo par de pontos distintos é topologicamente distinguível:

$$\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \mathcal{V}_x \neq \mathcal{V}_y.$$

$\vdash$ **Proposição 14.31.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes as seguintes propriedades:*

1. *$\boldsymbol{X}$ é distinguível.*

2. $\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \mathcal{V}_x \nsubseteq \mathcal{V}_y$ *ou* $\mathcal{V}_y \nsubseteq \mathcal{V}_x$.
3. $\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad x \notin \overline{\{y\}}$ *ou* $y \notin \overline{\{x\}}$.
4. $\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \overline{\{x\}} \neq \overline{\{y\}}$.

⊢ **Proposição 14.32.** *Sejam* $\boldsymbol{X}$ *um espaço topológico distinguível e* $Y \subseteq X$. *Então* $\boldsymbol{Y}$ *é um espaço topológico distinguível.*

⊢ **Proposição 14.33.** *Seja* $(X_i)_{i \in I}$ *uma família não vazia de espaços topológicos não vazios. O espaço produto* $\prod_{i \in I} \boldsymbol{X_i}$ *é distinguível se, e somente se, todos os espaços* $\boldsymbol{X}_i$ *são distinguíveis.*

14.2.3 Espaços acessíveis

:⊢ **Definição 14.25** (T_1)**.** Um espaço topológico *acessível* é um espaço topológico $\boldsymbol{X}$ em que

$$\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \mathcal{V}_x \nsubseteq \mathcal{V}_y \text{ e } \mathcal{V}_y \nsubseteq \mathcal{V}_x.$$

⊢ **Proposição 14.34** ($T_1 \Rightarrow T_0$)**.** *Seja* $\boldsymbol{X}$ *um espaço topológico. Se* $\boldsymbol{X}$ *é acessível, então* $\boldsymbol{X}$ *é distinguível.*

□ *Demonstração.* Se $\boldsymbol{X}$ é acessível, então, para todos $x, y \in X$ tais que $x \neq y$, temos que $\mathcal{V}_x \nsubseteq \mathcal{V}_y$ e $\mathcal{V}_y \nsubseteq \mathcal{V}_x$. Mas isso implica que $\mathcal{V}_x \nsubseteq \mathcal{V}_y$ ou $\mathcal{V}_y \nsubseteq \mathcal{V}_x$, o que é equivalente a $\mathcal{V}_x \neq \mathcal{V}_y$. Logo $\boldsymbol{X}$ é distinguível. ■

⊢ **Proposição 14.35.** *Seja* $\boldsymbol{X}$ *um espaço topológico. São equivalentes as seguintes propriedades:*

1. $\boldsymbol{X}$ *é acessível.*
2. *Todo par de pontos distintos é separado:*

$$\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad x \notin \overline{\{y\}} \text{ e } y \notin \overline{\{x\}}.$$

3. $\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \exists U \in \mathcal{V}_x,\ V \in \mathcal{V}_y \qquad y \notin U$ *e* $x \notin V$.
4. $\forall x \in X \quad \overline{\{x\}} = \{x\}$.

⊢ **Proposição 14.36.** *Sejam* $\boldsymbol{X}$ *um espaço topológico acessível e* $Y \subseteq X$. *Então* $\boldsymbol{Y}$ *é um espaço topológico acessível.*

⊢ **Proposição 14.37.** *Seja* $(X_i)_{i \in I}$ *uma família não vazia de espaços topológicos não vazios. O espaço produto* $\prod_{i \in I} \boldsymbol{X_i}$ *é acessível se, e somente se, todos os espaços* $\boldsymbol{X}_i$ *são acessíveis.*

14.2.4 Espaços separados

:⊢ **Definição 14.26** (T_2)**.** Um espaço topológico *separado (por vizinhanças)* é um espaço topológico $\boldsymbol{X}$ em que todo par de pontos distintos é separado por vizinhanças:

$$\forall x, y \in X \quad x \neq y \quad \Rightarrow \quad \exists U \in \mathcal{V}_x,\ V \in \mathcal{V}_y \quad U \cap V = \emptyset.$$

Esses espaços também são conhecidos como espaços de Hausdorff.

⊢ **Proposição 14.38** ($T_2 \Rightarrow T_1$)**.** *Seja $\boldsymbol{X}$ um espaço topológico. Se $\boldsymbol{X}$ é separado, então $\boldsymbol{X}$ é acessível.*

⊢ **Proposição 14.39.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes as seguintes propriedades:*

1. *$\boldsymbol{X}$ é separado.*
2. *Toda rede convergente em $\boldsymbol{X}$ tem limite único.*
3. *Todo filtro convergente em $\boldsymbol{X}$ tem limite único.*

⊢ **Proposição 14.40.** *Sejam $\boldsymbol{X}$ um espaço topológico separado e $Y \subseteq X$. Então $\boldsymbol{Y}$ é um espaço topológico separado.*

⊢ **Proposição 14.41.** *Seja $(\boldsymbol{X}_i)_{i \in I}$ uma família não vazia de espaços topológicos não vazios. O espaço produto $\prod_{i \in I} \boldsymbol{X_i}$ é separado se, e somente se, todos os espaços $\boldsymbol{X}_i$ são separados.*

⊢ **Proposição 14.42.** *Sejam $\boldsymbol{X}$ um espaço topológico, $\boldsymbol{Y}$ um espaço topológico separado, $D \subseteq X$ um subconjunto denso em X e $f, g : X \to Y$ funções contínuas tais que $f|_D = g|_D$. Então $f = g$.*

14.2.5 Espaços regulares

:⊢ **Definição 14.27.** Um espaço topológico *regular* é um espaço topológico em que é possível separar por vizinhanças qualquer ponto de qualquer conjunto fechado que não o contém.

Espaços separados regulares são também chamados de T_3.

⊢ **Proposição 14.43.** *Seja $\boldsymbol{X}$ um espaço topológico regular. Então $\boldsymbol{X}$ é separado por vizinhanças se, e somente se, é distinguível.*

□ *Demonstração.* Sabemos que todo espaço separado por vizinhanças é distinguível. Para demonstrar a recíproca, supondo que $\boldsymbol{X}$ é distinguível, então para todos pontos $x, y \in X$, $x \notin \overline{\{y\}}$ ou $y \notin \overline{\{x\}}$. Sem perda de generalidade, assumamos o primeiro caso. Seja $F := \overline{\{y\}}$. Então F é fechado e $x \notin F$. Da regularidade de $\boldsymbol{X}$, segue que existem $U \in \mathcal{V}_x$ e $V \in \mathcal{V}_F$ tais que $U \cap V = \emptyset$. Como $y \in F$, então $V \in \mathcal{V}_y$, o que implica que $\boldsymbol{X}$ é separado por vizinhanças. ■

$\vdash$ **Proposição 14.44.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes as seguintes propriedades:*

1. *$\boldsymbol{X}$ é regular.*
2. *Para todos ponto $x \in X$ e aberto $A \in \mathcal{V}_x$, existe aberto $V \in \mathcal{V}_x$ tal que*

$$x \in V \subseteq \overline{V} \subseteq A.$$

$\vdash$ **Proposição 14.45.** *Sejam $\boldsymbol{X}$ um espaço topológico regular e $Y \subseteq X$. Então $\boldsymbol{Y}$ é um espaço topológico regular.*

$\vdash$ **Proposição 14.46.** *Seja $(X_i)_{i \in I}$ uma família não vazia de espaços topológicos não vazios. O espaço produto $\prod_{i \in I} \boldsymbol{X_i}$ é regular se, e somente se, todos os espaços $\boldsymbol{X}_i$ são regulares.*

14.2.6 Espaços completamente regulares

$:\vdash$ **Definição 14.28.** Um espaço topológico *completamente regular* é um espaço topológico em que é possível separar por função contínua qualquer ponto de qualquer conjunto fechado que não o contém.

$\vdash$ **Proposição 14.47.** *Seja $\boldsymbol{X}$ um espaço topológico. Se $\boldsymbol{X}$ é completamente regular, então $\boldsymbol{X}$ é regular.*

$\vdash$ **Proposição 14.48.** *Sejam $\boldsymbol{X}$ um espaço topológico completamente regular e $Y \subseteq X$. Então $\boldsymbol{Y}$ é um espaço topológico completamente regular.*

$\vdash$ **Proposição 14.49.** *Seja $(X_i)_{i \in I}$ uma família não vazia de espaços topológicos não vazios. O espaço produto $\prod_{i \in I} \boldsymbol{X_i}$ é completamente regular se, e somente se, todos os espaços $\boldsymbol{X}_i$ são completamente regulares.*

14.2.7 Espaços normais

$:\vdash$ **Definição 14.29.** Um espaço topológico *normal* é um espaço topológico em que todo par de conjuntos fechados disjuntos é separado por vizinhanças.

Espaços normais separados por vizinhanças são também chamados de T_4.

$\vdash$ **Proposição 14.50.** *Seja $\boldsymbol{X}$ um espaço topológico normal. Então $\boldsymbol{X}$ é separado por vizinhanças se, e somente se, é acessível.*

$\square$ *Demonstração.* Sabemos que todo espaço separado por vizinhanças é distinguível. Para demonstrar a recíproca, suponhamos que $\boldsymbol{X}$ é acessível e sejam $x, y \in X$ tais que $x \neq y$. Então vale que $\overline{\{x\}} = \{x\}$ e $\overline{\{y\}} = \{y\}$. Como $x \neq y$, da normalidade de $\boldsymbol{X}$ existem $V_x \in \mathcal{V}_x$ e $V_y \in \mathcal{V}_y$ tais que $V_x \cap V_y = \emptyset$; ou seja, $\boldsymbol{X}$ é separado por vizinhanças. $\blacksquare$

⊢ **Proposição 14.51.** *Sejam $\boldsymbol{X}$ um espaço topológico normal e $Y \subseteq X$ fechado. Então $\boldsymbol{Y}$ é um espaço topológico normal.*

⊢ **Proposição 14.52.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes as seguintes propriedades:*

1. *$\boldsymbol{X}$ é normal.*
2. *Para todos fechado F e aberto $A \in \mathcal{V}_F$, existe aberto $V \in \mathcal{V}_F$ tal que*

$$F \subseteq V \subseteq \overline{V} \subseteq A.$$

⊢ **Proposição 14.53** (Lema de Urysohn)**.** *Um espaço topológico $\boldsymbol{X}$ é normal se, e somente se, todo par de conjuntos fechados disjuntos é separado por função contínua.*

□ *Demonstração.* Se todo par de conjuntos fechados é separado por função contínua, segue da proposição 14.29 que eles são separados por vizinhanças.

Para demonstrar a recíproca, suponhamos que $\boldsymbol{X}$ é normal. Sejam $F_0, F_1 \subseteq X$ conjuntos fechados disjuntos. Seja $Q := \mathbb{Q} \cap [0,1]$. Construiremos uma família de abertos $(A_q)_{q \in Q}$ com a seguinte propriedade:

- Se $p, q \in Q$ são racionais tais que $p < q$, então $\overline{A_p} \subseteq A_q$.

Consideremos uma enumeração de $r : \mathbb{N} \to Q$ tal que $r(0) = 0$ e $r(1) = 1$, de modo que possamos fazer indução nos elementos de Q. Definimos $A_1 := {F_1}^{\complement}$ e notamos que $F_0 \subseteq A_1$, pois F_0 e F_1 são disjuntos. Como $\boldsymbol{X}$ é normal e $F_0 \subseteq A_1$, existe aberto A_0 tal que

$$F_0 \subseteq A_0 \subseteq \overline{A_0} \subseteq A_1.$$

Mostremos por indução que para todo $q \in Q$ existe A_q com a propriedade enunciada. O caso base já está mostrado pela construção de A_0 e A_1, então consideremos o passo indutivo. Seja $Q_n := \{r(k) \mid k \in \{0, \ldots, n\}\}$ e suponha que $\overline{A_p} \subseteq A_q$ para todos racionais $p, q \in Q_n$ tais que $p < q$. Consideremos $r = r(n+1) \in Q$. O conjunto Q_{n+1} é um subconjunto finito de Q e, com essa ordem, todo elemento do conjunto tem um antecessor e um sucessor imediatos. Sejam $p, q \in Q_{n+1}$ tais racionais satisfazendo $p < r < q$. Os conjuntos abertos A_p e A_q já estão definidos pela hipótese de indução, então pela normalidade segue que existe aberto $A_r \subseteq X$ tal que

$$\overline{A_p} \subseteq A_r \subseteq \overline{A_r} \subseteq A_q.$$

Mostremos que a propriedade vale para todos elementos de Q_{n+1}. Se $p, q \in Q_n$, a propriedade vale. Consideremos $r = r(n+1)$ e $s \in Q_n$. Se $s < r(n+1)$, então $s \leq p$, portanto $\overline{A_s} \subseteq \overline{A_p} \subseteq A_r$, e se $r(n+1) < s$, então $q \leq s$, portanto $\overline{A_r} \subseteq A_q \subseteq A_s$.

Assim isso vale para todo Q_n e portanto para Q, e construímos uma família $(A_q)_{q\in Q}$ satisfazendo a propriedade.

Definamos agora a função

$$\begin{aligned} f\colon X &\longrightarrow [0,1] \\ x &\longmapsto \inf\{q \in Q \mid x \in A_q\}. \end{aligned}$$

A função f separa F_0 e F_1: por definição, $F_0 \subseteq A_0$, portanto $f(F_0) = \{0\}$; ainda, por definição, para todo $q \in Q$, tem-se $A_q \subseteq A_1 = F_1^{\complement}$, portanto $f(F_1) = \{1\}$. Para mostrar que f é contínua, notemos antes dois fatos. (1) Para todo $q \in Q$, se $x \in \overline{A_q}$ então $f(x) \leq q$. Isso ocorre porque, se $x \in \overline{A_q}$, então $x \in A_r$ para todo $r \in \mathbb{Q} \cap]q, 1]$, portanto $\{q \in Q \mid x \in A_q\} \subseteq \mathbb{Q} \cap]q, 1]$ o que implica $f(x) \leq q$ por definição de f. (2) Para todo $q \in Q$, se $x \notin A_q$ então $f(x) \geq q$. Isso ocorre porque, se $x \notin A_q$, então $x \notin A_r$ para todo $r \in \mathbb{Q} \cap [0, q[$, portanto $\{q \in Q \mid x \in A_q\}\, \mathbb{Q} \cap [0, q[$ o que implica $f(x) \geq q$ por definição de f.

Agora que provamos esses fatos, seja $x \in]a, b[\subseteq [0, 1]$. Mostremos que existe vizinhança $A \subseteq X$ de x tal que $f(A) \subseteq]a, b[$. Para isso, sejam $p, q \in Q$ tais que $a < p < f(x) < q < b$ e definamos $A := A_q \smallsetminus \overline{A_p}$. Então, pelos fatos acima, $x \in A_q$ porque $f(x) < q$ e $x \notin \overline{A_p}$ porque $f(x) > p$, o que mostra que $x \in A$. Por fim, seja $x' \in A$. Então $x' \in A_q \subseteq \overline{A_q}$, portanto $f(x') \leq q$ e $x' \notin \overline{A_p} \supseteq A_p$, portanto $f(x') \geq p$, o que implica $f(x) \subseteq [p, q] \subseteq]a, b[$. Isso mostra que f é contínua. ■

14.3 Convergência

14.3.1 Redes

⊢ **Proposição 14.54.** *Sejam* $\boldsymbol{X}$ *um espaço topológico e* $C \subseteq X$*. Então* $(\mathcal{V}_C, \supseteq)$*, o conjunto de vizinhanças de* C *com ordem de contenção invertida, é um conjunto direcionado.*

□ *Demonstração.* Sabemos que $\supseteq$ é uma ordem parcial. Sejam $V_0, V_1 \in \mathcal{V}_C$. Então $V := V_0 \cap V_1$ é uma vizinhança de C tal que $V_0 \supseteq V$ e $V_1 \supseteq V$. ■

:⊢ **Definição 14.30.** Sejam X um conjunto e $(\Lambda, \leq)$ um conjunto direcionado. Uma *rede* de elementos de X indexados por Λ é uma função $x : \Lambda \to X$. O conjunto Λ é o *conjunto de índices* da rede. Denota-se $(x_\lambda)_{\lambda\in\Lambda}$ e a imagem de $\lambda \in \Lambda$ por x é denotada x_λ e chamada de λ-*ésimo membro* da rede.

Note que uma sequência é uma rede cujo conjunto direcionado é $(\mathbb{N}, \leq)$.

:⊢ **Definição 14.31** (Limite e convergência)**.** Sejam $\boldsymbol{X}$ um espaço topológico e $(x_\lambda)_{\lambda\in\Lambda}$ uma rede de X. Um *limite* de $(x_\lambda)_{\lambda\in\Lambda}$ em $\boldsymbol{X}$ é um ponto $\ell \in X$ que satisfaz: para toda vizinhança V de ℓ, existe $\lambda \in \Lambda$ tal que, para todo $\mu \geq \lambda$, $x_\mu \in V$. Denota-se $(x_\lambda)_{\lambda\in\Lambda} \to \ell$. Uma rede *convergente* é uma rede que tem limite.

$\vdash$ **Proposição 14.55.** *Um espaço topológico $\boldsymbol{X}$ é separado por vizinhanças (T_2) se, e somente se, toda rede convergente $(x_\lambda)_{\lambda\in\Lambda}$ de X tem limite único.*

$\square$ *Demonstração.* ($\Rightarrow$) Sejam ℓ_0 e ℓ_1 limites de $(x_\lambda)_{\lambda\in\Lambda}$. Se ℓ_0 e ℓ_1 são distintos, então por $\boldsymbol{X}$ ser T_2 existem vizinhanças disjuntas V_0 de ℓ_0 e V_1 de ℓ_1. Da convergência da rede, existem λ_0 e $\lambda_1 \in \Lambda$ tais que, para todo $\mu \geq \lambda_0$, $x_\mu \in V_0$ e, para todo $\mu \geq \lambda_1$, $x_\mu \in V_1$. Como $(\Lambda, \leq)$ é direcionado, existe $\lambda \in \Lambda$ tal que $\lambda_0 \leq \lambda$ e $\lambda_1 \leq \lambda$, o que implica pela convergência da rede que $x_\lambda \in V_0 \cap V_1$, que é uma contradição. Portanto $\ell_0 = \ell_1$.

($\Leftarrow$) Suponhamos que $\boldsymbol{X}$ não é T_2. Então existem pontos distintos x_0 e $x_1 \in X$ tais que, para todas vizinhanças V_0 de x_0 e V_1 de x_1, $V_0 \cap V_1 \neq \emptyset$. Definamos $\mathcal{V} := \mathcal{V}_{\{x_0,x_1\}}$ e notemos que $(\mathcal{V}_{\{x_0,x_1\}}, \supseteq)$ é um conjunto direcionado. Para todos vizinhanças $V_0 \in \mathcal{V}_{x_0}$ e $V_1 \in \mathcal{V}_{x_1}$, tomamos $x_{V_0\cap V_1} \in V_0 \cap V_1$, que existe pois o conjunto $V_0 \cap V_1$ não é vazio. Mostraremos que a rede $(x_V)_{V\in\mathcal{V}}$ então converge para x_0 e para x_1. Sejam V_0 uma vizinhança de x_0 e V_1 uma vizinhança de x_1 e defina $V := V_0 \cap V_1$. Então, para toda vizinhança de $U \in \mathcal{V}$ tal que $U \subseteq V$, segue que $x_U \in U \subseteq V$, portanto a rede $(x_V)_{V\in\mathcal{V}}$ converge para x_0 e para x_1. ■

Usamos o axioma da escolha para construir a rede na volta da demonstração, pois a rede é elemento do produto de todas as vizinhanças de $\mathcal{V}$.

$\vdash$ **Proposição 14.56.** *Sejam $\boldsymbol{X}_0$ e $\boldsymbol{X}_1$ espaços topológicos e $f : X_0 \to X_1$ uma função. Então f é contínua se, e somente se, para todo $\ell \in X_0$ e toda rede $(x_\lambda)_{\lambda\in\Lambda}$ que converge para $\ell \in X_0$, a rede $(f(x_\lambda))_{\lambda\in\Lambda}$ converge para $f(\ell) \in X_1$.*

$\vdash$ **Proposição 14.57.** *Sejam $(\boldsymbol{X}_i)_{i\in I}$ uma família de espaços topológicos e $(x_\lambda)_{\lambda\in\Lambda}$ uma rede de $\boldsymbol{X} = \prod_{i\in I} \boldsymbol{X}_i$. Então $(x_\lambda)_{\lambda\in\Lambda} \to \ell \in X$ se, e somente se, para todo $i \in I$, $(\pi_i(x_\lambda))_{\lambda\in\Lambda} \to \pi_i(\ell) \in X_i$.*

$\square$ *Demonstração.*

($\Rightarrow$) Segue da continuidade de π_i.

($\Leftarrow$) Seja V uma vizinhança de $(\ell_i)_{i\in I}$. Então existe um aberto básico $A = \bigcap_{i\in J} {\pi_i}^{-1}(A_i)$ tal que $A \subseteq V$, $J \subseteq I$ é um conjunto finito e $A_i \subseteq X_i$ é um aberto. Sendo assim, para cada $i \in J$, existe $\lambda_i \in \Lambda$ tal que, para todo $\mu \geq \lambda_i$, $x_{i,\mu} \in A_i$. Portanto, como Λ é um conjunto direcionado (e J é finito), existe λ tal que, para todo $i \in J$, $\lambda_i \leq \lambda$, o que implica que, para todo $\mu \geq \lambda$, $(x_{i,\mu})_{i\in I} \in A \subseteq V$. Isso mostra que $((x_{i,\lambda})_{i\in I})_{\lambda\in\Lambda} \to (\ell_i)_{i\in I}$. ■

14.4 Conexidade

14.4.1 Conexidades

Definimos o conceito de desconexo antes por ele ser mais intuitivo de ser enunciado. Ser desconexo é ter como separar o espaço $\boldsymbol{X}$ em dois conjuntos disjuntos, aberto

e não triviais (não são $\emptyset$ nem X). Esses abertos cobrem todo espaço e, como são abertos, o separam pela definição de separação de conjuntos.

:⊢ **Definição 14.32.** Um espaço topológico *desconexo* é um espaço topológico $\boldsymbol{X}$ tal que X admite partição por dois conjuntos abertos. Um espaço topológico *conexo* é um espaço topológico que não é desconexo. Um subconjunto de X é conexo ou desconexo de acordo com sua topologia induzida.

Uma partição por dois abertos é equivalente a uma cobertura por dois conjuntos abertos, disjuntos e não triviais. Notemos que, por essa definição, o espaço topológico $\boldsymbol{\emptyset}$ é conexo, pois todos subconjuntos de $\emptyset$ são triviais, logo $\emptyset$ não admite partições. A conexidade de $\emptyset$ não é consenso entre autores, mas adotaremos esse resultado aqui. É importante notar que, na definição, poderíamos escolher uma partição por conjuntos fechados, já que, como cada conjunto da partição é o complementar do outro, ambos são fechados, pois ambos são abertos. A definição de separação de conjuntos vale nesse caso, os conjuntos da partição são separados no sentido que cada um é disjunto do fecho do outro, já que são fechados e disjuntos. Isso sugere que conexidade está relacionada com o problema de confusão semântica clássica em topologia: nem todo conjunto aberto é necessariamente não fechado, e vice-versa. Claro que $\emptyset$ e X são sempre abertos e fechados, mas em espaços conexos eles são os únicos. Com o conceito de conexidade, temos o seguinte resultado.

⊢ **Proposição 14.58.** *Um espaço topológico $\boldsymbol{X}$ é conexo se, e somente se, não existe um conjunto não trivial que é aberto e fechado.*

□ *Demonstração.* Vamos mostrar a ida e a volta pela contrapositiva. Se $\boldsymbol{X}$ é desconexo, os conjuntos que formam sua partição por abertos são ambos abertos e fechados. Reciprocamente, se $A \subseteq X$ é não trivial, aberto e fechado, $A^{\complement}$ também é não trivial e aberto (e fechado), logo $\{A, A^{\complement}\}$ é uma partição de X por dois conjuntos abertos. ■

As noções de conexidade de um espaço topológico e de um subconjunto de um espaço topológico são, de fato, a mesma, mas alguns cuidados devem ser tomados para escolher onde os conjuntos são abertos ou fechados, se na topologia original ou na induzida. A proposição a seguir oferece um critério para conjuntos conexos.

⊢ **Proposição 14.59.** *Seja $\boldsymbol{X}$ um espaço topológico. Então $S \subseteq X$ é conexo se, e somente se, não existem conjuntos não triviais separados em $\boldsymbol{X}$ cuja união é S.*

□ *Demonstração.* Se S é desconexo, existe uma partição $\{A, B\}$ de S por abertos de $\boldsymbol{S}$. Então $A \cup B = S$,

$$A \cap B^{\bullet_X} = (A \cap S) \cap B^{\bullet_X} = A \cap (S \cap B^{\bullet_X}) = A \cap B^{\bullet_S} = A \cap B = \emptyset$$

e, similarmente, $A^{\bullet_X} \cap B = \emptyset$, o que mostra que A e B são separados em $\boldsymbol{X}$. Reciprocamente, se existem $A, B \subseteq X$ conjuntos não triviais separados em $\boldsymbol{X}$ tais que $A \cup B = S$, então

$$A^{\bullet_S} = S \cap A^{\bullet_X} = (A \cap B) \cap A^{\bullet_X} = (A \cap A^{\bullet_X}) \cup (B \cap A^{\bullet_X}) = A \cup \emptyset = A$$

e, similarmente, $B^{\bullet_S} = B$, portanto $\{A, B\}$ é partição de S por fechados de $\boldsymbol{S}$, o que mostra que $\boldsymbol{S}$ é desconexo. ■

A seguir, enunciamos uma equivalência da definição de conexidade e um resultado trivial, mas importantíssimo na topologia: conexidade é um invariante topológico.

⊢ **Proposição 14.60.** *Sejam $\boldsymbol{X}$ um espaço topológico, e $\boldsymbol{\{0,1\}}$ espaço topológico com a topologia discreta. Então $\boldsymbol{X}$ é conexo se, e somente se, toda função contínua $f : \boldsymbol{X} \to \boldsymbol{\{0,1\}}$ é constante.*

□ *Demonstração.* Demonstraremos a ida e a volta pela contrapositiva. Se $\boldsymbol{X}$ é desconexo, seja $\{A_0, A_1\}$ uma partição de X por dois conjuntos abertos. Definamos

$$\begin{aligned} f\colon X &\longrightarrow \{0,1\} \\ x &\longmapsto \begin{cases} 0, & x \in A_0 \\ 1, & x \in A_1. \end{cases} \end{aligned}$$

Então $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\{0\}) = A_0$, $f^{-1}(\{1\}) = A_1$ e $f^{-1}(\{0,1\}) = X$. Portanto f é contínua, mas não é constante. Reciprocamente, seja $f : X \to \{0,1\}$ uma função contínua não constante. Então $f^{-1}(\{0\})$ e $f^{-1}(\{1\})$ são abertos, pois f é contínua, e formam uma partição de X: (1) não são vazios, pois f não é constante, (2) são disjuntos e (3) sua união é X. Logo $\boldsymbol{X}$ é desconexo. ■

⊢ **Proposição 14.61.** *Sejam $\boldsymbol{X}$ e $\boldsymbol{Y}$ espaços topológicos e $f : \boldsymbol{X} \to \boldsymbol{Y}$ uma função contínua. Se $\boldsymbol{X}$ é conexo, então $\boldsymbol{Y}$ é conexo.*

□ *Demonstração.* Se $\boldsymbol{Y}$ fosse desconexo, existiria uma partição de Y por abertos $\{A, B\}$, e como f é contínua, $\{f^{-1}(A), f^{-1}(B)\}$ seria uma partição de X por abertos, contradizendo a conexidade de $\boldsymbol{X}$. ■

A seguir, definimos uma função localmente constante e provamos um resultado simples, mas útil.

:⊢ **Definição 14.33.** Sejam $\boldsymbol{X}$ um espaço topológico e C um conjunto. Uma função *localmente constante* de X para C é uma função $f\colon X \longrightarrow C$ tal que, para todo $x \in X$, alguma vizinhança V de x satisfaz que $f|_V$ é constante.

⊢ **Proposição 14.62.** *Sejam $\boldsymbol{X}$ um espaço topológico conexo não vazio e $\boldsymbol{X}'$ um espaço topológico acessível. Se $f\colon X \longrightarrow X'$ é uma função contínua localmente constante, então f é constante.*

□ *Demonstração.* Como X é não vazio, $f(X)$ é não vazio. Seja $c \in f(X)$. Como X' é acessível, $\{c\}$ é fechado; como f é contínua, $f^{-1}(c)$ é fechado. Como f é localmente constante, para todo $x \in f^{-1}(c)$, alguma vizinhança aberta V de x satisfaz $f|_V$ constante, ou seja, $V \subseteq f^{-1}(c)$, o que mostra que $f^{-1}(c)$ é aberto. Assim, segue de X é conexo e, por definição $f^{-1}(c) \neq \emptyset$, concluímos que $f^{-1}(c) = X$, logo f é constante. ■

14.4.1.1 Componentes conexas

⊢ **Proposição 14.63.** *Seja $\boldsymbol{X}$ um espaço topológico e $(C_i)_{i\in I}$ uma família de conjuntos conexos tais que $\bigcap_{i\in I} C_i \neq \emptyset$. Então $\bigcup_{i\in I} C_i$ é conexo.*

□ *Demonstração.* Se $\bigcup_{i\in I} C_i$ for desconexo, existe partição $\{A, B\}$ por abertos de $\bigcup_{i\in I} C_i$. Como $p \in \bigcap_{i\in I} C_i \neq \emptyset$, existe $p \in \bigcap_{i\in I} C_i$ e, portanto, sem perda de generalidade, suponhamos $p \in A$. Seja $q \in B$. Então existe $i \in I$ tal que $q \in C_i$ e, como $p \in C_i$, temos que $p \in A \cap C_i$ e $q \in B \cap C_i$. Então $A \cap C_i$ e $B \cap C_i$ são não vazios e são abertos de C_i, o que implica que eles são uma partição de C_i por conjuntos abertos, e isso contradiz sua conexidade. ■

:⊢ **Definição 14.34.** Sejam $\boldsymbol{X}$ um espaço topológico, $p \in X$ e $(C_i)_{i\in I}$ uma indexação dos conjuntos conexos que contêm p. A *componente conexa* de $\boldsymbol{X}$ em p é o conjunto conexo

$$\Gamma_p := \bigcup_{i\in I} C_i.$$

O conjunto Γ_p é conexo pela proposição 14.63, pois a interseção de $(C_i)_{i\in I}$ contém p. A componente conexa em um ponto é o maior conjunto conexo que contém o ponto. A relação de estar na mesma componente conexa é uma relação de equivalência, como mostra a proposição a seguir, pois o conjunto de componentes conexas é uma partição de X.

⊢ **Proposição 14.64.** *As componentes conexas de um espaço topológico $\boldsymbol{X}$ são uma partição de X.*

□ *Demonstração.* Claramente nenhuma componente conexa é vazia, pois contém o próprio ponto. Ainda, a união de todas as componentes conexas é X, pelo mesmo motivo. Falta mostrar que elas são disjuntas duas a duas. Sejam $p, q \in X$ pontos distintos. Vamos mostrar que $\Gamma_p = \Gamma_q$ ou $\Gamma_p \cap \Gamma_q = \emptyset$. Se $C_p \neq C_q$ e $C_p \cap C_q \neq \emptyset$, então $C_P \cup C_q$ seria um conjunto conexo estritamente maior que C_p, contradizendo a maximalidade da componente conexa. ■

Na prática, podemos sempre reduzir o estudo de um espaço topológico ao estudo das suas componentes conexas, já que elas são abertos e definir funções contínuas no espaço é equivalente a definir em abertos que cobrem o espaço.

14.4.1.2 Conexidade por caminhos

:⊢ **Definição 14.35.** Seja $\boldsymbol{X}$ um espaço topológico. Um *caminho* em $\boldsymbol{X}$ é uma função contínua $c\colon [a, a'] \longrightarrow X$, em que $[a, a'] \subseteq \mathbb{R}$ é um intervalo compacto não degenerado ($a \neq a'$). O *ponto inicial* de c é $c(a)$ e o *ponto final* de c é $c(a')$. Sejam $x, x' \in X$. Um *caminho de x para x'* é um caminho c em $\boldsymbol{X}$ cujo ponto inicial é x e cujo ponto final é x'. Nesse caso o caminho c *conecta* x a x'.

:⊢ **Definição 14.36.** Sejam $\boldsymbol{X}$ um espaço topológico e $c\colon [a, a'] \longrightarrow X$ um caminho em $\boldsymbol{X}$. Uma *reparametrização* de c é um caminho $c'\colon [b, b'] \longrightarrow X$ tal que, para algum homeomorfismo $h\colon [b, b'] \longrightarrow [a, a']$,

$$c' = c \circ h.$$

:⊢ **Definição 14.37.** Sejam $\boldsymbol{X}$ um espaço topológico e $c\colon [a, a'] \longrightarrow X$, $c'\colon [a', a''] \longrightarrow X$ caminhos em $\boldsymbol{X}$ tais que $c(a') = c'(a')$. A *concatenação* de c com c' é o caminho

$$\begin{aligned} c \smile c' \colon [a, a''] &\longrightarrow X \\ t &\longmapsto \begin{cases} c(t), & t \in [a, a'] \\ c'(t), & t \in [a', a''] \end{cases}. \end{aligned}$$

:⊢ **Definição 14.38.** Um espaço topológico *conexo por caminhos* é um espaço topológico $\boldsymbol{X}$ tal que, para todos $x, x' \in X$, algum caminho em $\boldsymbol{X}$ conecta x a x'.

▷ **Exercício 14.4.** *Um espaço topológico conexo por caminhos é conexo.*

14.5 Compacidade

As propriedade de compacidade de um espaço topológico são propriedades relacionadas a coberturas de um espaço. De certa forma, a compacidade é uma noção de 'finitude topológica'.

14.5.1 Compacidade

:⊢ **Definição 14.39.** Um espaço topológico *compacto* é um espaço topológico $\boldsymbol{X}$ em que toda cobertura aberta de $\boldsymbol{X}$ tem subcobertura finita. Um subconjunto *compacto* de $\boldsymbol{X}$ é um subconjunto de X que é compacto com a topologia de subespaço.

▷ **Exercício 14.5.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes*

1. *$\boldsymbol{X}$ é compacto;*
2. *Toda rede em $\boldsymbol{X}$ tem sub-rede convergente.*

▷ **Exercício 14.6.** *Seja $\boldsymbol{X}$ um espaço topológico.*

1. *Para todo $C \subseteq X$ e todo $F \subseteq X$, se C é compacto e F é fechado, então F é compacto;*
2. *Para todo $S \subseteq X$ e todo $C \subseteq X'$, C é compacto em $\boldsymbol{S}$ se, e somente se, é compacto em $\boldsymbol{X}$.*

⊢ **Proposição 14.65.** *Seja $\boldsymbol{X}$ um espaço topológico separado. Para todo compacto $C \subseteq X$ e todo ponto $x \in X \smallsetminus C$, existem vizinhanças abertas V' de C e V de x que são disjuntas.*

□ *Demonstração.* Como $\boldsymbol{X}$ é separado, para todo $c \in C$ existem vizinhanças abertas V'_c de c e V_c de x que são disjuntas. Como C é compacto e $\{V'_c\}_{c \in C}$ é cobertura de C, existem $c_0, \cdots, c_{n-1} \in C$ tais que

$$C \subseteq \bigcup_{i \in [n]} V'_{c_i}$$

Definindo $V' := \bigcup_{i \in [n]} V'_{c_i}$ e $V := \bigcap_{i \in [n]} V_{c_i}$, segue que V' e V são vizinhanças abertas de C e x, respectivamente, e $V' \cap V = \emptyset$. ■

▷ **Exercício 14.7.** *Seja $\boldsymbol{X}$ um espaço topológico separado.*

1. *Para todo $C \subseteq X$, se C é compacto, então C é fechado;*
2. *Para todos $C, F \subseteq X$, se C é compacto e F é fechado, então $C \cap F$ é compacto.*

⊢ **Proposição 14.66.** *Sejam $\boldsymbol{X}_0$ e $\boldsymbol{X}_1$ espaços topológicos e $f \colon X_0 \to X_1$ uma função contínua. Se $\boldsymbol{X}_0$ é compacto, então $f(\boldsymbol{X}_0)$ é compacto.*

□ *Demonstração.* Seja $(C_I)_{i \in I}$ uma cobertura aberta de $f(\boldsymbol{X}_0)$. A família

$$(f^{-1}(C_i))_{i \in I}$$

é uma cobertura de X_1, pois

$$\bigcup_{i \in I} f^{-1}(C_i) = f^{-1}\left(\bigcup_{i \in I} C_i\right) = f^{-1}(X_1) = X_0.$$

A cobertura é aberta pois f é contínua. Como $\boldsymbol{X}_0$ é compacto, existem i_0, ..., $i_{n-1} \in I$ tal que $(f^{-1}(A_{i_k}))_{k\in[n]}$ é uma cobertura aberta de $\boldsymbol{X}_0$. Então $(A_{i_k})_{k\in[n]}$ é uma cobertura aberta de $f(\boldsymbol{X}_0)$, pois

$$f(X_0) = f\left(\bigcup_{k\in[n]} f^{-1}(A_{i_k})\right) = \bigcup_{k\in[n]} f(f^{-1}(A_{i_k})) \subseteq \bigcup_{k\in[n]} A_{i_k}.$$

Isso mostra que $f(\boldsymbol{X}_0)$ é compacto. ■

:⊢ **Definição 14.40.** Sejam $\boldsymbol{X}$ um espaço topológico e $\boldsymbol{L}$ um espaço linear topológico. O espaço das funções contínuas de $\boldsymbol{X}$ para $\boldsymbol{L}$ com suporte compacto é denotado $\mathscr{C}_c(X, L)$.

14.5.1.1 Compacidade relativa

:⊢ **Definição 14.41.** Seja $\boldsymbol{X}$ um espaço topológico. Um subconjunto *relativamente compacto* de $\boldsymbol{X}$ é um subconjunto de X cujo fecho em $\boldsymbol{X}$ é compacto.

14.5.2 Compacidade contável

:⊢ **Definição 14.42.** Um espaço topológico *contavelmente compacto*[1] é um espaço topológico $\boldsymbol{X}$ em que toda cobertura aberta $\mathcal{C}$ de $\boldsymbol{X}$ tem subcobertura contável.

Alguns autores definem compacidade contável como a propriedade de que toda cobertura aberta contável admite subcobertura finita.

14.5.3 Compacidade regional e local

:⊢ **Definição 14.43.** Um espaço topológico *regionalmente compacto* é um espaço topológico $\boldsymbol{X}$ em que todo ponto tem uma vizinhança compacta.

Um espaço topológico *localmente compacto* é um espaço topológico $\boldsymbol{X}$ em que todo ponto tem uma base de vizinhanças compacta.

Um espaço topológico *localmente relativamente compacto* é um espaço topológico $\boldsymbol{X}$ em que todo ponto tem uma base de vizinhanças relativamente compactas.

▷ **Exercício 14.8.** *Seja $\boldsymbol{X}$ um espaço topológico.*

1. *Se $\boldsymbol{X}$ é localmente compacto, então é regionalmente compacto;*
2. *Se $\boldsymbol{X}$ é localmente relativamente compacto, então é regionalmente compacto.*

▷ **Exercício 14.9.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes:*

[1] Conhecido na literatura por 'Lindelöf'.

1. *$\boldsymbol{X}$ é localmente compacto;*
2. *Toda vizinhança de um ponto de $\boldsymbol{X}$ contém uma uma vizinhanças compacta.*

▷ **Exercício 14.10.** *Seja $\boldsymbol{X}$ um espaço topológico. São equivalentes:*

1. *$\boldsymbol{X}$ é localmente relativamente compacto;*
2. *Todo ponto de $\boldsymbol{X}$ tem uma vizinhança relativamente compacta;*
3. *Todo ponto de $\boldsymbol{X}$ tem uma vizinhança fechada e compacta.*

As recíprocas são possíveis quando o espaço topológico é separado.

▷ **Exercício 14.11.** *Seja $\boldsymbol{X}$ um espaço topológico separado e regionalmente compacto. Então $\boldsymbol{X}$ é localmente compacto e localmente relativamente compacto.*

▷ **Exercício 14.12.** *Seja $\boldsymbol{X}$ um espaço topológico separado e regionalmente compacto.*

1. *Para todo compacto $C \subseteq X$ e toda vizinhança aberta $A \subseteq X$ de C, existe aberto $V \subseteq X$ com fecho compacto tal que*
$$C \subseteq V \subseteq \overline{V} \subseteq A.$$
2. *Para todo compacto C e todo aberto A de X, existe função $f \in \mathscr{C}_c(X, [0,1])$ que separa C e $A^{\complement}$.*

14.5.4 Paracompacidade

:⊢ **Definição 14.44.** Um espaço topológico *paracompacto* é um espaço topológico $\boldsymbol{X}$ em que toda cobertura aberta $\mathcal{C}$ de $\boldsymbol{X}$ tem refinamento aberto localmente finito.

14.5.5 Compacidade sequencial

:⊢ **Definição 14.45.** Um espaço topológico *sequencialmente compacto* é um espaço topológico $\boldsymbol{X}$ em que toda sequência convergente tem subsequência convergente.

14.6 Contabilidades

14.6.1 Base de vizinhanças contável (1º contável)

14.6.2 Base contável (2º contável)

14.7 Espaços de funções

Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos. Queremos construir uma topologia para o espaço $\mathscr{C}(\boldsymbol{X}, \boldsymbol{X}')$ das funções contínuas de $\boldsymbol{X}$ para $\boldsymbol{X}'$. Existem três principais topologias

que podemos considerar nesse conjunto; a primeira, chamada topologia pontual, ignora a estrutura topológica de $\boldsymbol{X}$ e a segunda e terceiras, chamadas topologia compacto-aberta e topologia fechado-aberto, respectivamente, a levam em consideração. Para mais detalhes, conferir [**art:Arens-TopologiesforHomeomorphismGroups**].

14.7.1 Topologia pontual

Para expor a primeira topologia, vamos generalizar um pouco o contexto e considerar o conjunto de funções de um conjunto C para um espaço topológico $\boldsymbol{X}$, o espaço de funções $\mathcal{F}(C, X)$. Esse espaço pode ser identificado com o espaço produto X^C e portanto podemos adotar sua topologia produto, cujos abertos sub-básicos são da forma $\mathcal{A}_{c_0} := \prod_{c \in C} A_x$, em que $c_0 \in C$, $A_{c_0} \subseteq X$ é aberto e, para todo $c \in c \smallsetminus \{c_0\}$, $A_c = X$. O conjunto de pontos desse aberto de X^C corresponde ao conjunto de funções $f \in \mathcal{F}(C, X)$ tais que $f(c_0) \in A_{c_0}$.

:⊢ **Definição 14.46.** Sejam C um conjunto e $\boldsymbol{X}$ um espaço topológico. A topologia *pontual* (ou *de convergência pontual* ou *finito-aberto*) em $\mathcal{F}(C, X)$ é a topologia gerada pelos abertos sub-básicos

$$\mathcal{A}_{F,A} := \{f \in \mathcal{F}(C, X) \mid f(F) \subseteq A\},$$

em que $F \subseteq C$ é um conjunto finito e $A \subseteq X$ é um conjunto aberto.

Pode-se mostrar que essa é a topologia produto de $\boldsymbol{X}^C$.

14.7.2 Topologia compacto-aberto

:⊢ **Definição 14.47.** Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos. A topologia *compacto-aberto* em $\mathscr{C}(\boldsymbol{X}, \boldsymbol{X}')$ é a topologia $\mathcal{T}_{\mathscr{C}}$ gerada pelos abertos sub-básicos

$$\mathscr{A}_{K,A} := \{f \in \mathscr{C}(\boldsymbol{X}, \boldsymbol{X}') \mid f(K) \subseteq A\},$$

em que $K \subseteq X$ é compacto e $A \subseteq X'$ é aberto.

14.7.3 Topologia compacto-cocompacto

:⊢ **Definição 14.48.** Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos. A topologia *compacto-cocompacto* em $\mathscr{C}(\boldsymbol{X}, \boldsymbol{X}')$ é a topologia $\mathcal{T}'_{\mathscr{C}}$ gerada pelos abertos sub-básicos

$$\mathscr{A}_{F,A} := \{f \in \mathscr{C}(\boldsymbol{X}, \boldsymbol{X}') \mid f(F) \subseteq A\},$$

em que $F \subseteq X$ é fechado, $A \subseteq X'$ é aberto e, ou F é compacto. ou $A^{\complement}$ é compacto.

14.8 Topologia algébrica

14.8.1 Homotopia

:⊢ **Definição 14.49.** Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos e $f, f' \in \mathscr{C}(X, X')$ funções contínuas. Uma *homotopia* de f para f' é uma função contínua

$$\begin{aligned} H\colon [0,1] \times X &\longrightarrow X' \\ (t,x) &\longmapsto H_t(x) \end{aligned}$$

tal que $H_0 = f$ e $H_1 = f'$. As funções f e f são *homotópicas* e denota-se $f \approx f$.

⊢ **Proposição 14.67.** *Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos. A relação de homotopia é uma equivalência no conjunto $\mathscr{C}(X, X')$ das funções contínuas de $\boldsymbol{X}$ para $\boldsymbol{X}'$.*

□ *Demonstração.*

1. (Reflexividade) Seja $f \in \mathscr{C}(X, X')$. Consideremos a função

$$\begin{aligned} H\colon [0,1] \times X &\longrightarrow X' \\ (t,x) &\longmapsto f(x) \end{aligned}$$

Então H é uma função contínua, pois f é contínua, e vale que $H_0 = H_1 = f$, portanto $f \approx f$.

2. (Simetria) Sejam $f, f' \in \mathscr{C}(X, X')$ tais que $f \approx f'$ e $H\colon [0,1] \times X \longrightarrow X'$ uma homotopia de f para f'. Consideremos a função

$$\begin{aligned} H'\colon [0,1] \times X &\longrightarrow X' \\ (t,x) &\longmapsto H_{1-t}(x). \end{aligned}$$

Como H e $1-t$ são contínuas, a função H' é contínua. Ainda, notemos que $H'_0 = H_1 = f'$ e $H'_1 = H_0 = f$, portanto $f' \approx f$.

3. (Transitividade) Sejam $f, f', f'' \in \mathscr{C}(X, X')$ tais que $f \approx f'$ e $f' \approx f''$ e $H\colon [0,1] \times X \longrightarrow X'$ uma homotopia de f para f' e $H'\colon [0,1] \times X \longrightarrow X'$ uma homotopia de f' para f''. Consideremos a função

$$\begin{aligned} H''\colon [0,1] \times X &\longrightarrow X' \\ (t,x) &\longmapsto \begin{cases} H_{2t}(x), & t \in \left[0, \frac{1}{2}\right] \\ H'_{2t-1}(x), & t \in \left[\frac{1}{2}, 1\right] \end{cases}. \end{aligned}$$

Como $H_1 = f' = H'_0$ e H e H' são contínuas, então H'' é contínua. Ainda, como H é uma homotopia de f para f', então $H''_0 = H_0 = f$ e, como H' é uma homotopia de f' para f'', então $H''_1 = H'_1 = f''$, portanto H'' é uma homotopia de f para f''.

■

$\vdash$ **Proposição 14.68.** *Sejam $\boldsymbol{X}$, $\boldsymbol{X}'$ e $\boldsymbol{X}''$ espaços topológicos, $f, f' \in \mathscr{C}(X, X')$ e $g, g' \in \mathscr{C}(X', X'')$ funções contínuas tais que $f \approx f'$ e $g \approx g'$. Então*

$$(g \circ f) \approx (g' \circ f').$$

$\square$ *Demonstração.* Sejam $H\colon [0,1] \times X \longrightarrow X'$ uma homotopia de f para f' e $H'\colon [0,1] \times X' \longrightarrow X''$ uma homotopia de g para g'. Consideremos a função

$$\begin{aligned} H''\colon [0,1] \times X &\longrightarrow X'' \\ (t, x) &\longmapsto H'_t \circ H_t(x). \end{aligned}$$

Como H e H' são contínuas, então H'' é contínua. Como H é homotopia de f para f' e H' é homotopia de g para g', então $H_0 = f$, $H_1 = f'$, $H'_0 = g$ e $H'_1 = g'$, o que implica

$$H''_0 = H'_0 \circ H_0 = g \circ f$$

e

$$H''_1 = H'_1 \circ H_1 = g' \circ f',$$

o que mostra que H'' é homotopia de $g \circ f$ para $g' \circ f'$. ■

$:\vdash$ **Definição 14.50.** Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços topológicos. Uma *equivalência homotópica* entre $\boldsymbol{X}$ e $\boldsymbol{X}'$ é uma par $(f, f') \in \mathscr{C}(X, X') \times \mathscr{C}(X', X)$ de funções contínuas tais que $f' \circ f \approx \mathrm{I}_X$ e $f \circ f' \approx \mathrm{I}_{X'}$. Os espaços $\boldsymbol{X}$ e $\boldsymbol{X}'$ são *homotopicamente equivalentes* e denota-se $X \approx X'$.

14.8.2 Caminhos e laços

$:\vdash$ **Definição 14.51** (Caminho e laço)**.** Seja $\boldsymbol{X}$ um espaço topológico. Um *caminho* em X é uma função contínua $c\colon [0,1] \longrightarrow X$. Um *laço* em X é uma função contínua $\ell\colon \mathbb{S}^1 \longrightarrow X$ e a *origem* desse laço é o ponto $\ell(0)$. Denotaremos o conjunto dos laços em X com origem em $x_0 \in X$ por $L(X, x_0)$

Note que um caminho c tal que $c(0) = c(1)$ gera um laço, e vice-versa.

$:\vdash$ **Definição 14.52.** Seja X um espaço métrico e $c\colon [0,1] \longrightarrow X$ um caminho em X. O *caminho inverso* de c é o caminho

$$\begin{aligned} c^{-1}\colon [0,1] &\longrightarrow X \\ s &\longmapsto c(1-s). \end{aligned}$$

$:\vdash$ **Definição 14.53.** Seja X um espaço métrico e $x_0 \in X$. O *caminho constante* em x_0 é o caminho

$$\begin{aligned} e_{x_0}\colon [0,1] &\longrightarrow X \\ s &\longmapsto x_0. \end{aligned}$$

:⊢ **Definição 14.54.** Sejam X um espaço métrico e $c_1, c_2 : [0,1] \to X$ caminhos em X tais que $c_1(1) = c_2(0)$. A *composição* dos caminhos c_1 e c_2 é o caminho

$$\begin{aligned}(c_1 \cdot c_2)\colon [0,1] &\longrightarrow X \\ s &\longmapsto \begin{cases} c_1(2s), & s \in \left[0, \frac{1}{2}\right] \\ c_2(2s-1), & s \in \left]\frac{1}{2}, 1\right] \end{cases}.\end{aligned}$$

14.8.3 Homotopia de caminhos

:⊢ **Definição 14.55.** Sejam X um espaço métrico e $c_1 : [0,1] \to X$ e $c_2 : [0,1] \to X$ caminhos em X tal que $c_1(0) = c_2(0)$ e $c_1(1) = c_2(1)$. Uma *homotopia de caminhos* entre c_1 e c_2 é uma homotopia H entre c_1 e c_2 tal que, para todo $t \in [0,1]$, $H(0,t) = c_1(0)$ e $H(1,t) = c_1(1)$. No caso de existir uma homotopia de caminhos entre c_1 e c_2, denota-se $c_1 \approx c_2$.

FIGURA 14.1: Ilustração de uma homotopia de caminhos.

⊢ **Proposição 14.69.** *Sejam X um espaço métrico e $x_0 \in X$. Então a relação $\approx$ de homotopia de caminhos é uma relação de equivalência em $L(X, x_0)$.*

□ *Demonstração.* Sejam $l_1, l_2, l_3 \in L(X, x_0)$. Então

1. Reflexividade: $l_1 \approx l_1$.
   Consideremos a função

$$\begin{aligned} H : S^1 \times [0,1] &\to X \\ (x,t) &\mapsto l(x). \end{aligned}$$

   Sabemos que H é uma homotopia entre l_1 e l_1. Basta notar que, para todo $t \in [0,1]$, $H(0,t) = H(1,t) = l(0)$, o que termina a demonstração de que H é uma homotopia de laços entre l_1 e l_1.

2. Simetria: $l_1 \approx l_2 \Rightarrow l_2 \approx l_1$.
   Seja $H : S^1 \times [0,1] \to X$ uma homotopia de laços entre l_1 e l_2. Então consideremos a função

$$\begin{aligned} H' : S^1 \times [0,1] &\to X \\ (x,t) &\mapsto H(x, 1-t). \end{aligned}$$

   Sabemos que H é uma homotopia entre l_2 e l_1. Basta notar que, para todo $t \in [0,1]$, $H'(0,t) = H(0,1-t) = l(0) = H(1,1-t) = H'(1,t)$, o que termina a demonstração de que H' é uma homotopia de laços entre l_2 e l_1.
3. Transitividade: $l_1 \approx l_2$ e $l_2 \approx l_3 \Rightarrow l_1 \approx l_3$.
   Sejam $H_1 : S^1 \times [0,1] \to X$ uma homotopia entre l_1 e l_2 e $H_2 : S^1 \times [0,1] \to X$ uma homotopia entre l_2 e l_3. Então consideremos a função

$$\begin{aligned} H : S^1 \times [0,1] &\to X \\ (x,t) &\mapsto \begin{cases} H_1(x, 2t) & \text{se } 0 \leq t \leq \frac{1}{2} \\ H_2(x, 2t-1) & \text{se } \frac{1}{2} \leq t \leq 1. \end{cases} \end{aligned}$$

   Sabemos que H é uma homotopia entre l_1 e l_3. Basta notar que, como H_1 é homotopia de laços, para todo $t \in [0, \frac{1}{2}]$, $H(0,t) = H_1(0,2t) = l_1(0)$ e, como H_2 é homotopia de laços entre l_2 e l_3, para todo $t \in [\frac{1}{2}, 1]$, $H(0,t) = H_2(0, 2t-1) = l_2(0) = l_1(0)$ o que termina a demonstração de que H é uma homotopia de laços entre l_1 e l_3.

■

⊢ **Proposição 14.70.** *Seja X um espaço métrico e $c_1, c_2, c_3 : [0,1] \to X$ caminhos em X tais que $c_1(0) = x_0$, $c_1(1) = c_2(0)$, $c_2(1) = c_3(0)$ e $c_3(0) = x_1$. Então*

1. $c_1 \cdot (c_1)^{-1} \approx e_{x_0}$;
2. $(c_1)^{-1} \cdot c_1 \approx e_{x_1}$;
3. $e_{x_0} \cdot c_1 \approx c_1 \approx c_1 \cdot e_{x_1}$;
4. $(c_1 \cdot c_2) \cdot c_3 \approx c_1 \cdot (c_2 \cdot c_3)$.

□ *Demonstração.* 1. Notemos que

$$c_1 \cdot (c_1)^{-1}(s) = \begin{cases} c_1(2s) & \text{se } s \in [0, \frac{1}{2}] \\ (c_1)^{-1}(2s-1) & \text{se } s \in [\frac{1}{2}, 1]. \end{cases} = \begin{cases} c_1(2s) & \text{se } s \in [0, \frac{1}{2}] \\ c_1(2-2s) & \text{se } s \in [\frac{1}{2}, 1]. \end{cases}$$

Assim, considerando a parametrização $\phi : [0,1] \to [0,1]$

$$\phi(s) = \begin{cases} 2s & \text{se } s \in [0, \frac{1}{2}] \\ 2-2s & \text{se } s \in [\frac{1}{2}, 1] \end{cases}$$

segue que $c_1 \cdot (c_1)^{-1}(s) = c_1(\phi(s))$. Consideremos, assim, a função

$$\begin{aligned} H : [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto c_1((1-t)\phi(s)). \end{aligned}$$

Temos que H é contínua, pois c_1, $1-t$ e ϕ são contínuas. Agora, notemos que, para todo $s, t \in [0,1]$, $1-t \in [0,1]$ e $\phi(s) \in [0,1]$, o que mostra que $(1-t)\phi(s) \in [0,1]$ e, portanto, que H está bem definida. Ainda, para todo $s \in [0,1]$, $H(s,0) = c_1(\phi(s)) = c_1 \cdot (c_1)^{-1}(s)$ e $H(s,1) = c_1(0) = x_0$. Portanto H é homotopia entre $c_1 \cdot (c_1)^{-1}$ e e_{x_0}. Para mostrar que H é homotopia de caminhos, note que, para todo $t \in [0,1]$, $H(0,t) = c_1((1-t)\phi(0)) = c_1(0)$ e $H(1,t) = c_1((1-t)\phi(1)) = c_1(0)$, o que termina a demonstração.

2. Análogo ao item anterior, mas considerando a parametrização $\phi : [0,1] \to [0,1]$

$$\phi(s) = \begin{cases} 1-2s & \text{se } s \in [0, \frac{1}{2}] \\ 2s-1 & \text{se } s \in [\frac{1}{2}, 1]. \end{cases}$$

3. Análogo ao item anterior, mas considerando as parametrizações ϕ, ϕ': $[0,1] \longrightarrow [0,1]$

$$\phi(s) = \begin{cases} 0 & \text{se } s \in [0, \frac{1}{2}] \\ 2s-1 & \text{se } s \in [\frac{1}{2}, 1]. \end{cases}$$

$$\phi'(s) = \begin{cases} 2s & \text{se } s \in [0, \frac{1}{2}] \\ 0 & \text{se } s \in [\frac{1}{2}, 1]. \end{cases}$$

4. Notemos que

$$(c_1 \cdot c_2) \cdot c_3 = \begin{cases} c_1(4s) & \text{se } s \in [0, \frac{1}{4}] \\ c_2(4s-1) & \text{se } s \in [\frac{1}{4}, \frac{1}{2}] \\ c_3(2s-1) & \text{se } s \in [\frac{1}{2}, 1] \end{cases}$$

$$c_1 \cdot (c_2 \cdot c_3) = \begin{cases} c_1(2s) & \text{se } s \in [0, \frac{1}{2}] \\ c_2(4s-2) & \text{se } s \in [\frac{1}{2}, \frac{3}{4}] \\ c_3(4s-3) & \text{se } s \in [\frac{3}{4}, 1]. \end{cases}$$

Assim, considerando a parametrização $\phi : [0,1] \to [0,1]$

$$\phi(s) = \begin{cases} 2s & \text{se } s \in [0, \frac{1}{4}] \\ s + \frac{1}{4} & \text{se } s \in [\frac{1}{4}, \frac{1}{2}] \\ \frac{s}{2} + \frac{1}{2} & \text{se } s \in [\frac{1}{2}, 1]. \end{cases}$$

segue que $((c_1 \cdot c_2) \cdot c_3)(s) = (c_1 \cdot (c_2 \cdot c_3))(\phi(s))$. Consideremos, assim, a função

$$\begin{aligned} H : [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto (c_1 \cdot c_2) \cdot c_3((1-t)s + t\phi(s)). \end{aligned}$$

Analogamente aos itens anteriores, mostra-se que H é homotopia de caminhos.

■

⊢ **Proposição 14.71.** *Sejam X um espaço métrico, $x_0 \in X$ e $c_1, c_2, c_1', c_2' : [0,1] \to X$ caminhos em X. Então*

1. *Se $c_1 \approx c_1'$ e $c_2 \approx c_2'$, então $c_1 \cdot c_2 \approx c_1' \cdot c_2'$.*
2. *$(c_1)^{-1} \approx (c_1')^{-1}$.*

□ *Demonstração.* 1. Seja H_1 a homotopia de caminhos entre c_1 e c_1' e H_2 a homotopia de caminhos entre c_2 e c_2'. Consideremos a função

$$\begin{aligned} H : [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto \begin{cases} H_1(2s,t) & \text{se } 0 \leq s \leq \frac{1}{2} \\ H_2(2s-1,t) & \text{se } \frac{1}{2} \leq s \leq 1. \end{cases} \end{aligned}$$

Primeiro, notemos que H é uma homotopia entre c_1 e c_2. Pa Como H_1 e H_2 são contínuas, basta mostrar que H é contínua nos pontos em que $s = \frac{1}{2}$. Para isso, notemos que, para todo $t \in [0,1]$, $H_1(2\frac{1}{2}, t) = H_1(1,t) = c_1'(1)$, pois H_1 é homotopia de caminhos, e $H_2(2\frac{1}{2} - 1, t) = H_2(0,t) = c_2(0)$, pois H_2 é homotopia de caminhos. Assim, segue que as funções nesses pontos são iguais e, portanto, H é contínua. Ainda, para todo $s \in [0,1]$,

$$H(s,0) = \begin{cases} H_1(2s,0) = c_1(2s) & \text{se } 0 \leq s \leq \frac{1}{2} \\ H_2(2s-1,0) = c_2(2s-1) & \text{se } \frac{1}{2} \leq s \leq 1, \end{cases}$$

o que mostra que $H(s,0) = (c_1 \cdot c_2)(s)$, e

$$H(s,0) = \begin{cases} H_1(2s,1) = c_1'(2s) & \text{se } 0 \leq s \leq \frac{1}{2} \\ H_2(2s-1,1) = c_2'(2s-1) & \text{se } \frac{1}{2} \leq s \leq 1, \end{cases}$$

o que mostra que $H(s,0) = (c_1' \cdot c_2')(s)$ e, portanto, que H é homotopia.
Agora, devemos mostrar que H é homotopia de caminhos. Notemos que, para todo $t \in [0,1]$, $H(0,t) = H_1(0,t) = c_1(0) = (c_1 \cdot c_2)(0)$, pois H_1 é homotopia de caminhos, e $H(1,t) = H_2(1,t) = c_2'(1) = (c_1' \cdot c_2')(1)$, o que mostra que H é homotopia de caminhos.

2. Seja H uma homotopia de caminhos entre c_1 e c_1'. Consideremos a função

$$\begin{aligned} H' : [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto H(1-s,t). \end{aligned}$$

Primeiro notemos que H' é uma homotopia entre $(c_1)^{-1}$ e $(c_1')^{-1}$. Claramente, H' é contínua, pois H e $1-s$ são contínuas. Ainda, para todo $s \in [0,1]$,

$$H'(s,0) = H(1-s,0) = c_1(1-s) = (c_1)^{-1}(s)$$

e

$$H'(s,1) = H(1-s,1) = c_1'(1-s) = (c_1')^{-1}(s),$$

pois H é homotopia. Assim, mostramos que H' é homotopia entre $(c_1)^{-1}$ e $(c_1')^{-1}$.
Agora, mostremos que H' é homotopia de caminhos. Para todo $t \in [0,1]$, $H'(0,t) = H(1,t) = c_1'(1) = (c_1')^{-1}(0)$ e $H'(1,t) = H(0,t) = c_1(0) = (c_1)^{-1}(1)$, pois H é homotopia de caminhos. Assim, mostramos que H' é homotopia de caminhos entre $(c_1)^{-1}$ e $(c_1')^{-1}$.

■

14.8.4 Grupo fundamental

Como $\approx$ é uma relação de equivalência em $L(X,x_0)$, podemos considerar o espaço quociente $L(X,x_0)/\approx$ das classes de equivalência de laços em $L(X,x_0)$.

:⊢ **Definição 14.56.** Sejam X um espaço métrico conexo por caminhos e $x_0 \in X$. Então o *grupo fundamental* de X com base em x_0 é o conjunto

$$\pi_1(X,x_0) := L(X,x_0)/\approx .$$

:⊢ **Definição 14.57.** Sejam X um espaço métrico conexo por caminhos. A *composição* de classes de equivalência de laços em $\pi_1(X)$ é a função

$$\begin{aligned} \cdot : \pi_1(X) \times \pi_1(X) &\to \pi_1(X) \\ ([l_1],[l_2]) &\mapsto [l_1 \cdot l_2]. \end{aligned}$$

Notemos que a composição de caminhos está bem definida por causa da proposição 14.71.

⊢ **Teorema 14.72.** *Sejam X um espaço métrico conexo por caminhos e $x_0 \in X$. Então $(\pi_1(X),\cdot)$ é um grupo.*

□ *Demonstração.* Segue direto das proposições 14.70 e 14.71. ■

Capítulo 15

Medida

15.1 Espaço mensurável

15.1.1 Sigma-álgebras e sub-sigma-álgebras

:⊢ **Definição 15.1.** Seja X um conjunto. Uma *sigma-álgebra* sobre X é um conjunto $\mathcal{M} \subseteq \wp(X)$ de subconjuntos X que satisfaz

1. (Vazio) $\emptyset \in \mathcal{M}$;
2. (Fechamento por complementação) Para todo $M \in \mathcal{M}$, $M^{\complement} \in \mathcal{M}$;
3. (Fechamento por união enumerável) Para toda sequência $(M_i)_{i\in\mathbb{N}}$ de conjuntos de $\mathcal{M}$,
$$\bigcup_{i\in\mathbb{N}} M_i \in \mathcal{M}.$$

Vale notar que uma sigma-álgebra $\mathcal{M}$ é uma álgebra booleana (3.38) e, portanto, todas propriedades de álgebras booleanas valem para uma sigma-álgebra. De fato, o *sigma* no nome vem da terceira propriedade das sigma-álgebras, pois veremos que essa propriedade tem a ver com um tipo de soma de medidas a ser definido adiante.

⊢ **Proposição 15.1.** *Seja X um conjunto não vazio e $\mathcal{M}$ uma sigma-álgebra sobre X. Então*

1. *(Universo) $X \in \mathcal{M}$;*
2. *(Fechamento por interseção enumerável) Para toda sequência $(M_i)_{i\in\mathbb{N}}$ de conjuntos de $\mathcal{M}$,*
$$\bigcap_{i\in\mathbb{N}} M_i \in \mathcal{M}.$$

□ *Demonstração.* 1. Da primeira propriedade de $\mathcal{M}$, tem-se que $\emptyset \in \mathcal{M}$. Da segunda propriedade de $\mathcal{M}$, tem-se que $X = \emptyset^{\complement} \in \mathcal{M}$.

2. Da segunda propriedade, tem-se que, para todo $i \in \mathbb{N}$, $M_i^{\complement} \in \mathcal{M}$. Da terceira propriedade de $\mathcal{M}$, tem-se que $\bigcup_{i\in\mathbb{N}} M_i^{\complement} \in \mathcal{M}$. Das Leis de De Morgan (3.47), tem-se que
$$\left(\bigcap_{i\in\mathbb{N}} M_i\right)^{\complement} = \bigcup_{i\in\mathbb{N}} (M_i)^{\complement} \in \mathcal{M},$$
e conclui-se que $\displaystyle\bigcap_{i\in\mathbb{N}} M_i \in \mathcal{M}$.

■

► **Exemplo 15.1.** $\mathcal{M} = \{\emptyset, X\}$ *e* $\mathcal{M} = \wp(X)$ *são sigma-álgebras sobre* X.

:⊢ **Definição 15.2.** Seja X um conjunto não vazio e $\mathcal{M}$ uma sigma-álgebra sobre X. Uma *sub-sigma-álgebra* de $\mathcal{M}$ é um conjunto $\mathcal{M}' \subseteq \mathcal{M}$ que é uma sigma-álgebra sobre X.

:⊢ **Definição 15.3.** Um *espaço mensurável* é um par $(X, \mathcal{M})$ em que X é um conjunto não vazio e $\mathcal{M}$ é uma sigma-álgebra sobre X. Um *conjunto mensurável* é um elemento da sigma-álgebra $\mathcal{M}$.

⊢ **Proposição 15.2.** *Seja* $(C_n)_{n\in\mathbb{N}}$ *uma sequência de conjuntos.*

1. *A sequência*
$$M_n := \bigcup_{k=0}^{n} C_k$$
é uma sequência crescente de conjuntos;
2. *A sequência* $D_0 := C_0$ *e, para* $n \in \mathbb{N}^*$,
$$D_n := C_n \setminus M_n$$
é uma sequência disjunta de conjuntos;
3. $$\bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} M_n = \bigcup_{n\in\mathbb{N}} D_n.$$

15.1.2 Sigma-álgebras geradas

⊢ **Proposição 15.3.** *Seja* X *um conjunto não vazio e* $(\mathcal{M}_i)_{i\in I}$ *uma família de sigma-álgebras sobre* X. *Então*
$$\mathcal{M} := \bigcap_{i\in I} \mathcal{M}_i$$
é uma sigma-álgebra sobre X.

□ *Demonstração.* Como $\mathcal{M}_i$ são sigma-álgebras, então $\emptyset \in \mathcal{M}_i$ para todo $i \in I$. Assim, segue que $\emptyset \in \mathcal{M}$. Ainda, se $A \in \mathcal{M}$, então $A \in \mathcal{M}_i$ para todo $i \in I$. Logo $A^{\complement} \in \mathcal{M}_i$ para todo $i \in I$, o que implica $A^{\complement} \in \mathcal{M}$. Por fim, se $(A_j)_{j\in\mathbb{N}}$ é uma sequência de conjuntos em $\mathcal{M}$, então $A_j \in \mathcal{M}$ para todo $j \in \mathbb{N}$. Mas isso implica que $A_j \in \mathcal{M}_i$ para todo $j \in \mathbb{N}$, $i \in I$, o que, por sua vez, implica que, para todo $i \in I$,

$$\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{M}_i.$$

Então conclui-se que $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{M}$ e, portanto, $\mathcal{M}$ é uma sigma-álgebra sobre X. ■

:⊢ **Definição 15.4.** Seja X um conjunto e $\mathcal{C} \in \wp(X)$ um conjunto de subconjuntos de X. A *sigma-álgebra gerada por* $\mathcal{C}$ é a interseção da família de todas as sigma-álgebras sobre X de que $\mathcal{C}$ é subconjunto, denotada $\langle \mathcal{C} \rangle$.

A sigma-álgebra gerada por um conjunto é a menor sigma-álgebra que contém esse conjunto no sentido que não existe subconjunto dessa sigma-álgebra que contenha o conjunto e também seja uma sigma-álgebra.

► **Exemplo 15.2.** *A sigma-álgebra sobre X gerada por $\emptyset$ é $\{\emptyset, X\}$.*

15.1.3 Limites de conjuntos

:⊢ **Definição 15.5.** Sejam X um conjunto e $(A_n)_{n\in\mathbb{N}}$ uma sequência de subconjuntos de X. O *limite inferior* de $(A_n)_{n\in\mathbb{N}}$ é o conjunto

$$\underline{\lim}\, A_n := \bigcup_{m=0}^{\infty} \left(\bigcap_{n=m}^{\infty} A_n \right).$$

Esse conjunto é o conjunto dos pontos que não pertencem todos menos finitos conjuntos A_n. O *limite superior* de $(A_n)_{n\in\mathbb{N}}$ é o conjunto

$$\overline{\lim}\, A_n := \bigcap_{m=0}^{\infty} \left(\bigcup_{n=m}^{\infty} A_n \right).$$

Esse conjunto é o conjunto dos pontos que pertencem a infinitos conjuntos A_n.

⊢ **Proposição 15.4.** *Sejam X um conjunto e $(A_n)_{n\in\mathbb{N}}$ uma sequência de subconjuntos de X. Então*

$$\emptyset \subseteq \underline{\lim}\, A_n \subseteq \overline{\lim}\, A_n \subseteq X.$$

⊢ **Proposição 15.5.** *Sejam X um conjunto e $(A_n)_{n\in\mathbb{N}}$ uma sequência de subconjuntos de X. Então*

1. *Se* $(A_n)_{n\in\mathbb{N}}$ *é monótona crescente,*

$$\underline{\lim}\, A_n = \bigcup_{n=0}^{\infty} A_n = \overline{\lim}\, A_n.$$

2. *Se* $(A_n)_{n\in\mathbb{N}}$ *é monótona decrescente,*

$$\underline{\lim}\, A_n = \bigcap_{n=0}^{\infty} A_n = \overline{\lim}\, A_n.$$

:⊢ **Definição 15.6.** Sejam X um conjunto e $(A_n)_{n\in\mathbb{N}}$ uma sequência de subconjuntos de X. Um *limite* de $(A_n)_{n\in\mathbb{N}}$ é um conjunto $\lim A_n$ tal que

$$\lim A_n = \underline{\lim}\, A_n = \overline{\lim}\, A_n.$$

15.2 Funções mensuráveis

:⊢ **Definição 15.7.** Sejam $\boldsymbol{X} = (X, \mathcal{M}_X)$ e $\boldsymbol{Y} = (Y, \mathcal{M}_Y)$ espaços mensuráveis. Uma *função mensurável* de $\boldsymbol{X}$ para $\boldsymbol{Y}$ é uma função $f\colon X \to Y$ tal que, para todo $M \in \mathcal{M}$,

$$f^{-1}(M) \in \mathcal{M}_X.$$

Denota-se $f\colon \boldsymbol{X} \to \boldsymbol{Y}$. O conjunto dessas funções é denotado $\mathscr{M}(\boldsymbol{X}, \boldsymbol{Y})$.

⊢ **Proposição 15.6.** *Seja* $\boldsymbol{X}$ *um espaço mensurável. A função* $\mathrm{I}_X : X \to X$ *é uma função mensurável.*

⊢ **Proposição 15.7.** *Sejam* $\boldsymbol{X_0}$*,* $\boldsymbol{X_1}$ *e* $\boldsymbol{X_2}$ *espaços mensuráveis e* $f_0 : \boldsymbol{X_0} \to \boldsymbol{X_1}$ *e* $f_1 : \boldsymbol{X_1} \to \boldsymbol{X_2}$ *funções mensuráveis. Então* $f_1 \circ f_0 : \boldsymbol{X_0} \to \boldsymbol{X_2}$ *é uma função mensurável.*

15.2.1 Sigma-álgebras puxadas e empurradas

:⊢ **Definição 15.8.** Sejam X um conjunto, $\boldsymbol{Y} = (Y, \mathcal{M}_Y)$ um espaço mensurável e $f : X \to Y$ uma função. A *sigma-álgebra puxada* por f é

$$f^*(\mathcal{M}_Y) := \left\{ f^{-1}(M) \;\middle|\; M \in \mathcal{M}_Y \right\}.$$

⊢ **Proposição 15.8.** *Sejam* X *um conjunto,* $\boldsymbol{Y} = (Y, \mathcal{M}_Y)$ *um espaço mensurável e* $f : X \to Y$ *uma função. Então* $\mathcal{M}_X := f^*(\mathcal{M}_Y)$*, a sigma-álgebra puxada por* f*, é uma sigma-álgebra sobre* X*.*

□ *Demonstração.* Primeiro, notemos que $\emptyset \in \mathcal{M}_X$, pois $\emptyset \in \mathcal{M}_Y$ e $f^{-1}(\emptyset) = \emptyset$ (3.13). Segundo, seja $B \in \mathcal{M}_X$. Então existe $A \in \mathcal{M}_Y$ tal que $B = f^{-1}(A)$. Como $\mathcal{M}_Y$ é uma sigma-álgebra, então $A^{\complement} \in \mathcal{M}_Y$, o que implica $f^{-1}(A^{\complement}) \in \mathcal{M}_X$. Mas $(f^{-1}(A))^{\complement} = f^{-1}(A^{\complement})$ (3.13). Então $B^{\complement} \in \mathcal{M}_X$. Terceiro, seja $(B_i)_{i\in\mathbb{N}}$ uma sequência de conjuntos de $\mathcal{M}_X$. Então, para todo $i \in I$, existe $A_i \in \mathcal{M}_Y$ tal que $B_i = f^{-1}(A_i)$. Como $\mathcal{M}_Y$ é uma sigma-álgebra, então $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{M}_Y$. Isso implica que $f^{-1}(\bigcup_{i\in\mathbb{N}} A_i) \in \mathcal{M}_X$. Mas $f^{-1}(\bigcup_{i\in\mathbb{N}} A_i) = \bigcup_{i\in\mathbb{N}} f^{-1}(A_i)$ (3.13). Então $\bigcup_{i\in\mathbb{N}} B_i \in \mathcal{M}_X$ e, assim, conclui-se que $\mathcal{M}_X$ é uma sigma-álgebra sobre X. ■

:⊢ **Definição 15.9.** Sejam $\boldsymbol{X} = (X, \mathcal{M}_X)$ um espaço mensurável, Y um conjunto e $f : X \to Y$ uma função. A *sigma-álgebra empurrada* por f é

$$f_*(\mathcal{M}_X) := \left\{ M \subseteq Y \;\middle|\; f^{-1}(M) \in \mathcal{M}_X \right\}.$$

⊢ **Proposição 15.9.** *Sejam $\boldsymbol{X} = (X, \mathcal{M}_X)$ um espaço mensurável, Y um conjunto e $f : X \to Y$ uma função. Então $\mathcal{M}_Y := f_*(\mathcal{M}_X)$, a sigma-álgebra empurrada por f, é uma sigma-álgebra sobre Y.*

□ *Demonstração.* Primeiro, notemos que $\emptyset \in \mathcal{M}_Y$, pois $\emptyset \in \mathcal{M}_X$ e $f^{-1}(\emptyset) = \emptyset$ (3.13). Segundo, seja $A \in \mathcal{M}_Y$. Então $f^{-1}(A) \in \mathcal{M}_X$, o que implica $(f^{-1}(A))^{\complement} \in \mathcal{M}_X$, pois $\mathcal{M}_X$ é sigma-álgebra. Mas $(f^{-1}(A))^{\complement} = f^{-1}(A^{\complement})$ (3.13), o que implica $f^{-1}(A^{\complement}) \in \mathcal{M}_X$ e, portanto, $A^{\complement} \in \mathcal{M}_Y$. Terceiro, seja $(A_i)_{i\in\mathbb{N}}$ uma sequência de conjuntos de $\mathcal{M}_Y$. Então, para todo $i \in \mathbb{N}$, $f^{-1}(A_i) \in \mathcal{M}_X$, o que implica que $\bigcup_{i\in\mathbb{N}} f^{-1}(A_i) \in \mathcal{M}_X$, pois $\mathcal{M}_X$ é uma sigma-álgebra. Mas, como $\bigcup_{i\in\mathbb{N}} f^{-1}(A_i) = f^{-1}(\bigcup_{i\in\mathbb{N}} A_i)$ (3.13), segue que $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{M}_Y$ e, assim, conclui-se que $\mathcal{M}_Y$ é uma sigma-álgebra sobre Y. ■

⊢ **Proposição 15.10.** *Sejam $\boldsymbol{X} = (X, \mathcal{M}_X)$ e $\boldsymbol{Y} = (Y, \mathcal{M}_Y)$ espaços mensuráveis. Uma função $f : X \to Y$ é função mensurável de $\boldsymbol{X}$ para $\boldsymbol{Y}$ se, e somente se, a sigma-álgebra $f^*(\mathcal{M}_Y)$ puxada por f é uma sub-sigma-álgebra de $\mathcal{M}_X$.*

$$f \in \mathscr{M}(\boldsymbol{X}, \boldsymbol{Y}) \quad \Longleftrightarrow \quad f^*(\mathcal{M}_Y) \subseteq \mathcal{M}_X.$$

□ *Demonstração.* Suponha que f é uma função mensurável. Seja $B \in f^*(\mathcal{M}_Y)$. Então existe $A \in \mathcal{M}_Y$ tal que $B = f^{-1}(A)$. Como f é mensurável, vale $f^{-1}(A) \in \mathcal{M}_X$, o que implica $B \in \mathcal{M}_X$ e, portanto, $f^*(\mathcal{M}_Y) \subseteq \mathcal{M}_X$. Como $f^*(\mathcal{M}_Y)$ é uma sigma-álgebra sobre X pela proposição acima, segue que $f^*(\mathcal{M}_Y)$ é uma sub-sigma-álgebra de $\mathcal{M}_X$. Reciprocamente, suponha que $f^*(\mathcal{M}_Y)$ é uma sub-sigma-álgebra de $\mathcal{M}_X$. Seja $A \in \mathcal{M}_Y$. Então $f^{-1}(A) \in f^*(\mathcal{M}_Y)$. Mas como $f^*(\mathcal{M}_Y)$ é uma sub-sigma-álgebra de $\mathcal{M}_X$, segue que $f^{-1}(A) \in \mathcal{M}_X$, o que mostra que f é mensurável. ■

15.3 Produto de espaços mensuráveis

:⊢ **Definição 15.10.** Seja $(\boldsymbol{X_i})_{i\in I} = (X_i, \mathcal{M}_i)_{i\in I}$ uma família de espaços mensuráveis. O *produto* da família $(\boldsymbol{X}_i)_{i\in I}$ é o par

$$\prod_{i\in I} \boldsymbol{X_i} := (X, \mathcal{M})$$

em que $X := \prod_{i\in I} X_i$ é o produto de conjuntos e

$$\mathcal{M} := \left\langle \bigcup_{i\in I} \pi_i^*(\mathcal{M}_i) \right\rangle.$$

⊢ **Proposição 15.11.** *Seja $(\boldsymbol{X}_i)_{i\in I} = (X_i, \mathcal{M}_i)_{i\in I}$ uma família de espaços mensuráveis. Então o produto $\prod_{i\in I} \boldsymbol{X}_i$ é um espaço mensurável.*

□ *Demonstração.* Sejam $X := \prod_{i\in I} X_i$ e $\mathcal{M} = \langle \bigcup_{i\in I} \pi_i^*(\mathcal{M}_i)\rangle$. Devemos somente argumentar que $\mathcal{M}$ é uma sigma-álgebra sobre $X = \prod_{i\in I} X_i$. Para isso, notemos que, para cada $i \in I$, a sigma-álgebras $\pi_i^*(\mathcal{M}_i)$ é a sigma-álgebras puxada por $\pi_i : X \to X_i$, portanto uma sigma-álgebra sobre X. Assim, sendo, $\bigcup_{i\in I} \pi_i^*(\mathcal{M}_i) \subseteq \wp(X)$ e, portanto, a sigma-álgebra $\mathcal{M}$ gerada por esse conjunto é uma sigma-álgebra sobre X. ■

⊢ **Proposição 15.12.** *Seja $(\boldsymbol{X}_i)_{i\in I} = (X_i, \mathcal{M}_i)_{i\in I}$ uma família de espaços mensuráveis. Para todo $i \in I$, a projeção canônica $\pi_i : \prod_{i\in I} X_i \to X_i$ é uma função mensurável.*

□ *Demonstração.* Sejam $i \in I$ e $M \in \mathcal{M}_i$. Então $\pi_i^{-1}(M) \in \pi_i^*(\mathcal{M}_i)$ e, portanto, $\pi_i^{-1}(M) \in \mathcal{M}$. ■

⊢ **Proposição 15.13** (Propriedade Universal)**.** *Seja $(\boldsymbol{X}_i)_{i\in I} = (X_i, \mathcal{M}_i)_{i\in I}$ uma família de espaços mensuráveis, $\boldsymbol{Y} = (Y, \mathcal{M}_Y)$ um espaço mensurável e, para todo $i \in I$, $f_i : \boldsymbol{Y} \to \boldsymbol{X_i}$ uma função mensurável. Então existe uma única função mensurável $f : \boldsymbol{Y} \to \prod_{i\in I} \boldsymbol{X_i}$ tal que, para todo $i \in I$, $\pi_i \circ f = f_i$ (o diagrama comuta).*

$$\begin{tikzcd}
 & \prod_{i\in I} \boldsymbol{X_i} \arrow[d, "\pi_i"] \\
\boldsymbol{Y} \arrow[ur, dashed, "f"] \arrow[r, "f_i"'] & \boldsymbol{X_i}
\end{tikzcd}$$

□ *Demonstração.* Defina a função

$$\begin{aligned} f\colon Y &\longrightarrow \prod_{i\in I} X_i \\ y &\longmapsto (f_i(y))_{i\in I}. \end{aligned}$$

Da propriedade universal para o produto de conjuntos, f é a única função tal que, para todo $i \in I$, $\pi_i \circ f = f_i$. Basta mostrar que f é uma função mensurável. Para simplificar a notação, definamos $(X, \mathcal{M}) := \prod_{i\in I} \boldsymbol{X_i}$. Todo elemento de $\mathcal{M}$ é formado a partir de complementos e uniões de elementos de $\bigcup_{i\in I} \pi_i^*(\mathcal{M})$. Sendo assim, como f^{-1} preserva complemento e união, e $f^{-1}(\emptyset) = \emptyset$, se mostrarmos que, para todo $M \in \bigcup_{i\in I} \pi_i^*(\mathcal{M}_i)$, $f^{-1}(M) \in \mathcal{M}_Y$, seguirá que, para todo $M \in \mathcal{M}$, $f^{-1}(M) \in \mathcal{M}_Y$. Seja $M \in \bigcup_{i\in I} \pi_i^*(\mathcal{M}_i)$. Então existe $i \in I$ tal que $M \in \pi_i^*(\mathcal{M}_i)$ e, portanto, existe $M_i \in \mathcal{M}_i$ tal que $M = \pi_i^{-1}(M_i)$. Então segue que

$$f^{-1}(M) = f^{-1}(\pi_i^{-1}(M_i)) = (\pi_i \circ f)^{-1}(M_i) = f_i^{-1}(M_i)$$

e portanto, como f_i é mensurável, $f_i^{-1}(M_i) \in \mathcal{M}_Y$, portanto $f^{-1}(M) \in \mathcal{M}_Y$. Isso prova, pelos comentários anteriores, que para todo $M \in \mathcal{M}$, $f^{-1}(M) \in \mathcal{M}$ e, portanto, f é mensurável. ■

15.4 Espaços mensuráveis com estrutura adicional

15.4.1 Espaços mensuráveis topológicos

:⊢ **Definição 15.11.** Seja $\boldsymbol{X} = (X, \mathcal{T})$ um espaço topológico. A *σ-álgebra topológica* de $\boldsymbol{X}$ é

$$\mathcal{M}_{\mathcal{T}} := \langle \mathcal{T} \rangle ,$$

a σ-álgebra gerada pela topologia $\mathcal{T}$.

Essa σ-álgebra é comumente chamada de *σ-álgebra de Borel* e seus conjuntos mensuráveis de *borelianos* em homenagem ao matemático francês Émile Borel[1].

⊢ **Proposição 15.14.** *Sejam $\boldsymbol{X} = (X, \mathcal{T})$ um espaço topológico e $\mathcal{B} \subseteq \mathcal{T}$ uma base da topologia. Então*

$$\mathcal{M}_{\mathcal{T}} = \langle \mathcal{B} \rangle .$$

15.4.2 Funções mensuráveis com valores vetoriais

Tratamos brevemente de funções mensuráveis a valores vetoriais. De fato poderíamos somente considerar espaços lineares mensuráveis (nos quais as operações do espaço linear são mensuráveis, mas esse caso é mais geral do que o necessário e não será definido). Pode-se perceber, no entanto, que a demonstração seria a mesma.

⊢ **Proposição 15.15.** *Sejam $\boldsymbol{X}$ um espaço mensurável e $\boldsymbol{L}$ um espaço linear topológico sobre um corpo $\boldsymbol{C}$. O espaço $\mathscr{M}(\boldsymbol{X}, \boldsymbol{L})$ de funções mensuráveis de $\boldsymbol{X}$ para $(L, \mathcal{M}_{\mathcal{T}})$ é um subespaço linear de $\boldsymbol{L}^X$.*

□ *Demonstração.* Sejam $c \in C$ e $f, f' \in \mathscr{M}(X, L)$. Como as operações $+$ e $\cdot$ do espaço linear são mensuráveis, pois são contínuas, e f, f' e a função constante c são mensuráveis, segue que cf e $f + f'$ são mensuráveis. ■

15.4.3 Funções mensuráveis com valores em espaços métricos

⊢ **Proposição 15.16.** *Sejam $\boldsymbol{X} = (X, \mathcal{M}_X)$ um espaço mensurável, $\boldsymbol{Y} = (Y, |\cdot, \cdot|)$ um espaço métrico e $(f_n)_{n \in \mathbb{N}}$ uma sequência de funções mensuráveis de $(X, \mathcal{M}_X)$ para $(Y, \mathcal{M}_{\mathcal{T}})$ que converge pontualmente para $f \colon X \to Y$. Então f é mensurável.*

[1] Félix Édouard Justin Émile Borel (07/01/1871 – 03/02/1956)

□ *Demonstração.* Como a σ-álgebra topológica de Y é gerada por $\mathfrak{T}$, basta mostrar que f puxa abertos para mensuráveis. Primeiro notemos que, se $A \subseteq Y$ é um conjunto aberto, então, para todo $x \in f^{-1}(A)$, existe $N \in \mathbb{N}$ tal que, para todo $n \geq N$, $x \in f_n^{-1}(A)$. Portanto, para todo $m \in \mathbb{N}$,

$$f^{-1}(A) \subseteq \bigcup_{n=m}^{\infty} {f_n}^{-1}(A)$$

e, consequentemente,

$$f^{-1}(A) \subseteq \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} {f_n}^{-1}(A) = \overline{\lim_{n\in\mathbb{N}}}\, {f_n}^{-1}(A).$$

Segundo, seja $F \subseteq Y$ um conjunto fechado. Suponha que, para todo $m \in \mathbb{N}$, $x \in \bigcup_{n=m}^{\infty} {f_n}^{-1}(F)$. Então existe $N \in \mathbb{N}$ tal que, para todo $n \geq N$, $f_n(x) \in F$, portanto $f(x) \in F$ pois F é fechado. Segue então a inclusão contrária

$$\overline{\lim_{n\in\mathbb{N}}}\, {f_n}^{-1}(F) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} {f_n}^{-1}(F) \subseteq f^{-1}(F).$$

Finalmente, seja agora $A \subseteq Y$ um conjunto aberto, e defina, para todo $n \in \mathbb{N}$, os conjuntos abertos

$$A_n := \left\{ y \in Y \;\middle|\; \left| y, A^{\complement} \right| > n^{-1} \right\}.$$

e os conjuntos fechados

$$F_n := \left\{ y \in Y \;\middle|\; \left| y, A^{\complement} \right| \geq n^{-1} \right\},$$

Claramente $A_n \subseteq F_n$ e

$$A = \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} F_n.$$

Temos então as inclusões

$$f^{-1}(A) = \bigcup_{N\in\mathbb{N}} f^{-1}(F_N) \supseteq \bigcup_{N\in\mathbb{N}} \overline{\lim_{n\in\mathbb{N}}}\, {f_n}^{-1}(F_N) \supseteq \bigcup_{N\in\mathbb{N}} \overline{\lim_{n\in\mathbb{N}}}\, {f_n}^{-1}(A_N)$$

e

$$f^{-1}(A) = \bigcup_{N\in\mathbb{N}} f^{-1}(A_N) \subseteq \bigcup_{N\in\mathbb{N}} \overline{\lim_{n\in\mathbb{N}}}\, {f_n}^{-1}(A_N),$$

logo $f^{-1}(A) = \bigcup_{n\in\mathbb{N}} \overline{\lim}_{n\in\mathbb{N}}\, {f_n}^{-1}(A_N)$, e como f_n são mensuráveis e A_N são abertos, $f^{-1}(A)$ é mensurável. ■

15.5 Medida e espaço de medida

Comentário sobre o Monoide de Números Reais Positivos com Infinito Adotaremos nesta seção as definições de que, em $[0,\infty]$, a adição é definida por

$$\begin{aligned} +\colon [0,\infty]\times[0,\infty] &\longrightarrow [0,\infty] \\ (c,c') &\longmapsto \begin{cases} c+c, & c \neq \infty \ \text{ e } \ c \neq \infty \\ \infty, & c=\infty \ \text{ ou } \ c'=\infty \end{cases} \end{aligned}$$

e a multiplicação é definida por

$$\begin{aligned} \times\colon [0,\infty]\times[0,\infty] &\longrightarrow [0,\infty] \\ (c,c') &\longmapsto \begin{cases} 0, & c=0 \ \text{ ou } \ c'=0 \\ c\times c, & c\in\,]0,\infty[\ \text{ e } \ c'\in\,]0,\infty[\\ \infty, & (c=\infty \ \text{ e } \ c'\neq 0) \ \text{ ou } \ (c\neq 0 \ \text{ e } \ c'=\infty). \end{cases} \end{aligned}$$

15.5.1 Medidas

Isso faz de $[0,\infty]$ um monoide com operação binária de adição $+$ e identidade 0 e um monoide com operação binária de multiplicação $+$ e identidade 1.

:⊢ **Definição 15.12.** Seja $\boldsymbol{X}=(X,\mathcal{M})$ um espaço mensurável. Uma *medida* sobre $\boldsymbol{X}$ é uma função $m\colon \mathcal{M}\to[0,\infty]$ que satisfaz

1. $m(\emptyset)=0$;
2. Para toda sequência $(M_i)_{i\in\mathbb{N}}$ de conjuntos mensuráveis disjuntos aos pares,

$$m\left(\bigcup_{i\in\mathbb{N}} M_i\right) = \mathop{+}_{i\in\mathbb{N}} m(M_i).$$

:⊢ **Definição 15.13.** Sejam X um conjunto e $\wp(X)$ a σ-álgebra das partes de X. A *medida de contagem* sobre $(X,\wp(X))$ é a função

$$\begin{aligned} \#\colon \wp(X) &\longrightarrow [0,\infty] \\ C &\longmapsto \begin{cases} \#(M), & \#(M)<\#(\mathbb{N}) \\ \infty, & \#(M)\geq\#(\mathbb{N}). \end{cases} \end{aligned}$$

:⊢ **Definição 15.14.** Sejam $\boldsymbol{X}$ um espaço mensurável e $x\in X$. A *medida atômica*[2]

[2]Essa medida é conhecida como 'medida de Dirac', em homenagem ao físico britânico *Paul Dirac* (08/08/1902 – 20/10/1984). https://en.wikipedia.org/wiki/Dirac_measure.

sobre $\boldsymbol{X}$ em x é a função

$$\begin{aligned} \mathbf{1}|_x \colon \mathcal{M} &\longrightarrow [0,\infty] \\ M &\longmapsto \begin{cases} 1, & x \in M \\ 0, & x \notin M. \end{cases} \end{aligned}$$

Em geral, a notação adotada é δ_x. Note que $\mathbf{1}|_x(M) = \mathbf{1}_M(x)$, portanto adotamos essa notação.

► **Exemplo 15.3.** *Prove que as medida de contagem e atômica são medidas.*

:⊢ **Definição 15.15.** Um *espaço de medida* é uma tripla $(X, \mathcal{M}, m)$ em que $(X, \mathcal{M})$ é um espaço mensurável e m é uma medida sobre $(X, \mathcal{M})$.

⊢ **Proposição 15.17.** *Sejam $(X, \mathcal{M}, m)$ um espaço de medida e $M_1, M_2 \in \mathcal{M}$ conjuntos mensuráveis tais que $M_1 \subseteq M_2$. Então*

1. $m(M_1) \leq m(M_2)$;
2. $m(M_1) < +\infty \quad \Longrightarrow \quad m(M_2 \smallsetminus M_1) = m(M_2) - m(M_1)$.

□ *Demonstração.* Como $M_2 = M_1 \cup (M_2 \smallsetminus M_1)$ e $M_1 \cap (M_2 \smallsetminus M_1) = \emptyset$, segue que

$$m(M_2) = m(M_1) + m(M_2 \smallsetminus M_1).$$

1. Daí, como $m(M_2 \smallsetminus M_1) \geq 0$, segue que $m(M_2) \geq m(M_1)$.
2. Se $m(M_1) < +\infty$, então, subtraindo-a dos dois lados da equação, temos $m(M_2) - m(M_1) = m(M_2 \smallsetminus M_1)$.

■

⊢ **Proposição 15.18.** *Sejam $(X, \mathcal{M}, m)$ um espaço de medida e $(M_n)_{n\in\mathbb{N}}$ uma sequência de conjuntos mensuráveis.*

1. *Se $(M_n)_{n\in\mathbb{N}}$ é crescente, então*

$$m\left(\bigcup_{n\in\mathbb{N}} M_n\right) = \lim_{n\to+\infty} m(M_n);$$

2. *Se $(M_n)_{n\in\mathbb{N}}$ é decrescente e, para algum $n_0 \in \mathbb{N}$ vale $m(M_{n_0}) < +\infty$, então*

$$m\left(\bigcap_{n\in\mathbb{N}} M_n\right) = \lim_{n\to+\infty} m(M_n).$$

□ *Demonstração.* 1. Se para algum $n \in \mathbb{N}$ vale $m(M_n) = \infty$, então ambos lados da equação serão ∞. Resta provar a equação para o caso em que, para todo $n \in \mathbb{N}$, $m(M_n) < \infty$. Definamos $M'_0 := M_0$ e, para todo $n \in \mathbb{N} \smallsetminus \{0\}$,

$$M'_n := M_n \smallsetminus M_{n-1}.$$

A sequência $(M'_n)_{n\in\mathbb{N}}$ é disjunta aos pares, pois $(M_n)_{n\in\mathbb{N}}$ é crescente, e vale

$$\bigcup_{n\in\mathbb{N}} M'_n = \bigcup_{n\in\mathbb{N}} M_n.$$

Assim, segue da σ-aditividade de m que

$$m\left(\bigcup_{n\in\mathbb{N}} M_n\right) = m\left(\bigcup_{n\in\mathbb{N}} M'_n\right) = \mathop{+}_{n\in\mathbb{N}} m(M'_n) = \lim_{n\to\infty} \mathop{+}_{n'=0}^{n} m(M'_{n'}).$$

Pela proposição 15.17, segue que $m(M'_{n'}) = m(M_{n'}) - m(M_{n'-1})$, pois $m(M_{n'}) < \infty$, portanto

$$\mathop{+}_{n'=0}^{n} m(M'_{n'}) = \mathop{+}_{n'=0}^{n} (m(M_{n'}) - m(M_{n'-1})) = m(M_n),$$

logo

$$m\left(\bigcup_{n\in\mathbb{N}} M_n\right) = \lim_{n\to\infty} m(M_n).$$

2. Seja $n_0 \in \mathbb{N}$ tal que $m(M_{n_0}) < +\infty$ e definamos, para todo $n \in \mathbb{N}$, $M'_n := M_{n_0} \smallsetminus M_n$. Isso implica que para $n < n_0$ vale $M'_n = \emptyset$ e que a sequência $(M'_n)_{n\in\mathbb{N}}$ é crescente, portanto segue do item anterior que

$$\begin{aligned} m\left(\bigcup_{n\in\mathbb{N}} M'_n\right) &= \lim_{n\to\infty} m(M'_n) \\ &= \lim_{n\to\infty} (m(M_{n_0}) - m(M_n)) \\ &= m(M_{n_0}) - \lim_{n\to\infty} m(M_n). \end{aligned}$$

Mas $\bigcup_{n\in\mathbb{N}} M'_n = M_{n_0} \smallsetminus (\bigcap_{n\in\mathbb{N}} M_n)$, portanto

$$m\left(\bigcup_{n\in\mathbb{N}} M'_n\right) = m(M_{n_0}) - m\left(\bigcap_{n\in\mathbb{N}} M_n\right).$$

Assim, concluímos que

$$m\left(\bigcap_{n\in\mathbb{N}} M_n\right) = \lim_{n\to\infty} m(M_n).$$

■

15.5.2 Medida exterior

O conceito de medida exterior é um conceito mais abrangente que nos permite atribuir a cada subconjunto de um conjunto X um número real positivo. No entanto, nem sempre essa função será uma contavelmente aditiva, o que significa que ela nem sempre é uma medida. Mostraremos um jeito de restringir essa função a subconjuntos propícios de modo a termos uma sigma-álgebra de mensuráveis restrita à qual a medida exterior é uma medida. As medidas exteriores recebem esse nome exatamente porque medem conjuntos fora dessa sigma-álgebra, conjuntos não-mensuráveis.

:⊢ **Definição 15.16.** Seja X um conjunto. Uma *medida exterior* sobre $\boldsymbol{X}$ é uma função $m\colon \wp(X) \to [0, \infty]$ que satisfaz

1. $m(\emptyset) = 0$;
2. Para todos $C, C' \subseteq X$ tais que $C \subset C'$,
$$m(C) \leq m(C');$$
3. Para toda sequência $(C_i)_{i\in\mathbb{N}}$ de subconjuntos de X,
$$m\left(\bigcup_{i\in\mathbb{N}} C_i\right) \leq \mathop{+}_{i\in\mathbb{N}} m(M_i).$$

:⊢ **Definição 15.17.** Sejam X um conjunto e $m\colon \wp(X) \to [0, \infty]$ uma medida exterior sobre X. Um conjunto *mensurável* por m (ou *m-mensurável*) é um conjunto $M \subseteq X$ tal que, para todo conjunto $C \subseteq X$,

$$m(C) = m(C \cap M) + m(C \cap M^{\complement}).$$

O conjunto de conjuntos mensuráveis por m é denotado $\mathcal{M}_m$.

⊢ **Proposição 15.19.** *Sejam X um conjunto e $m\colon \wp(X) \to [0, \infty]$ uma medida exterior sobre X. A tripla $(X, \mathcal{M}_m, m|_{\mathcal{M}_m})$ é um espaço de medida.*

Esse método nos permite traduzir o problema de construir medida ao problema de construir medidas exteriores. Agora, vamos delinear um método de construir medidas exteriores em X usando somente um conjunto de subconjuntos de X em que uma função real positiva está definida. Essa medida é construída a partir de coberturas e por isso é chamada de medida exterior de cobertura.

:⊢ **Definição 15.18.** Seja X um conjunto. Um *pré-sistema de medida* sobre X é um par $(\mathcal{E}, \rho)$, em que $\mathcal{E} \subseteq \wp(X)$, $\emptyset \in \mathcal{E}$ e $\rho\colon \mathcal{E} \to [0, \infty]$ é uma função tal que $\rho(\emptyset) = 0$.

Na definição seguinte, consideramos $\inf(\emptyset) = \infty$, pois o ínfimo é considerado no conjunto $[0, \infty]$.

:⊢ **Definição 15.19.** Sejam X um conjunto e $(\mathcal{E}, \rho)$ um pré-sistema de medida sobre X. A *medida exterior de cobertura* induzida por $(\mathcal{E}, \rho)$ é a função

$$\begin{aligned} m^{(\mathcal{E},\rho)} \colon \wp(X) &\longrightarrow [0, \infty] \\ C &\longmapsto \inf \left\{ \mathop{+}_{i \in \mathbb{N}} \rho(C_i) \;\middle|\; \{C_i\}_{i \in \mathbb{N}} \subseteq \mathcal{E},\ C \subseteq \bigcup_{i \in \mathbb{N}} C_i \right\}. \end{aligned}$$

⊢ **Proposição 15.20.** *Sejam X um conjunto e $(\mathcal{E}, \rho)$ um pré-sistema de medida sobre X. A medida exterior de cobertura $m^{(\mathcal{E},\rho)}$ induzida por $(\mathcal{E}, \rho)$ é a uma medida exterior sobre X.*

□ *Demonstração.*

- (Conjunto vazio) $m^{(\mathcal{E},\rho)}(\emptyset) = 0$, pois se tomamos a cobertura vazia $(\emptyset)_{i \in \mathbb{N}}$, temos que $\emptyset \subseteq \bigcup_{i \in \mathbb{N}} \emptyset$ e $\rho(\emptyset) = 0$;
- (Monotonicidade crescente) Para todos $C, C' \subseteq M$ tais que $C \subseteq C'$, temos que uma cobertura de C' por conjuntos de $\mathcal{E}$ é uma cobertura de C por conjuntos de $\mathcal{E}$, logo $m^{(\mathcal{E},\rho)}(C) \leq m^{(\mathcal{E},\rho)}(C')$;
- (Subaditividade contável) Seja $(C_i)_{i \in \mathbb{N}}$ uma sequência de subconjuntos de M. Para todos $i \in \mathbb{N}$ e $\varepsilon \in \,]0, \infty[$, seja $U^i = (U_{i,j})_{j \in \mathbb{N}}$ é uma cobertura de C_i por conjuntos de $\mathcal{E}$ tal que

$$\mathop{+}_{j \in \mathbb{N}} \rho(U_{i,j}) \leq m^{(\mathcal{E},\rho)}(C_i) + \frac{\varepsilon}{2^{i+1}}.$$

Essa cobertura existe porque $m^{(\mathcal{E},\rho)}(C_i)$ é um ínfimo. Então $(U_{i,j})_{(i,j) \in \mathbb{N}^2}$ é uma cobertura de $\bigcup_{i \in \mathbb{N}} C_i$ por conjuntos de $\mathcal{E}$ e segue que

$$\begin{aligned} m^{(\mathcal{E},\rho)}\left(\bigcup_{i \in \mathbb{N}} C_i\right) &\leq m^{(\mathcal{E},\rho)}\left(\bigcup_{(i,j) \in \mathbb{N}^2} U_{i,j}\right) \\ &\leq \mathop{+}_{(i,j) \in \mathbb{N}^2} \rho(U_{i,j}) \\ &\leq \mathop{+}_{i \in \mathbb{N}} \left(m^{(\mathcal{E},\rho)}(C_i) + \frac{\varepsilon}{2^{i+1}}\right) \\ &= \mathop{+}_{i \in \mathbb{N}} \left(m^{(\mathcal{E},\rho)}(C_i)\right) + \mathop{+}_{i \in \mathbb{N}} \frac{\varepsilon}{2^{i+1}} \\ &= \mathop{+}_{i \in \mathbb{N}} \left(m^{(\mathcal{E},\rho)}(C_i)\right) + \varepsilon. \end{aligned}$$

A primeira desigualdade vem da monotonicidade de $m^{(\mathcal{E},\rho)}$, a segunda de $m^{(\mathcal{E},\rho)}$ ser ínfimo, e a terceira vem da condição para as coberturas $(U_{i,j})_{j\in\mathbb{N}}$. Como isso vale para qualquer ε, segue que

$$m^{(\mathcal{E},\rho)}\left(\bigcup_{i\in\mathbb{N}} C_i\right) \leq \mathop{+}_{i\in\mathbb{N}} m^{(\mathcal{E},\rho)}(C_i). \qquad \blacksquare$$

$\vdash$ **Proposição 15.21.** *Sejam X um conjunto e $(\mathcal{E},\rho)$ e $(\mathcal{E}',\rho')$ pré-sistemas de medida tais que $\mathcal{E}' \subseteq \mathcal{E}$ e $\rho \leq \rho'$. Então $m^{(\mathcal{E},\rho)} \leq m^{(\mathcal{E}',\rho')}$.*

$\square$ *Demonstração.* Vamos mostrar que $m^{(\mathcal{E},\rho)} \leq m^{(\mathcal{E}',\rho)}$ e então que $m^{(\mathcal{E}',\rho)} \leq m^{(\mathcal{E}',\rho')}$. Seja $C \subseteq X$. Como toda cobertura de C por de $\mathcal{E}'$ é uma cobertura de C por conjuntos de $\mathcal{E}$. Isso implica que

$$\begin{aligned} m^{(\mathcal{E},\rho)}(C) &= \inf\left\{ \mathop{+}_{i\in\mathbb{N}} \rho(C_i) \;\middle|\; \{C_i\}_{i\in\mathbb{N}} \subseteq \mathcal{E},\ C \subseteq \bigcup_{i\in\mathbb{N}} C_i \right\} \\ &\leq \inf\left\{ \mathop{+}_{i\in\mathbb{N}} \rho(C_i) \;\middle|\; \{C_i\}_{i\in\mathbb{N}} \subseteq \mathcal{E}',\ C \subseteq \bigcup_{i\in\mathbb{N}} C_i \right\} \\ &= m^{(\mathcal{E}',\rho)}(C), \end{aligned}$$

o que mostra que $m^{(\mathcal{E},\rho)} \leq m^{(\mathcal{E}',\rho)}$.

Agora, como $\rho(C) \leq \rho'(C)$, segue que, para toda cobertura de C por conjuntos de $\mathcal{E}$,

$$\mathop{+}_{i\in\mathbb{N}} \rho(C_i) \leq \mathop{+}_{i\in\mathbb{N}} \rho'(C_i),$$

portanto

$$\begin{aligned} m^{(\mathcal{E}',\rho)}(C) &= \inf\left\{ \mathop{+}_{i\in\mathbb{N}} \rho(C_i) \;\middle|\; \{C_i\}_{i\in\mathbb{N}} \subseteq \mathcal{E}',\ C \subseteq \bigcup_{i\in\mathbb{N}} C_i \right\} \\ &\leq \inf\left\{ \mathop{+}_{i\in\mathbb{N}} \rho'(C_i) \;\middle|\; \{C_i\}_{i\in\mathbb{N}} \subseteq \mathcal{E}',\ C \subseteq \bigcup_{i\in\mathbb{N}} C_i \right\} \\ &= m^{(\mathcal{E}',\rho')}(C), \end{aligned}$$

o que mostra que $m^{(\mathcal{E}',\rho)} \leq m^{(\mathcal{E}',\rho')}$ e, finalmente, que

$$m^{(\mathcal{E},\rho)} \leq m^{(\mathcal{E}',\rho')}.$$

$\blacksquare$

15.6 Medida em espaços topológicos

Consideraremos nesta seção medidas em espaços topológicos. No entanto, em geral supomos que o espaço topológico é separado, isto é, T_2, pois isso garante um mínimo de propriedades de separação entre os pontos do espaço, a separação de pontos por vizinhanças abertas.

Consideraremos, em geral, σ-álgebras que contêm a topologia do espaço, o que é equivalente a dizer que contêm a σ-álgebra topológica do espaço, que é, por definição, a menor σ-álgebra que contém sua topologia.

15.6.1 Medidas regulares

:⊢ **Definição 15.20.** Sejam $\boldsymbol{X} = (X, \mathcal{T})$ um espaço topológico separado, $\mathcal{M}$ uma σ-álgebra sobre X que contém $\mathcal{T}$ e m uma medida sobre $(X, \mathcal{M})$. Um conjunto *interiormente regular* com respeito a m é um conjunto mensurável $M \subseteq X$ tal que

$$m(M) = \sup \{m(C) \mid C \subseteq M,\ C \text{ é compacto}\}.$$

Uma medida *interiormente regular* sobre $(X, \mathcal{M})$ é uma medida m sobre $(X, \mathcal{M})$ para a qual todo conjunto mensurável de X é interiormente regular com respeito a m.

Um conjunto *exteriormente regular* com respeito a m é um conjunto mensurável $M \subseteq X$ tal que

$$m(M) = \sup \{m(C) \mid C \subseteq M, M \in \mathcal{T}\}.$$

Uma medida *exteriormente regular* sobre $(X, \mathcal{M})$ é uma medida m sobre $(X, \mathcal{M})$ para a qual todo conjunto mensurável de X é exteriormente regular com respeito a m.

Um conjunto *regular* com respeito a m é um conjunto que é interior e exteriormente regular com respeito a m. Uma medida *regular* sobre $(X, \mathcal{M})$ é uma medida m sobre $(X, \mathcal{M})$ que é interiormente e exteriormente regular.

15.6.2 Medidas localmente finitas

:⊢ **Definição 15.21.** Sejam $\boldsymbol{X} = (X, \mathcal{T})$ um espaço topológico separado e $\mathcal{M}$ uma σ-álgebra sobre X que contém $\mathcal{T}$. Uma medida *localmente finita* sobre $(X, \mathcal{M})$ é uma m sobre $(X, \mathcal{M})$ tal que, para todo $x \in X$, existe vizinhança aberta $A \subseteq X$ de x tal que $m(A) < \infty$.

15.7 Espaço linear de medidas

Podemos expandir o conceito de medida para obtermos mais estrutura algébrica entre as medidas. Vamos modificar o conceito de medida para admitir valores negativos, mas em contrapartida temos que impedir que a medida de um conjunto seja infinita. Para isso, podemos definir a medida como uma função assumindo valores em $\mathbb{R}$ mas, de modo mais geral, basta considerarmos valores em um corpo topológico, o que permite que tenhamos medidas com valores em $\mathbb{C}$, ou podemos até mesmo generalizar para espaços lineares topológicos. A estrutura linear é o que garante que o espaço de medidas será também um espaço linear e a topologia é o que garante que podemos falar de convergência de uma sequência de elementos do espaço linear, o que é necessário para que definamos a propriedade de σ-aditividade de uma medida.

:⊢ **Definição 15.22.** Sejam $\boldsymbol{X} = (X, \mathcal{M})$ um espaço mensurável e $\boldsymbol{L}$ um espaço linear topológico sobre um corpo topológico $\boldsymbol{C}$. Uma *medida* sobre $\boldsymbol{X}$ a valores em $\boldsymbol{L}$ é uma função $m\colon \mathcal{M} \to L$ que satisfaz:

1. $m(\emptyset) = 0$;
2. Para toda sequência $(M_i)_{i\in\mathbb{N}}$ de conjuntos mensuráveis disjuntos,

$$m\left(\bigcup_{i\in\mathbb{N}} M_i\right) = \mathop{+}_{i\in\mathbb{N}} m(M_i).$$

O conjunto das medidas sobre $\boldsymbol{X}$ a valores em $\boldsymbol{L}$ é denotado $\mathfrak{M}(\boldsymbol{X}, \boldsymbol{L})$. Caso não haja ambiguidade, denotamos $\mathfrak{M}$.

O conjunto $\mathfrak{M}(\boldsymbol{X})$ das medidas sobre $\boldsymbol{X}$ é um espaço de funções do conjunto $\mathcal{M}$ para o corpo (ou espaço linear) $\mathbb{R}$. Nesse sentido, podemos entender $\mathfrak{M}(\boldsymbol{X})$ como um subconjunto de $\mathbb{R}^{\mathcal{M}}$ e, de fato, esse conjunto forma um subespaço com respeito à adição e à multiplicação pontuais, bem como inversas e identidades de cada.

⊢ **Proposição 15.22.** *Seja $\boldsymbol{X}$ um espaço mensurável. O espaço $\mathfrak{M}(\boldsymbol{X})$ é um subespaço linear de $\mathbb{R}^{\mathcal{M}}$.*

□ *Demonstração.* Temos que mostrar que $\mathfrak{M}(\boldsymbol{X})$ é fechado pela adição e multiplicação. Sejam $c \in \mathbb{R}$ e $m, m' \in \mathfrak{M}(\boldsymbol{X})$. Então

1. $(cm + m')(\emptyset) = cm(\emptyset) + m'(\emptyset) = 0$;

2. Seja $(M_i)_{i\in\mathbb{N}}$ uma sequência de conjuntos mensuráveis disjuntos. Então

$$\begin{aligned}(cm+m')\left(\bigcup_{i\in\mathbb{N}} M_i\right) &= cm\left(\bigcup_{i\in\mathbb{N}} M_i\right) + m'\left(\bigcup_{i\in\mathbb{N}} M_i\right)\\ &= \mathop{+}_{i\in\mathbb{N}} cm(M_i) + \mathop{+}_{i\in\mathbb{N}} m'(M_i)\\ &= \mathop{+}_{i\in\mathbb{N}} (cm+m')(M_i).\end{aligned}$$

■

Quando $L=\mathbb{R}$, podemos quase recuperar a definição que tínhamos de medida finita considerando as *medidas positivas*. Esse é o conjunto das medida reais que só assumem valores em $\mathbb{R}_{\geq 0} = [0,\infty[$ e é denotado $\mathfrak{M}_{\geq 0}$. Note que $\mathfrak{M}_{\geq 0}$ é um cone em $\mathfrak{M}$, pois para todo $c \in \mathbb{R}_{\geq 0}$ e toda $m \in \mathfrak{M}_{\geq 0}$, temos $cm \in \mathfrak{M}_{\geq 0}$. Além disso, se $m \in \mathfrak{M}_{\geq 0} \cap -\mathfrak{M}_{\geq 0}$, então, para todo $M \in \mathcal{M}$, $m(M) \geq 0$ e $-m(M) \geq 0$, portanto $m = 0$, o que mostra que $\mathfrak{M}_{\geq 0}(\boldsymbol{X})$ é um cone agudo, ou seja, $\mathfrak{M}_{\geq 0}(\boldsymbol{X}) \cap -\mathfrak{M}_{\geq 0}(\boldsymbol{X}) = \{0\}$.

De modo mais geral, se temos um cone X em um espaço linear $\boldsymbol{L}$, então o espaço $\mathfrak{M}|_X$ é um cone e, se X é agudo, então $\mathfrak{M}$.

15.8 Quase

15.8.1 Quase todo

A estrutura de medida nos permite definir um novo quantificador, que formaliza o conceito de que quase todo ponto satisfaz alguma propriedade, ou seja, de que, a menos de pontos em um conjunto de medida nula, todos os pontos satisfazem tal propriedade.

:⊢ **Definição 15.23.** Seja $\boldsymbol{X}$ um espaço de medida. Uma propriedade P de elementos de X vale para *quase todo* ponto se existe um conjunto $M \in \mathcal{M}$ com $m(M) = 0$ tal que, para todo $x \in X \smallsetminus M$, vale a propriedade P:

$$\overset{\circ}{\forall} x \mathrm{P}x \equiv \exists M \forall x (M \in \mathcal{M} \wedge m(M) = 0 \wedge (x \in X \smallsetminus M \to \mathrm{P}x)).$$

A partir desse quantificador, podemos definir o quantificador *quase existe* da seguinte forma

$$\overset{\circ}{\exists} x \mathrm{P}x \equiv \neg \overset{\circ}{\forall} x \neg \mathrm{P}x$$

$$\begin{aligned}\neg\overset{\circ}{\forall}x\neg \mathrm{P}x &\equiv \neg\exists M\forall x(M \in \mathcal{M} \wedge m(M) = 0 \wedge (x \in X \smallsetminus M \to \neg \mathrm{P}x))\\ &\equiv \forall M\exists x(M \notin \mathcal{M} \vee m(M) > 0 \vee (x \in X \smallsetminus M \wedge \mathrm{P}x))\\ &\equiv \forall M\exists x(\neg(M \notin \mathcal{M} \vee m(M) > 0) \to (x \in X \smallsetminus M \wedge \mathrm{P}x))\\ &\equiv \forall M\exists x(M \in \mathcal{M} \wedge m(M) = 0 \to (x \in X \smallsetminus M \wedge \mathrm{P}x))\end{aligned}$$

pois

$$\neg(A \to B) \equiv A \wedge \neg B$$

$$A \vee B \equiv \neg A \to B$$

⊢ **Proposição 15.23.** *Sejam* $\boldsymbol{X}$ *um espaço de medida e* P *e* Q *propriedades de elementos de* X*. Então*

1. $\overset{\circ}{\forall}x\mathrm{P}x \wedge \overset{\circ}{\forall}x\mathrm{Q}x \to \overset{\circ}{\forall}x(\mathrm{P}x \wedge \mathrm{Q}x)$*;*
2. $\forall x\mathrm{P}x \to \overset{\circ}{\forall}x\mathrm{P}x$*;*
3. $\overset{\circ}{\exists}x\mathrm{P}x \to \exists x\mathrm{P}x$*;*
4. $\neg\overset{\circ}{\exists}x\neg \mathrm{P}x \equiv \overset{\circ}{\forall}x\mathrm{P}x$*;*

□ *Demonstração.*

1. Sejam $M, N \in \mathcal{M}$ os conjuntos de medida 0 em cujos complementares P e Q falham, respectivamente. Então $M \cup N \in \mathcal{M}$ tem medida 0, portanto P e Q falham no complementar de $M \cup N$ e segue o teorema.
2. Tomando $M = \emptyset$, temos $X \smallsetminus M = X$ e concluímo que para todo $x \in X \smallsetminus M$ vale $\mathrm{P}x$.
3. Tomando $M = \emptyset$, temos $X \smallsetminus M = X$ e concluímos que existe $x \in X$ tal que $\mathrm{P}x$.
4. Exercício.

■

15.8.2 Quase igualdade de conjuntos

:⊢ **Definição 15.24.** Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida e $A \in \mathcal{M}$. Um conjunto *quase contido* em A com respeito a $\boldsymbol{X}$ é um conjunto $B \in \mathcal{M}$ tal que $m(B \smallsetminus A) = 0$. Denota-se $B \overset{\circ}{\subseteq} A$.

Um conjunto *quase igual* a A com respeito a $\boldsymbol{X}$ é um conjunto $B \in \mathcal{M}$ tal que $B \overset{\circ}{\subseteq} A$ e $A \overset{\circ}{\subseteq} B$. Denota-se $B \overset{\circ}{=} A$. Conjuntos *quase vazios* em $\boldsymbol{X}$ são conjuntos quase iguais ao conjunto vazio ($A \overset{\circ}{=} \emptyset$) e conjuntos *quase totais* em $\boldsymbol{X}$ são conjuntos quase iguais a X ($A \overset{\circ}{=} X$).

⊢ **Proposição 15.24.** *Seja* $\boldsymbol{X}$ *um espaço de medida.*

1. *A relação* $\overset{\circ}{\subseteq}$ *em* $\mathcal{M}$ *é uma relação de ordem parcial (com respeito a* $\overset{\circ}{=}$*);*
2. *A relação* $\overset{\circ}{=}$ *em* $\mathcal{M}$ *é uma relação de equivalência.*

□ *Demonstração.* 1. (Reflexividade) Seja $A \in \mathcal{M}$. Então $A \overset{\circ}{\subseteq} A$, pois $m(A \setminus A) = m(\emptyset) = 0$. (Antissimetria) Sejam $A, B \in \mathcal{M}$ tais que $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} A$. Então $A \overset{\circ}{=} B$ por definição. (Transitividade) Sejam $A, B, C \in \mathcal{M}$ tais que $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} C$. Então $m(A \setminus B) = 0$ e $m(B \setminus C) = 0$. Mas $A \setminus C = (A \cap B \setminus C) \cup (A \setminus (B \cup C))$; como $(A \cap B \setminus C) \subseteq B \setminus C$ e $(A \setminus (B \cup C)) \subseteq A \setminus B$, segue que

$$m(A \setminus C) = m(A \cap B \setminus C) + m(A \setminus (B \cup C)) \leq 0,$$

logo $A \overset{\circ}{\subseteq} C$;

2. (Reflexividade) Seja $A \in \mathcal{M}$. Então $A \overset{\circ}{=} A$, pois $A \overset{\circ}{\subseteq} A$. (Simetria) Sejam $A, B \in \mathcal{M}$ tais que $A \overset{\circ}{=} B$. Então $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} A$, portanto $B \overset{\circ}{=} A$. (Transitividade) Sejam $A, B, C \in \mathcal{M}$ tais que $A \overset{\circ}{=} B$ e $B \overset{\circ}{=} C$. Então $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} C$, portanto $A \overset{\circ}{\subseteq} C$. Analogamente, $C \overset{\circ}{\subseteq} A$, portanto $A \overset{\circ}{=} C$.

■

⊢ **Proposição 15.25.** *Sejam* $\boldsymbol{X}$ *um espaço de medida e* $A, B \in \mathcal{M}$.

1. $A \overset{\circ}{=} B$ *se, e somente se,* $m(A \triangle B) = 0$*;*
2. *Se* $A \overset{\circ}{=} B$*, então* $m(A) = m(B)$*;*
3. *Se* $A \subseteq B$*, então* $A \overset{\circ}{=} B$ *implica* $m(A) = m(B)$*; se* $A \subseteq B$ *e* m *é finita, então* $m(A) = m(B)$ *implica* $A \overset{\circ}{=} B$*;*
4. $A \overset{\circ}{=} \emptyset$ *se, e somente se,* $m(A) = 0$*;*
5. $A \overset{\circ}{=} X$ *se, e somente se,* $m(A^{\complement}) = 0$*. Se* m *é finita,* $A \overset{\circ}{=} X$ *se, e somente se,* $m(A) = m(X)$*;*
6. $A \overset{\circ}{=} \emptyset$ *se, e somente se,* $A^{\complement} \overset{\circ}{=} X$*;*
7. *Se* $A \overset{\circ}{=} X$ *e* $B \overset{\circ}{=} X$*, então* $A \cap B \overset{\circ}{=} X$*.*
8. *Se* $A \overset{\circ}{=} X$ *e* m *é finita, então, para todo* $B \subseteq A$*, vale* $m(B) = m(A \cap B)$*;*

□ *Demonstração.* 1. Como $A \setminus B$ e $B \setminus A$ são disjuntos, temos $m(A \triangle B) = m(A \setminus B) + m(B \setminus A)$, portanto $m(A \setminus B) = m(B \setminus A) = 0$ se, e somente se, $m(A \triangle B) = 0$;

2. Como $A \setminus B$ e $B \setminus A$ são disjuntos, temos $m(A \triangle B) = m(A \setminus B) + m(B \setminus A)$, portanto $m(A \setminus B) = m(B \setminus A) = 0$. Como $A = (A \cap B) \cup (A \setminus B)$ e $A \cap B$ e $A \setminus B$ são disjuntos, $m(A) = m(A \cap B) + m(A \setminus B) = m(A \cap B)$. Analogamente $m(B) = m(B \cap A)$, portanto $m(A) = m(B)$.

3. Como $A \subseteq B$, então $A \cup B = A$ e $A \cap B = B$, logo de $m(A \triangle B) = 0$ segue

$$m(A) = m(A \cup B) = m(A \cap B) + m(A \triangle B) = m(B).$$

Se m é finita, da mesma igualdade segue que

$$m(A \triangle B) = m(A \cup B) - m(A \cap B) = m(A) - m(B) = 0.$$

4. A ida segue do item acima. Como $A \cup \emptyset = A$ e $A \cap \emptyset = \emptyset$, segue que $A \triangle \emptyset = A$, logo $m(A \triangle \emptyset) = m(A)$. Assim segue de $m(A) = 0$ que $A \mathrel{\mathring{=}} \emptyset$.
5. A ida segue do item acima e de $m(X) = m(A) + m(A^{\complement})$. Como $A \cup X = X$ e $A \cap X = A$, segue que $A \triangle X = A^{\complement}$, logo $m(A \triangle X) = m(A^{\complement})$. Assim segue de $m(A^{\complement}) = 0$ que $A \mathrel{\mathring{=}} X$. Se m é finita, $m(A^{\complement}) = m(X) - m(A)$. Assim segue de $m(A) = m(X)$ se, e somente se, $m(A^{\complement}) = 0$.
6. Consequência dos itens anteriores.
7. Como $A \mathrel{\mathring{=}} X$ e $B \mathrel{\mathring{=}} X$, então $m(A^{\complement}) = m(B^{\complement}) = 0$, portanto

$$m\left((A \cap B)^{\complement}\right) = m\left(A^{\complement} \cup B^{\complement}\right) \leq m\left(A^{\complement}\right) + m\left(B^{\complement}\right) = 0,$$

   o que mostra que $A \cap B \mathrel{\mathring{=}} X$.
8. Como $A \subseteq A \cup B$, então $m(A \cup B) = m(A \cap B) + m(A \triangle B) = 1$ e $m(A) = m(A \cap B) + m(A \smallsetminus B) = 1$, e, por m ser finita, igualando as expressões segue que

$$m(A \triangle B) = m(A \smallsetminus B).$$

   Como

$$m(A \triangle B) = m(A \smallsetminus B) = m(B \smallsetminus A),$$

   segue que $m(B \smallsetminus A) = 0$. Portanto, como

$$m(B) = m(B \cap A) + m(B \smallsetminus A)$$

   concluímos que $m(B) = m(B \cap A)$.

■

⊢ **Proposição 15.26.** *Seja* $\boldsymbol{X}$ *um espaço de medida.*

1. *Os conjuntos* $\{M \in \mathcal{M} \mid M \mathrel{\mathring{=}} \emptyset\}$ *e* $\{M \in \mathcal{M} \mid M \mathrel{\mathring{=}} X\}$ *são fechados sob união e interseção enumeráveis;*
2. *O conjunto* $\{M \in \mathcal{M} \mid M \mathrel{\mathring{=}} \emptyset \;\; ou \;\; M \mathrel{\mathring{=}} X\}$ *é uma sigma-álgebra.*

□ *Demonstração.* 1. Seja $(M_i)_{i \in \mathbb{N}}$ uma sequência de conjuntos quase vazios. Então

$$m\left(\bigcup_{i \in \mathbb{N}} M_i\right) \leq \mathop{+}_{i \in \mathbb{N}} m(M_i) = 0$$

e, para algum $i \in \mathbb{N}$

$$m\left(\bigcap_{i \in \mathbb{N}} M_i\right) \leq m(M_i) = 0.$$

Seja $(M_i)_{i\in\mathbb{N}}$ uma sequência de conjuntos quase totais. O mesmo vale já que $M \mathrel{\mathring{=}} X$ se, e somente, $M^{\complement} \mathrel{\mathring{=}} \emptyset$.

2. Segue do item anterior, do fato que $M \mathrel{\mathring{=}} X$ se, e somente, $M^{\complement} \mathrel{\mathring{=}} \emptyset$, e de $\emptyset$ ser quase vazio.

■

15.8.3 Quase igualdade de funções

Existem dois modos de generalizar o conceito de função a partir de uma estrutura de medida, de modo a considerar equivalentes funções que diferem em um conjunto de medida zero. O método padrão consiste em construir classes de equivalência de funções definidas no espaço todo. O outro é considerar funções que não estão definidas no domínio todo do espaço. Descreveremos a abordagem padrão e numa subseção posterior esboçaremos algumas construções da segunda abordagem.

:⊢ **Definição 15.25.** Sejam $\boldsymbol{X}$ espaço de medida e X' conjunto. Funções de $\boldsymbol{X}$ para X' *quase iguais* são funções (mensuráveis) $f\colon X \to X'$ e $f'\colon X \to X'$ tais que, para quase todo $x \in X$, $f(x) = f'(x)$. Denota-se $f \mathrel{\mathring{=}} f'$.

Explicitamente, a definição significa que existe $N \in \mathcal{M}$ tal que $N \mathrel{\mathring{=}} \emptyset$ e, para todo $x \in X \smallsetminus N$, $f(x) = f'(x)$; ou ainda, tal que $f|_{X\smallsetminus N} = f'|_{X\smallsetminus N}$.

⊢ **Proposição 15.27.** *Sejam $\boldsymbol{X}$ espaço de medida e X' conjunto. A relação $\mathring{=}$ de quase igualdade de funções é uma relação de equivalência no conjunto de funções de X para X'.*

□ *Demonstração.* Sejam $f, f', f''\colon X \to X'$. (Reflexividade) Claramente, $f \mathrel{\mathring{=}} f$. (Simetria) Claramente $f \mathrel{\mathring{=}} f'$ implica $f' \mathrel{\mathring{=}} f$. (Transitividade) Se $f \mathrel{\mathring{=}} f'$ e $f' \mathrel{\mathring{=}} f''$, então existem conjuntos N e N' quase vazios tais que $f|_{X\smallsetminus N} = f'|_{X\smallsetminus N}$ e $f'|_{X\smallsetminus N'} = f'|_{X\smallsetminus N}$. Então, como $N \cap N'$ é quase vazio, segue que $f|_{X\smallsetminus(N\cap N')} = f'|_{X\smallsetminus(N\cap N')} = f''|_{X\smallsetminus(N\cap N')}$, portanto $f \mathrel{\mathring{=}} f''$. ■

Embora não seja necessário na definição que as funções sejam mensuráveis, consideraremos somente funções mensuráveis. Lembremos que, dados $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços de medida, $\mathscr{M}(\boldsymbol{X}, \boldsymbol{X}')$ é o conjunto de funções mensuráveis de $\boldsymbol{X}$ para $\boldsymbol{X}'$. Como $\mathring{=}$ é uma equivalência, podemos quocientar esse espaço pela equivalência, obtendo o conjunto de classes de equivalência de funções mensuráveis

:⊢ **Definição 15.26.** Sejam $\boldsymbol{X}$ e $\boldsymbol{X}'$ espaços de medida. Uma *quase função* é uma classe de equivalência

$$[f] = \{f' \in \mathscr{M}(\boldsymbol{X}, \boldsymbol{X}') \mid f' \mathrel{\mathring{=}} f\}\,.$$

O conjunto de *quase funções* de $\boldsymbol{X}$ para $\boldsymbol{X}'$ é o conjunto quociente

$$\mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{X}') := \mathscr{M}(\boldsymbol{X}, \boldsymbol{X}')\big/_{\mathring{=}}\,.$$

Quando conveniente, denotaremos uma quase função $[f]$ cujo representante é f por f, para simplificar a notação.

:⊢ **Definição 15.27.** Sejam $\boldsymbol{X}$, $\boldsymbol{X}'$ e $\boldsymbol{X}''$ espaços de medida. A *composição* de quase funções é definida como

$$\begin{aligned} \circ\colon \mathscr{M}_{\mathring{=}}(\boldsymbol{X}', \boldsymbol{X}'') \times \mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{X}') &\longrightarrow \mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{X}'') \\ ([f'], [f]) &\longmapsto [f' \circ f]. \end{aligned}$$

⊢ **Proposição 15.28.** *Sejam $\boldsymbol{X}$, $\boldsymbol{X}'$, $\boldsymbol{X}''$ e $\boldsymbol{X}'''$ espaços de medida.*

1. *Para toda $[f] \in \mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{X}')$,*

$$[\mathrm{I}_{X'}] \circ [f] = [f] = [f] \circ [\mathrm{I}_X];$$

2. *Para todas $[f] \in \mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{X}')$, $[f'] \in \mathscr{M}_{\mathring{=}}(\boldsymbol{X}', \boldsymbol{X}'')$ e $[f''] \in \mathscr{M}_{\mathring{=}}(\boldsymbol{X}'', \boldsymbol{X}''')$,*

$$[f''] \circ ([f'] \circ [f]) = ([f''] \circ [f']) \circ [f].$$

15.8.3.1 Abordagem alternativa

Nesta subseção, definimos outros objetos com o nome de quase funções. Eles têm uma relação direta com as quase funções da seção anterior, mas fora desta seção, uma quase função sempre será entendida como uma classe de equivalência de funções que diferem somente em um conjunto de medida zero, como na seção anterior.

:⊢ **Definição 15.28.** Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços de medida. Uma *quase função* de $\boldsymbol{X}_1$ para $\boldsymbol{X}_2$ é uma função $f\colon C \to X_2$ tal que $C \mathrel{\mathring{=}} X_1$. Denota-se $f\colon \boldsymbol{X}_1 \overset{\circ}{\to} \boldsymbol{X}_2$.

:⊢ **Definição 15.29.** Sejam X_1, C_1, X_2, C_2 e X_3 conjuntos tais que $C_1 \subseteq X_1$ e $C_2 \subseteq X_2$, e $f_1\colon C_1 \to X_2$ e $f_2\colon C_2 \to X_3$ funções. A *composição* (generalizada) de f_1 com f_2 é a função

$$\begin{aligned} f_2 \circ f_1\colon C_1 \cap {f_1}^{-1}(C_2) &\longrightarrow X_3 \\ x &\longmapsto f_2(f_1(x)). \end{aligned}$$

⊢ **Proposição 15.29** (Composição de quase função)**.** *Sejam $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ e $\boldsymbol{X}_3$ espaços de medida e $f_1\colon \boldsymbol{X}_1 \overset{\circ}{\to} \boldsymbol{X}_2$ e $f_2\colon \boldsymbol{X}_2 \overset{\circ}{\to} \boldsymbol{X}_3$ quase funções, com respectivos domínios C_1 e C_2, que preservam medida. Então $f_2 \circ f_1\colon \boldsymbol{X}_1 \overset{\circ}{\to} \boldsymbol{X}_3$ é uma quase função, com domínio $C_1 \cap {f_1}^{-1}(C_2)$, que preserva medida.*

□ *Demonstração.* Como f_1 e f_2 preservam medida, $f_2 \circ f_1$ preserva medida. Basta mostrar que seu domínio é quase total. Como f_1 é quase função, $m_1({C_1}^{\complement}) = 0$, e como f_2 é quase função, $m_2({C_2}^{\complement}) = 0$. Assim, como f_1 preserva medida, $m_1({f_1}^{-1}((C_2)^{\complement})) = m_2((C_2)^{\complement}) = 0$. Por fim segue que

$$\begin{aligned} m_1\left(\left(C_1 \cap {f_1}^{-1}(C_2)\right)^{\complement}\right) &= m_1\left({C_1}^{\complement} \cup {f_1}^{-1}({C_2}^{\complement})\right) \\ &\leq m_1\left({C_1}^{\complement}\right) + m_1\left({f_1}^{-1}({C_2}^{\complement})\right) \\ &= 0 + 0 = 0, \end{aligned}$$

portanto $C_1 \cap {f_1}^{-1}(C_2) \mathrel{\mathring{=}} X_1$, e concluímos que $f_2 \circ f_1$ é uma quase função. ■

⊢ **Proposição 15.30** (Associatividade)**.** *Sejam $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ e $\boldsymbol{X}_4$ espaços de medida e $f_1\colon \boldsymbol{X}_1 \mathrel{\mathring{\to}} \boldsymbol{X}_2$, $f_2\colon \boldsymbol{X}_2 \mathrel{\mathring{\to}} \boldsymbol{X}_3$ e $f_3\colon \boldsymbol{X}_3 \mathrel{\mathring{\to}} \boldsymbol{X}_4$ quase funções, com respectivos domínios C_1, C_2 e C_3, que preservam medida. Então*

$$f_3 \circ (f_2 \circ f_1) = (f_3 \circ f_2) \circ f_1.$$

□ *Demonstração.* O domínio de $f_3 \circ (f_2 \circ f_1)$ é o conjunto

$$(C_1 \cap {f_1}^{-1}(C_2)) \cap (f_2 \circ f_1)^{-1}(C_3)$$

e o domínio de $(f_3 \circ f_2) \circ f_1$ é o conjunto

$$C_1 \cap {f_1}^{-1}(C_2 \cap {f_2}^{-1}(C_3)).$$

Vamos mostrar que esses conjuntos são iguais.

$$\begin{aligned} (C_1 \cap {f_1}^{-1}(C_2)) \cap (f_2 \circ f_1)^{-1}(C_3) &= C_1 \cap {f_1}^{-1}(C_2) \cap (f_2 \circ f_1)^{-1}(C_3) \\ &= C_1 \cap {f_1}^{-1}(C_2) \cap {f_1}^{-1}({f_2}^{-1}(C_3)) \\ &= C_1 \cap {f_1}^{-1}(C_2 \cap {f_2}^{-1}(C_3)). \end{aligned}$$

■

:⊢ **Definição 15.30.** Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços de medida. Quase funções de $\boldsymbol{X}_1$ para $\boldsymbol{X}_2$ *quase iguais* são quase funções $f_1\colon \boldsymbol{X}_1 \mathrel{\mathring{\to}} \boldsymbol{X}_2$ e $f_2\colon \boldsymbol{X}_1 \mathrel{\mathring{\to}} \boldsymbol{X}_2$ satisfazendo $f_1|_C = f_2|_C$ para algum conjunto $C \mathrel{\mathring{=}} X_1$. Denota-se $f_1 \mathrel{\mathring{=}} f_2$.

⊢ **Proposição 15.31.** *A relação $\mathring{=}$ de quase igualdade de quase funções é uma relação de equivalência.*

□ *Demonstração.* Sejam $f, f', f''\colon X \to X'$. (Reflexividade) Claramente, $f \mathrel{\mathring{=}} f$. (Simetria) Claramente $f \mathrel{\mathring{=}} f'$ implica $f' \mathrel{\mathring{=}} f$. (Transitividade) Se $f \mathrel{\mathring{=}} f'$ e $f' \mathrel{\mathring{=}} f''$, então existem conjuntos C e C' quase totais tais que $f|_C = f'|_C$ e $f'|_{C'} = f'|_{C'}$. Então, como $C \cap C'$ é quase total, segue que $f|_{C\cap C'} = f'|_{C\cap C'} = f''|_{C\cap C'}$, portanto $f \mathrel{\mathring{=}} f''$. ■

15.9 Integração

15.9.1 Integral de funções mensuráveis simples

Lembremos que a função indicadora em um conjunto X é a função

$$\begin{aligned} \mathbf{1}\colon \wp(X) &\longrightarrow 2^X \\ C &\longmapsto \mathbf{1}_C\colon X \longrightarrow \{0,1\} \\ & \qquad\qquad x \longmapsto \begin{cases} 1, & x \in C \\ 0, & x \notin C. \end{cases} \end{aligned}$$

Na proposição a seguir mostramos que funções indicadoras são mensuráveis se, e somente se, o conjunto que elas indicam são. Para fazer sentido uma função indicadora ser mensurável, precisamos dar uma estrutura mensurável para $\{0,1\}$, e a estrutura que escolhemos é a álgebra discreta. Essa é a álgebra induzida se consideramos $\{0,1\}$ como subconjunto de $\mathbb{R}$, logo a função indicadora será mensurável também como uma função para $\mathbb{R}$.

$\vdash$ **Proposição 15.32.** *Seja $(X,\mathcal{M})$ um espaço mensurável. Um conjunto $M \subseteq X$ é mensurável se, e somente se, $\mathbf{1}_M \in \mathscr{M}(\boldsymbol{X},\{0,1\})$ é mensurável.*

$\square$ *Demonstração.* Basta notar que $\mathbf{1}_M^{-1}(\{1\}) = M$ e $\mathbf{1}_M^{-1}(\{0\}) = M^{\complement}$. Se M é mensurável, então $M^{\complement}$ é mensurável, logo $\mathbf{1}_M$ também é. Reciprocamente, se $\mathbf{1}_M$ é mensurável, então $M = \mathbf{1}_M^{-1}(\{1\})$ é mensurável, pois $\{1\}$ é mensurável. ■

Usaremos funções indicadoras na teoria de integração. Elas permitem cancelar funções $f\colon X \to \mathbb{R}$ em um conjunto C se multiplicamos f por $\mathbf{1}_C$ (com a multiplicação definida pontualmente). Nesse caso, temos a função

$$\begin{aligned} \mathbf{1}_C f\colon X &\longrightarrow \mathbb{R} \\ x &\longmapsto \mathbf{1}_C(x)f(x) = \begin{cases} f(x), & x \in C \\ 0, & x \notin C. \end{cases} \end{aligned}$$

Consideraremos primeiro funções que têm um número finito de valores. Essas funções são chamadas simples.

$:\vdash$ **Definição 15.31.** Seja $\boldsymbol{X} = (X,\mathcal{M},m)$ um espaço de medida. Uma função *simples* em X é uma função $f\colon X \to \mathbb{R}$ tal que $f(X)$ é finito. A *partição por níveis* de f é o conjunto

$$\mathcal{P}_f := \{f^{-1}(c)\}_{c\in f(X)}.$$

O conjunto das funções simples mensuráveis de X para $\mathbb{R}$ é denotado $\mathscr{M}_s(\boldsymbol{X},\mathbb{R})$.

$\vdash$ **Proposição 15.33.** *Sejam $\boldsymbol{X}$ um espaço de medida e f, $f' \in \mathscr{M}(\boldsymbol{X}, [0, \infty])$ funções simples mensuráveis.*

1. *A partição por níveis $\mathcal{P}_f$ é uma partição por medida de X e*
$$f = \mathop{+}_{c \in f(X)} c\mathbf{1}_{f^{-1}(c)}.$$
2. *A partição por níveis de $f + f'$ é mais grossa que o refinamento das partições por níveis de f e f':*
$$\mathcal{P}_{f+f'} \leq \mathcal{P}_f \vee \mathcal{P}_{f'}.$$

$:\vdash$ **Definição 15.32.** Sejam $\boldsymbol{X}$ um espaço de medida e $f \in \mathscr{M}(\boldsymbol{X}, [0, \infty])$ uma função simples mensurável. A *integral* de f em $\boldsymbol{X}$ é
$$\int f \mathrm{d}m := \mathop{+}_{c \in f(X)} c\, m(f^{-1}(c)).$$
Para todo conjunto mensurável $M \in \mathcal{M}$, a *integral* de f sobre M em $\boldsymbol{X}$ é
$$\int_M f \mathrm{d}m := \int \mathbf{1}_M f \mathrm{d}m.$$
Quando não for necessário explicitar a medida m, escreveremos
$$\int f.$$
Quando for necessário explicitar a variável da função f, escreveremos
$$\int f(x) \mathrm{d}m(x).$$

Para denotar a integral, a notação
$$\int\limits_{x \in X} f(x)$$
também poderia ser usada, e teria a vantagem de se assemelhar mais com a notação de somatório
$$\mathop{+}_{i \in I} f_i.$$
A notação da definição, no entanto, tem a vantagem de evitar escrever a variável x, que é de fato desnecessária na maioria dos contextos. Essa notação não é usual e não será usada aqui.

$\vdash$ **Proposição 15.34.** *Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida.*

1. *Para todo* $M \in \mathcal{M}$, $\displaystyle\int \mathbf{1}_M = m(M)$.
2. *Para todo* $M \in \mathcal{M}$ *e toda função simples mensurável* $f\colon X \to \mathbb{R}$,
$$\int_M f = \mathop{+}_{c \in f(X)} c\, m(M \cap f^{-1}(c)).$$
3. *Se* $f \in \mathscr{M}(\boldsymbol{X}, [0, \infty])$ *é uma função simples mensurável,* $\mathcal{P}$ *é uma partição tal que* $\mathcal{P}_f \leq \mathcal{P}$ *e, para todo* $P \in \mathcal{P}$, $f|_P = c_P$, *então*
$$\int f = \mathop{+}_{P \in \mathcal{P}} c_P\, m(P).$$

□ *Demonstração.* 1. Como $\mathbf{1}_M(X) = \{0, 1\}$, $\mathbf{1}_M^{-1}(1) = M$ e $\mathbf{1}_M^{-1}(0) = M^{\complement}$,
$$\int \mathbf{1}_M = \mathop{+}_{c \in \mathbf{1}_M(X)} c\, m(\mathbf{1}_M^{-1}(c)) = 1m(M) + 0m(M^{\complement}) = m(M).$$

■

⊢ **Proposição 15.35.** *Sejam* $(X, \mathcal{M}, m)$ *um espaço de medida,* $M \in \mathcal{M}$ *um conjunto mensurável* $f\colon X \to \mathbb{R}$ *e* $f'\colon X \to \mathbb{R}$ *funções simples e* $a \in \mathbb{R}$. *Então*

1. *A função* af *é função simples mensurável e* $\displaystyle\int_M af = a \int_M f$.
2. *A função* $f + f'$ *é função simples mensurável e* $\displaystyle\int_M (f + f') = \int_M f + \int_M f'$.

□ *Demonstração.* 1. Notemos que $(af)(X) = af(X)$, pois todo elemento de $(af)(X)$ é da forma ac, para $c \in f(X)$. Isso significa que $(af)(X)$ é finito, logo af é simples. Agora, separamos em dois casos. Se $a = 0$, então $af = 0f = 0$, logo $0f(X) = \{0\}$ e $(0f)^{-1}(0) = X$, portanto
$$\int_M 0f = \mathop{+}_{c \in (0f)(X)} cm(M \cap (0f)^{-1}(c)) = 0m(M \cap X) = 0 = 0 \int_M f.$$

Se $a \neq 0$, então
$$(af)^{-1}(ac) = \{x \in X \mid af(x) = ac\} = \{x \in X \mid f(x) = c\} = f^{-1}(c).$$

Nesse caso segue que

$$\begin{aligned}\int_M af &= \mathop{+}_{c\in(af)(X)} cm(M\cap(af)^{-1}(c))\\ &= \mathop{+}_{c\in f(X)} acm(M\cap(af)^{-1}(ac))\\ &= \mathop{+}_{c\in f(X)} acm(M\cap f^{-1}(c))\\ &= a\mathop{+}_{c\in f(X)} cm(M\cap f^{-1}(c))\\ &= a\int_M f.\end{aligned}$$

2. Notemos que

$$\begin{aligned}(f+f')(X) &= \{f(x)+f'(x)\mid x\in X\}\\ &= \left\{c+c' \;\middle|\; (c,c')\in f(X)\times f'(X), f^{-1}(c)\cap(f')^{-1}(c')\neq\emptyset\right\}\\ &\subseteq f(X)+f'(X).\end{aligned}$$

Como $f(X)+f'(X)$ é finito, então $(f+f')(X)$ é finito, logo $f+f'$ é simples. Para todo $x\in X$, existem únicos $c\in f(X)$ e $c'\in f'(X)$ tais que $x\in f^{-1}(c)$ e $x\in(f')^{-1}(c')$, pois $\{f^{-1}(c)\}_{c\in f(X)}$ e $\{(f')^{-1}(c)\}_{c\in f'(X)}$ são partições de X. Logo $(f+f')(x)=f(x)+f'(x)=c+c'=(c+c')\mathbf{1}_{f^{-1}(c)\cap f^{-1}(c')}(x)$, o que mostra que

$$f+f' = \mathop{+}_{c\in f(X)}\mathop{+}_{c'\in f(X)}(c+c')\mathbf{1}_{f^{-1}(c)\cap f^{-1}(c')},$$

Como $\{f^{-1}(c)\cap f^{-1}(c')\}_{(c,c')\in f(X)\times f(X)} = \{f^{-1}(c)\}_{c\in f(X)}\vee\{(f')^{-1}(c)\}_{c\in f'(X)}$ é o refinamento comum da partição por níveis de $f+f$, segue que

$$\begin{aligned}\int f+f' &= \mathop{+}_{c\in f(X)}\mathop{+}_{c'\in f(X)}(c+c')m\left(f^{-1}(c)\cap f^{-1}(c')\right)\\ &= \mathop{+}_{c\in f(X)}\mathop{+}_{c'\in f(X)} c\,m\left(f^{-1}(c)\cap f^{-1}(c')\right)+c'\,m\left(f^{-1}(c)\cap f^{-1}(c')\right)\\ &= \mathop{+}_{c\in f(X)} c\left(\mathop{+}_{c'\in f(X)} m\left(f^{-1}(c)\cap f^{-1}(c')\right)\right)\\ &\qquad + \mathop{+}_{c'\in f(X)} c'\left(\mathop{+}_{c\in f(X)} m\left(f^{-1}(c)\cap f^{-1}(c')\right)\right)\\ &= \mathop{+}_{c\in f(X)} c\,m(f^{-1}(c)) + \mathop{+}_{c'\in f'(X)} c'\,m((f')^{-1}(c'))\\ &= \int f+\int f'.\end{aligned}$$

■

□ *Demonstração.* Sejam $f = +_{i=1}^{n} c_i \mathbf{1}_{P_i}$, $g = +_{j=1}^{m} d_j \mathbf{1}_{Q_j}$, $I := \{1, \cdots, n\}$ e $J := \{1, \cdots, m\}$.

1. Se $c = 0$, vale a igualdade, pois $0f = 0\mathbf{1}_X$, logo

$$\int_M 0f = 0m(M \cap X) = 0 = 0 \int_M f.$$

Se $c \neq 0$, então $cf(X) = \{cc_1, \cdots, cc_n\}$ e as constantes $cc_1, \cdots, cc_n$ são todas distintas. Definindo, para todo $i \in I$, $R_i := \{x \in X \mid cf(x) = cc_i\}$, os conjuntos $R_1, \cdots, R_n$ formam uma partição de X em conjuntos mensuráveis. Além disso, temos $R_i = P_i$ para todo $i \in I$ porque, como $c \neq 0$, segue que $f(x) = c_i$ se, e somente se, $cf(x) = cc_i$. Portanto

$$\int_M cf = \mathop{+}_{i=1}^{n} cc_i m(R_i \cap M) = c \mathop{+}_{i=1}^{n} c_i m(P_i \cap M) = c \int_M f.$$

2. Como $f(X)$ e $g(X)$ são conjuntos finitos,

$$(f + g)(X) := \{c_i + d_j \mid (i, j) \in I \times J\}$$

é um conjunto finito. No entanto, não necessariamente $(f + g)(X)$ tem mn elementos, pois podem existir $(i_1, j_1), (i_2, j_2) \in I \times J$ distintos tais que $c_{i_1} + d_{j_1} = c_{i_2} + d_{j_2}$. Sejam $e_1, \cdots, e_l \in \mathbb{R}$ as constantes distintas tais que $(f + g)(X) = \{e_1, \cdots, e_l\}$, $K := \{1, \cdots, l\}$ e $R_k := \{x \in X \mid (f + g)(x) = e_k\}$. Nesse caso, $\{R_k \mid k \in K\}$ é uma partição de X em conjuntos mensuráveis e

$$f + g = \mathop{+}_{k=1}^{l} e_k \mathbf{1}_{R_k}.$$

Por outro lado, temos

$$\begin{aligned}(f + g)(x) &= f(x) + g(x) \\ &= \mathop{+}_{i=1}^{n} c_i \mathbf{1}_{P_i}(x) + \mathop{+}_{j=1}^{m} d_j \mathbf{1}_{Q_j}(x) \\ &= \end{aligned}$$

$$f + g = \mathop{+}_{i=1}^{n} \mathop{+}_{j=1}^{m} (c_i + d_j) \mathbf{1}_{P_i \cap Q_j}.$$

Isso significa que os conjuntos

■

15.9.2 Integral de funções mensuráveis positivas

15.9.3 Integral de funções mensuráveis

15.9.4 Teoremas de convergências

15.9.5 Mudança de variáveis na integração

Lembremos que, se $T\colon (X,\mathcal{M}) \to (X',\mathcal{M}')$ é uma função mensurável e m é uma medida sobre $(X,\mathcal{M})$, então a medida $T_\star m$ empurrada de m por T é a medida dada por

$$T_\star m = m \circ T^{-1}.$$

Ainda, se $f\colon X' \to X''$ é uma função, a função puxada de f por T é a função $T^\star f\colon X \to X''$ dada por

$$T^\star f = f \circ T.$$

⊢ **Proposição 15.36.** *Sejam $\boldsymbol{X} = (X,\mathcal{M},m)$ e $\boldsymbol{X}' = (X',\mathcal{M}',m')$ espaços de medida e $T\colon X \to X'$ uma função mensurável tal que $m' = T_\star m$. Para toda $f \in \mathscr{M}(X',\mathbb{R})$, $T^\star f \in \mathscr{I}^1(X,\mathbb{R})$ se, e somente se, $f \in \mathscr{I}^1(X',\mathbb{R})$ e, nesse caso,*

$$\int T^\star f \mathrm{d}m = \int f \mathrm{d}T_\star m.$$

□ *Demonstração.* A demonstração é evidente para funções simples, e pelo teorema da convergência monótona segue para qualquer função. Para qualquer conjunto mensurável $M \in \mathcal{M}'$, notemos que

$$T^\star \mathbf{1}_M = \mathbf{1}_M \circ T = \mathbf{1}_{T^{-1}(M)},$$

logo

$$\int T^\star \mathbf{1}_M \mathrm{d}m = \int \mathbf{1}_{T^{-1}(M)} \mathrm{d}m = m(T^{-1}(M)) = T_\star m(M) = \int \mathbf{1}_M \mathrm{d}T_\star m.$$

Da linearidade de $T^\star$, segue que para qualquer função simples $f\colon X' \to \mathbb{R}$,

$$T^\star f = T^\star \left(\mathop{+}_{c \in f(X')} c\mathbf{1}_{f^{-1}(c)} \right) = \mathop{+}_{c \in f(X')} cT^\star \mathbf{1}_{f^{-1}(c)} = \mathop{+}_{c \in (T^\star f)(X)} c\mathbf{1}_{(T^\star f)^{-1}(c)}$$

é uma função simples. Da linearidade da integral segue então que

$$\begin{aligned}\int T^\star f \mathrm{d}m &= \int \mathop{+}_{c\in(T^\star f)(X)} c\mathbf{1}_{(T^\star f)^{-1}(c)}\mathrm{d}m \\ &= \int \mathop{+}_{c\in f(X')} cT^\star\mathbf{1}_{f^{-1}(c)}\mathrm{d}m \\ &= \mathop{+}_{c\in f(X')} c\int T^\star\mathbf{1}_{f^{-1}(c)}\mathrm{d}m \\ &= \mathop{+}_{c\in f(X')} c\int \mathbf{1}_{f^{-1}(c)}\mathrm{d}T_\star m \\ &= \int \mathop{+}_{c\in f(X')} c\mathbf{1}_{f^{-1}(c)}\mathrm{d}T_\star m \\ &= \int f\mathrm{d}T_\star m.\end{aligned}$$

Para uma função mensurável positiva $f\colon X' \to [0,\infty[$, existe uma sequência crescente $(f_n)_{n\in\mathbb{N}}$ de funções mensuráveis simples positivas que converge para f. Segue que $(T^\star f_n)_{n\in\mathbb{N}}$ é uma sequência crescente de funções mensuráveis simples positivas que converge para $T^\star f$ e, pela convergência monótona da integral,

$$\int T^\star f\mathrm{d}m = \lim_{n\to\infty}\int T^\star f_n\mathrm{d}m = \lim_{n\to\infty}\int f_n\mathrm{d}T_\star m = \int f\mathrm{d}T_\star m.$$

Agora, para qualquer função mensurável $f\colon X' \to \mathbb{R}$, vale que $(T^\star f)^+ = T^\star f^+$ e $(T^\star f)^- = T^\star f^-$, portanto segue que

$$\begin{aligned}\int T^\star f\mathrm{d}m &= \int (T^\star f)^+\mathrm{d}m - \int (T^\star f)^-\mathrm{d}m \\ &= \int T^\star f^+\mathrm{d}m - \int T^\star f^-\mathrm{d}m \\ &= \int f^+\mathrm{d}T_\star m - \int f^-\mathrm{d}T_\star m \\ &= \int f\mathrm{d}T_\star m.\end{aligned}$$

■

15.9.6 Integral em espaços normados completos

Queremos definir o conceito de integração de funções de espaços de medida para espaços normados completos. Isso generaliza o caso de funções reais e complexas, pois de fato tudo que se precisa para integração são linearidade, norma e completude. Para estudar funções mensuráveis de um espaço de medida para um espaço normado, consideraremos no espaço normado a álgebra de mensuráveis topológica, gerada pelos abertos da topologia da norma. Começamos pelas funções simples.

15.9.7 Funções simples

:⊢ **Definição 15.33.** Sejam $\boldsymbol{X}$ um espaço de medida e $\mathbb{E}$ um espaço normado real. Uma *função simples* de $\boldsymbol{X}$ para $\boldsymbol{E}$ é uma função $f\colon X \to E$ tal que $f(X)$ é finito. A *partição por níveis* de f é o conjunto

$$\mathcal{P}_f := \{f^{-1}(c)\}_{c \in f(X)}.$$

O conjunto das funções simples mensuráveis de X para E é denotado $\mathscr{M}_s(X, E)$.

⊢ **Proposição 15.37.** *Sejam $\boldsymbol{X}$ um espaço de medida, $\mathbb{E}$ um espaço normado real e $f, f'\colon X \to E$ funções simples.*

1. *Uma função simples $f\colon X \to E$ é mensurável se, e somente se, sua partição por níveis $\mathcal{P}_f$ é mensurável;*
2. *A função f pode ser decomposta em funções indicadoras como*

$$f = \mathop{+}_{v \in f(X)} \mathbf{1}_{f^{-1}(v)} v$$

e, para toda partição $\mathcal{P} \geq \mathcal{P}_f$, definindo $\{v_P\} := f(P)$,

$$f = \mathop{+}_{P \in \mathcal{P}} \mathbf{1}_P v_P;$$

3. *Para todas funções simples $f, f'\colon X \to E$, a partição por níveis de $f + f'$ é mais grossa que o refinamento das partições por níveis de f e f':*

$$\mathcal{P}_{f+f'} \leq \mathcal{P}_f \vee \mathcal{P}_{f'}.$$

□ *Demonstração.* 1. Suponha que $f\colon X \to E$ é uma função simples mensurável. Seja $v \in f(X)$. Como $\{v\}$ é fechado e f é mensurável, $f^{-1}(\{v\}) = f^{-1}(v)$ é mensurável.

Reciprocamente, seja $M \subseteq E$ um conjunto mensurável. Como $f(X)$ é finito, segue que $M \cap f(X)$ é finito. Seja $v_0, \cdots, v_{n-1} \in E$ tais que $M \cap f(X) = \bigcup_{i \in [n]} \{v_i\}$. Então

$$f^{-1}(M) = f^{-1}(M \cap f(X)) = f^{-1}\left(\bigcup_{i \in [n]} \{v_i\}\right) = \bigcup_{i \in [n]} f^{-1}(\{v_0\}).$$

Como $f^{-1}(\{v_i\})$ são mensuráveis, pois são elementos da partição por níveis de f, segue que $f^{-1}(M)$ é mensurável, pois é união de mensuráveis, o que mostra que f é mensurável.

2. Seja $x \in X$. Como $\mathcal{P}_f$ é partição de X, existe $v \in f(X)$ tal que $x \in f^{-1}(v)$. Nesse caso, $f(x) = v$, $\mathbf{1}_{f^{-1}(v)}(x) = 1$ e, para todo $v' \in f(X) \smallsetminus \{v\}$, $\mathbf{1}_{f^{-1}(v')}(x) = 0$, logo
$$f(x) = 1v = \mathbf{1}_{f^{-1}(v)}(x)v = \mathop{\bigplus}_{v \in f(X)} \mathbf{1}_{f^{-1}(v)}(x)v,$$
portanto $f = \bigplus_{v \in f(X)} \mathbf{1}_{f^{-1}(v)} v$. A outra igualdade é semelhante.
3. Exercício.

■

⊢ **Proposição 15.38.** *Sejam $\boldsymbol{X}$ um espaço de medida e $\mathbb{E}$ um espaço normado real. O conjunto $\mathscr{M}_s(\boldsymbol{X}, \boldsymbol{E})$ de funções simples mensuráveis é um subespaço linear real de $\mathscr{M}(\boldsymbol{X}, \boldsymbol{E})$. Se $f \colon X \to E$ é uma função simples mensurável, então $\|f\| : X \to \mathbb{R}$ também é.*

□ *Demonstração.* Para ver que $\mathscr{M}_s(\boldsymbol{X}, \boldsymbol{E})$ é um espaço linear, basta mostrar que ele subespaço linear de $\mathscr{M}(\boldsymbol{X}, \boldsymbol{E})$. Sejam $c \in \mathbb{R}$ e $f, f' \in \mathscr{M}_s(X, E)$. Como $f(X)$ e $f'(X)$ são finitos, e $(cf + f')(X) \subseteq cf(X) + f'(X)$, claramente $cf + f'$ é simples. Ainda, como f, f' são mensuráveis, segue que $cf + f'$ é mensurável, o que mostra que $cf + f' \in \mathscr{M}_s(X, E)$.

Para mostrar que $\|f\|$ é simples, basta notar que, como $f(X)$ é finito, claramente $\|f\|(X)$ é finito, e, como $\|\cdot\|$ é mensurável (pois é contínua), a composição $\|f\| = \|\cdot\| \circ f$ é mensurável. ■

⊢ **Proposição 15.39.** *Sejam $\boldsymbol{X}$ um espaço de medida e $\mathbb{E}$ um espaço normado real de dimensão finita. Uma função $f \colon X \to E$ é mensurável se, e somente se, existe uma sequência $(f_n)_{n \in \mathbb{N}}$ de funções simples de $\mathscr{M}_s(X, E)$ que convergem pontualmente para f.*

□ *Demonstração.* Como $\mathbb{E}$ tem dimensão finita, basta mostrar a proposição para $\mathbb{R}$. Suponhamos que f é mensurável. Para cada $n \in \mathbb{N}^*$, particionamos o intervalo $[-n, n[$ em intervalos de tamanho $\frac{1}{n}$ e definimos, para cada $k \in [-n^2 - 1, n^2 + 1[\cap \mathbb{Z}$, os conjuntos
$$X_k := \begin{cases} f^{-1}\left(]\infty, -n[\right), & k = -n^2 - 1 \\ f^{-1}\left(\left[\dfrac{k}{n}, \dfrac{k+1}{n}\right[\right), & k \in [-n^2, n^2[\cap \mathbb{Z} \\ f^{-1}\left([n, \infty[\right), & k = n^2. \end{cases}$$
Os conjuntos $X_k \subseteq X$ são mensuráveis e particionam X. Definimos $f_0 := 0$ e, para cada $n \in \mathbb{N}^*$, as funções $f_n \colon X \to \mathbb{R}$ por
$$f_n := \mathbf{1}_{X_{-n^2-1}}(-n) + \mathop{\bigplus}_{k=-n^2}^{n^2} \mathbf{1}_{X_k} \frac{k}{n}.$$

As funções f_n são funções simples e mensuráveis, e convergem pontualmente para f, o que termina a demonstração.

A recíproca é consequência de 15.16. ■

15.9.8 Desintegração de medida

Seja $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida. Nessa seção consideraremos partições mensuráveis e partições por medida (contavelmente gerada). Dada uma partição $\mathcal{P}$ de $\boldsymbol{X}$, podemos induzir em $\mathcal{P}$ uma σ-álgebra e uma medida. Para isso, definimos a projeção

$$\begin{aligned} \mathrm{p}\colon X &\longrightarrow \mathcal{P} \\ x &\longmapsto [x] = P_x. \end{aligned}$$

Como $\mathcal{P}$ é partição por medida (contavelmente gerada), a projeção p não está definida para todo $x \in X$, as para quase todo, portanto é uma quase função. Agora, a σ-álgebra sobre $\mathcal{P}$ é $\mathrm{p}_\star\mathcal{M}$, a σ-álgebra empurrada por p, e a medida sobre $(\mathcal{P}, \mathrm{p}_\star\mathcal{M})$ é $\mathrm{p}_\star m$, a medida empurrada por p. A tripla $(\mathcal{P}, \mathrm{p}_\star\mathcal{M}, \mathrm{p}_\star m)$ é um espaço de medida. Usaremos esse espaço de medida para falar sobre desintegração de medidas.

:⊢ **Definição 15.34.** Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida (finita) e $\mathcal{P}$ uma partição por medida (contavelmente gerada) de $\boldsymbol{X}$. Uma *desintegração* de μ relativa a $\mathcal{P}$ é uma família $(m_P)_{P\in\mathcal{P}}$ tal que

1. Para quase todo $P \in \mathcal{P}$,
$$m_P(P) = 1;$$
2. Para todo $M \in \mathcal{M}$, a função
$$\begin{aligned} m_{(\cdot)}(M)\colon \mathcal{P} &\longrightarrow \mathbb{R} \\ P &\longmapsto m_P(M) \end{aligned}$$
é mensurável;
3. Para todo $M \in \mathcal{M}$,
$$m(M) = \int m_P(M)\mathrm{d}\mathrm{p}_\star m(P).$$
$$m(M) = \int_{P\in\mathcal{P}} m_P(M)\mathrm{d}\mathrm{p}_\star m.$$

⊢ **Proposição 15.40** (Unicidade de Desintegração)**.** *Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida (finita) e $\mathcal{P}$ uma partição por medida (contavelmente gerada) de $\boldsymbol{X}$. Se $(m_P)_{P\in\mathcal{P}}$ e $(m'_P)_{P\in\mathcal{P}}$ são desintegrações de m relativas a $\mathcal{P}$, então, para quase todo $P \in \mathcal{P}$, $m_P = m'_P$.*

Capítulo 16

Espaços métricos

16.1 Espaço métrico

16.1.1 Métricas

:⊢ **Definição 16.1.** Seja M um conjunto. Uma *métrica* (ou *função distância*) em M é uma função

$$\begin{aligned} |\cdot,\cdot| \colon M \times M &\longrightarrow \mathbb{R} \\ (p,p') &\longmapsto |p,p'| \end{aligned}$$

que satisfaz

1. (Separação) Para todos $p, p' \in M$,

$$|p,p'| = 0 \quad \Longleftrightarrow \quad p = p';$$

2. (Simetria) Para todos $p, p' \in M$

$$|p,p'| = |p',p|\,;$$

3. (Desigualdade triangular) Para todos $p, p', p'' \in M$,

$$|p,p''| \leq |p,p'| + |p',p''|\,.$$

A *distância* entre p e p' é o número real $|p,p'|$.

Na definição da função distância geralmente se assume que o contradomínio é $[0,\infty[$. No entanto, pode-se mostrar que qualquer função real que satisfaz separação, simetria e desigualdade triangular é positiva. Por isso a proposição seguinte.

:⊢ **Definição 16.2.** Um *espaço métrico* é um par $\boldsymbol{M} = (M, |\cdot,\cdot|)$ em que M é um conjunto e $|\cdot,\cdot|$ é uma métrica em M. Os elementos de M são *pontos*. Um *subespaço métrico* de $\boldsymbol{M}$ é o par $\boldsymbol{S} = (S, |\cdot,\cdot|\,|_{S\times S})$.

⊢ **Proposição 16.1.** *Seja $\boldsymbol{M}$ um espaço métrico. Então*

1. *(Positividade) Para todos $p, p' \in M$, $|p, p'| \geq 0$.*
2. *(Desigualdade triangular generalizada) Para todos $p_0, \ldots, p_n \in M$,*

$$|p_0, p_n| \leq \mathop{\mathbin{+}}_{i=1}^{n-1} |p_i, p_{i+1}|$$

□ *Demonstração.* 1. Sejam $p, p' \in M$. Da separação, desigualdade triangular e da simetria de $|\cdot,\cdot|$, segue que

$$0 = \frac{|p,p|}{2} \leq \frac{|p,p'| + |p',p|}{2} = |p,p'|\,.$$

2. Para $n = 1$, seja $p_1 \in M$; então $|p_1, p_1| = 0$ e $\mathop{\mathbin{+}}_{i=1}^{0} |p_i, p_{i+1}| = 0$, pois a soma é vazia. Para $n = 2$, sejam $p_1, p_2 \in M$; então $|p_1, p_2|$ e $\mathop{\mathbin{+}}_{i=1}^{1} |p_i, p_{i+1}| = |p_1, p_2|$, e vale a propriedade. Para $n = 3$, sejam $p_1, p_2, p_3 \in M$; então a propriedade é a desigualdade triangular. Agora, sejam $n \geq 4$, $p_1, \ldots, p_n \in M$ e assumamos que a propriedade vale para todo $k \in \mathbb{N}$, tal que $3 \leq k \leq n-1$. Então

$$|p_1, p_n| \leq \mathop{\mathbin{+}}_{i=1}^{n-3} |p_i, p_{i+1}| + |p_{n-2}, p_n|\,,$$

pois essa soma tem $n-1$ termos e vale a hipótese de indução. Pela desigualdade triangular, vale que $|p_{n-2}, p_n| \leq |p_{n-2}, p_{n-1}| + |p_{n-1}, p_n|$, e, portanto,

$$\begin{aligned} |p_1, p_n| &\leq \mathop{\mathbin{+}}_{i=1}^{n-3} |p_i, p_{i+1}| + |p_{n-2}, p_n| \\ &\leq \mathop{\mathbin{+}}_{i=1}^{n-3} |p_i, p_{i+1}| + |p_{n-2}, p_{n-1}| + |p_{n-1}, p_n| \\ &= \mathop{\mathbin{+}}_{i=1}^{n-1} |p_i, p_{i+1}|\,. \end{aligned}$$

■

Alguns exemplos de métricas seguem. Podemos definir distâncias a partir de distâncias já conhecidas no espaço.

▷ **Exercício 16.1.** *Seja M um conjunto.*

1. *A* métrica discreta

$$\begin{aligned} |\cdot,\cdot|_d \colon M \times M &\longrightarrow \mathbb{R} \\ (p,p') &\longmapsto \begin{cases} 0, & p = p' \\ 1, & p \neq p' \end{cases} \end{aligned}$$

   é uma métrica sobre M.

2. *Se $|\cdot,\cdot|$ é uma métrica em M, então*

$$\begin{aligned} |\cdot,\cdot|' \colon M \times M &\longrightarrow [0,\infty[\\ (p,p') &\longmapsto \frac{|p,p'|}{1+|p,p'|} \end{aligned}$$

   é uma métrica sobre M.

⊢ **Proposição 16.2.** *Sejam M um conjunto e $|\cdot,\cdot|_0, \ldots, |\cdot,\cdot|_{n-1}$ métricas em M. Então a função*

$$\begin{aligned} |\cdot,\cdot| \colon M \times M &\longrightarrow \mathbb{R} \\ (p,p') &\longmapsto \mathop{\vdots}_{i=0}^{n-1} |p,p'|_i \end{aligned}$$

é uma métrica em M.

□ *Demonstração.* 1. (Separação) Sejam $p, p' \in M$. Suponhamos que

$$|p,p'| = \mathop{\vdots}_{i=0}^{n} |p,p'|_i = 0.$$

Como, para todo $i \in [n]$, $|p,p'|_i \geq 0$, então, para todo $i \in [n]$, $|p,p'|_i = 0$. Logo $p = p'$. Reciprocamente, suponhamos $p = p'$. Então, para todo $i \in [n]$, $|p,p'|_i = 0$, o que implica

$$|p,p'| = \mathop{\vdots}_{i=0}^{n-1} 0 = 0.$$

2. (Simetria) Sejam $p, p' \in M$. Então, pela simetria de $|\cdot,\cdot|_i$ para todo $i \in [n]$,

$$|p,p'| = \mathop{\vdots}_{i=0}^{n-1} |p,p'|_i = \mathop{\vdots}_{i=1}^{n-1} |p,p'|_i = |p,p'|\,.$$

3. (Desigualdade Triangular) Sejam $p, p', p'' \in M$. Então, para todo $i \in [n]$, vale $|p, p''|_i \leq |p, p'|_i + |p', p''|_i$ pela desigualdade triangular de $|\cdot, \cdot|_i$, e segue que

$$\begin{aligned} |p, p''| &= \mathop{\mathbin{+}}_{i=0}^{n-1} |p, p''|_i \\ &\leq \mathop{\mathbin{+}}_{i=0}^{n-1} (|p, p'|_i + |p', p''|_i) \\ &= \mathop{\mathbin{+}}_{i=0}^{n-1} |p, p'|_i + \mathop{\mathbin{+}}_{i=0}^{n-1} |p', p''|_i \\ &= |p, p'| + |p', p''| . \end{aligned}$$

■

⊢ **Proposição 16.3.** *Sejam M um conjunto não vazio e $|\cdot, \cdot|_0, \ldots, |\cdot, \cdot|_{n-1}$ métricas em M. Então a função*

$$\begin{aligned} |\cdot, \cdot| \colon M \times M &\longrightarrow \mathbb{R} \\ (p, p') &\longmapsto \max \{|p, p'|_i \mid i \in [n]\} \end{aligned}$$

é uma métrica sobre M.

□ *Demonstração.* Demonstraremos para $n = 2$, pois o caso geral é análogo.

1. (Separação) Sejam $p, p' \in M$. Suponhamos que

$$|p, p'| = \max\{|p, p'|_1, |p, p'|_2\} = 0.$$

Então $|p, p'|_1 = 0$ ou $|p, p'|_2 = 0$. Em ambos os casos, temos $p = p'$. Reciprocamente, suponhamos que $p = p'$. Então $|p, p'|_1 = 0$ e $|p, p'|_2 = 0$, o que implica $|p, p'| = \max\{|p, p'|_1, |p, p'|_2\} = 0$.

2. (Simetria) Sejam $p, p' \in M$. Então

$$|p, p'| = \max\{|p, p'|_1, |p, p'|_2\} = \max\{|p', p|_1, |p', p|_2\} = |p', p| .$$

3. (Desigualdade Triangular) Sejam $p, p', p'' \in M$. Então $|p, p''| = |p, p''|_1$ ou $|p, p''| = |p, p''|_2$. No primeiro caso, segue que

$$|p, p''| = |p, p''|_1 \leq |p, p'|_1 + |p', p''|_1 \leq |p, p'| + |p', p''| .$$

No segundo caso, segue que

$$|p, p''| = |p, p''|_2 \leq |p, p'|_2 + |p', p''|_2 \leq |p, p'| + |p', p''| .$$

■

▷ **Exercício 16.2** (Métrica Limitada)**.** *Seja* $(M, |\cdot,\cdot|)$ *um espaço métrico. A função*

$$\begin{aligned} |\cdot,\cdot| \wedge 1 \colon M \times M &\longrightarrow \mathbb{R} \\ (p, p') &\longmapsto |p, p'| \wedge 1 \end{aligned}$$

é uma métrica sobre M*, a* métrica limitada induzida por $|\cdot,\cdot|$.

▷ **Exercício 16.3** (Métricas p)**.** *Sejam* $\boldsymbol{M}$ *e* $\boldsymbol{M'}$ *espaços métricos e* $p \in [1, \infty[$*. A função*

$$\begin{aligned} |\cdot,\cdot|_p \colon (M \times M') \times (M \times M') &\longrightarrow \mathbb{R} \\ ((x, x'), (y, y')) &\longmapsto (|x, y|^p + |x', y'|'^p)^{\frac{1}{p}} \end{aligned}$$

é uma métrica sobre $M \times M'$*.*

:⊢ **Definição 16.3.** Seja $d \in \mathbb{N}$. A *métrica reta* sobre $\mathbb{R}^d$ é a função

$$\begin{aligned} |\cdot,\cdot|_{\mathbb{R}^d} \colon \mathbb{R}^d &\longrightarrow \mathbb{R}^d \\ (x, x') &\longmapsto \left(\mathop{+}_{i \in [d]} |x_i - x'_i|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

16.1.2 Diâmetro, bolas e limitação

:⊢ **Definição 16.4.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$. O *diâmetro* de C é

$$\Theta(C) := \sup \{|p, p'| \mid p, p' \in C\}$$

se $\{|p, p'| \mid p, p' \in C\}$ é limitado superiormente, e ∞, caso contrário. Um *conjunto limitado em* $\boldsymbol{M}$ é um conjunto $C \subseteq M$ tal que $\Theta(C) < \infty$.

Na definição, adotamos a convenção de que $\sup \emptyset = 0$ e $\sup [0, \infty[= \infty$. Isso é só parcialmente uma convenção, pois a ambiguidade não está em que valor atribuir a $\sup \emptyset$, mas sim em qual conjunto parcialmente ordenado está sendo considerado. Por exemplo, quando o conjunto ordenado é $[0, \infty]$, $\sup_{[0,\infty]} \emptyset = 0$, e quando o conjunto ordenado é $[-\infty, +\infty]$, $\sup_{[-\infty,+\infty]} \emptyset = -\infty$.

Isso define uma função

$$\begin{aligned} \Theta \colon \wp(M) &\longrightarrow [0, \infty] \\ C &\longmapsto \Theta(C). \end{aligned}$$

:⊢ **Definição 16.5.** Sejam $\boldsymbol{M}$ um espaço métrico, $c \in M$ e $r \in [0, \infty[$. A *bola aberta* de centro c e raio r em M é o conjunto

$$\mathrm{B}\,(c; r) := \{p \in M \mid |c, p| < r\}\,.$$

A *bola fechada* de centro c e raio r em M é o conjunto

$$\bar{\mathrm{B}}\,(c; r) := \{p \in M \mid |c, p| \leq r\}\,.$$

No caso em que $M = \mathbb{R}$, o *intervalo aberto* de centro c e raio r é o conjunto

$$\mathrm{I}\,(c; r) := \mathrm{B}\,(c; r) =]c - r, c + r[$$

e o *intervalo fechado* de centro c e raio r é o conjunto

$$\bar{\mathrm{I}}\,(c; r) := \bar{\mathrm{B}}\,(c; r) = [c - r, c + r]\,.$$

FIGURA 16.1: Bolas aberta e fechada de centro c e raio r, respectivamente.

⊢ **Proposição 16.4.** *Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Então C é um conjunto limitado se, e somente se, existe bola $\bar{\mathrm{B}}\,(c; r)$ tal que $C \subseteq \bar{\mathrm{B}}\,(c; r)$.*

□ *Demonstração.* Se C é limitado, basta tomar $r = \Theta(C)$ e $c \in C$. Reciprocamente, se existe bola $\bar{\mathrm{B}}\,(c; r)$ tal que $C \subseteq \bar{\mathrm{B}}\,(c; r)$, então para todos $p, p' \in C$, segue da desigualdade triangular que

$$|p, p'| \leq |p, c| + |c, p'| \leq r + r = 2r,$$

portanto $\Theta(C) \leq 2r \in [0, \infty[$. ■

▷ **Exercício 16.4.** *Seja $\boldsymbol{M}$ um espaço métrico.*

1. *Para todos $C, C' \subseteq M$ tais que $C \subseteq C'$,*

$$\Theta(C) \leq \Theta(C').$$

2. *Para todos* $C, C' \subseteq M$,

$$\Theta(C \cup C') = \max\{\Theta(C), \Theta(C'), \sup\{|p, p'| \mid p \in C, p' \in C'\}\}.$$

3. *Para todo* $r \in [0, \infty[$ *e todo* $p \in M$.

$$\Theta(\mathrm{B}\,(p; r)) \leq 2r.$$

Note que o diâmetro da bola de raio r não necessariamente é $2r$, por exemplo para $r = 2$ na métrica discreta.

▷ **Exercício 16.5.** *Sejam* M *um conjunto,* $|\cdot, \cdot|$ *uma métrica em* M *e* $|\cdot, \cdot| \wedge 1$ *a métrica limitada sobre* M. *Todo subconjunto de* M *é limitado com respeito a* $|\cdot, \cdot| \wedge 1$.

:⊢ **Definição 16.6.** Sejam X um conjunto e $\boldsymbol{M}$ um espaço métrico. Uma função *limitada* de X para $\boldsymbol{M}$ é uma função $f\colon X \longrightarrow M$ de X para M cuja imagem $f(X)$ é limitada. O conjunto das funções limitadas de X para M é denotado $\mathscr{B}(X, M)$.

▷ **Exercício 16.6.** *Sejam* X *um conjunto e* $\boldsymbol{M}$ *e* $\boldsymbol{M}'$ *espaços métricos. Para todas funções limitadas* $f\colon X \longrightarrow M$ *e* $f'\colon M \longrightarrow M'$, *a composição* $f' \circ f\colon X \longrightarrow M'$ *é limitada.*

16.2 Topologia dos espaços métricos

16.2.1 Interior e pontos interiores

:⊢ **Definição 16.7.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Um *ponto interior* de C é um ponto $p \in C$ para o qual existe um número real $r > 0$ tal que $\mathrm{B}\,(p; r) \subseteq C$. O *interior* de C é o conjunto C° de todos pontos interiores de C. Um *conjunto aberto* de $\boldsymbol{M}$ é um conjunto $A \subseteq M$ tal que $A = A^\circ$. O conjunto dos conjuntos abertos de $\boldsymbol{M}$ é denotado $\mathcal{T}_{\boldsymbol{M}}$.

⊢ **Proposição 16.5.** *Seja* $\boldsymbol{M} = (M, |\cdot, \cdot|)$ *um espaço métrico. Então*

1. *Para todo* $c \in M$ *e para todo número real* $r > 0$, *a bola aberta* $\mathrm{B}\,(c; r)$ *é um conjunto aberto;*
2. *O conjunto* $\mathcal{T}_{\boldsymbol{M}}$ *é uma topologia de* M.

□ *Demonstração.* 1. Sejam $c \in M$ e $r \in]0, \infty[$. Queremos mostrar que $\mathrm{B}\,(c; r)$ é aberto. Para isso, seja $p \in \mathrm{B}\,(c; r)$. Então segue que $d := |c, p| < r$, pela definição de bola aberta, e, portanto, $r - d \in]0, \infty[$. Para mostrar que essa bola centrada em p está contida na bola maior centrada em c, seja

$p' \in \mathrm{B}(p; r-d)$. Então $|p, p'| < r - d$ e, pela desigualdade triangular, segue que

$$|c, p'| \leq |c, p| + |p, p'| < d + (r - d) = r,$$

o que mostra que $p' \in \mathrm{B}(c; r)$ e que, portanto, $\mathrm{B}(p; s) \subseteq \mathrm{B}(c; r)$. Assim, mostramos que $\mathrm{B}(p; r)$ é aberta.

2. 2.1. Podemos notar que $\emptyset$ é aberto por vacuidade, pois, se não fosse, existiria $p \in \emptyset$ para o qual não há $r \in]0, \infty[$ satisfazendo $\mathrm{B}(p; r) \subseteq A$, o que é absurdo. Para mostrar que M é aberto, sejam $p \in M$ e $r \in]0, \infty[$. Então $\mathrm{B}(p; r) \subseteq M$, pois qualquer bola aberta é subconjunto de M. Portanto M é aberto.

   2.2. Seja $(A_i)_{i \in I}$ uma família de abertos em $\boldsymbol{M}$ e seja $p \in (A_i)_{i \in I}$. Então existe $k \in I$ tal que $p \in A_k$. Como A_k é aberto, então existe $r \in]0, \infty[$ tal que $\mathrm{B}(p; r) \subseteq A_k$. Como $A_k \subseteq (A_i)_{i \in I}$, segue que $\mathrm{B}(p; r) \subseteq (A_i)_{i \in I}$ e que, portanto, $(A_i)_{i \in I}$ é aberto.

   2.3. Seja $(A_i)_{i \in [n]}$ uma sequência de abertos em $\boldsymbol{M}$ e seja $p \in (A_i)_{i \in [n]}$. Então, para todo $k \in [n]$, $p \in A_k$. Como, para todo $k \in [n]$, A_k é aberto, segue que existe $r_k \in]0, \infty[$ tal que $\mathrm{B}(p; r - k) \subseteq A_k$, Seja $r := \min\{r_k \mid k \in [n]\}$. Então, para todo $k \in [n]$, vale $\mathrm{B}(p; r) \subseteq \mathrm{B}(p; r - k)$, e segue que $\mathrm{B}(p; r) \subseteq A_k$ e, portanto, $\mathrm{B}(p; r) \subseteq (A_i)_{i \in [n]}$, o que mostra que $(A_i)_{i \in [n]}$ é aberto.

■

▷ **Exercício 16.7.** *Sejam M um conjunto e $|\cdot, \cdot|$ uma métrica em M. A métrica limitada $|\cdot, \cdot| \wedge 1$ sobre M induz a mesma topologia que $|\cdot, \cdot|$ sobre M.*

16.2.2 Limites e convergência de sequências

:⊢ **Definição 16.8** (Limite e convergência)**.** Sejam $\boldsymbol{M}$ um espaço métrico e $(p_n)_{n \in \mathbb{N}}$ uma sequência de pontos em M. Um *limite* da sequência $(p_n)_{n \in \mathbb{N}}$ é um ponto $p_\infty \in M$ que satisfaz:

- para todo $\varepsilon \in]0, \infty[$, algum $n_\varepsilon \in \mathbb{N}$ satisfaz que, para todo $n \in \mathbb{N}$ tal que $n \geq n_\varepsilon$,

$$p_n \in \mathrm{B}(p_\infty; \varepsilon).$$

Nesse caso, a sequência $(p_n)_{n \in \mathbb{N}}$ *converge* para p_∞ e denota-se $(p_n)_{n \in \mathbb{N}} \to p_\infty$ ou

$$(p_n)_{n \in \mathbb{N}} \underset{n \to \infty}{\longrightarrow} p_\infty.$$

Uma *sequência convergente* é uma sequência que tem limite. Uma sequência *divergente* é uma sequência que não tem limite.

$\vdash$ **Proposição 16.6.** *Todo espaço métrico* $\boldsymbol{M}$ *é um espaço topológico separado.*

$\square$ *Demonstração.* Sejam $p, p' \in M$ pontos distintos. Mostraremos que existe um número real r tal que $0 < r \leq \frac{1}{2}|p, p'|$, e que isso implica que $\mathrm{B}(p; r) \cap \mathrm{B}(p'; r) = \emptyset$. Como $p \neq p'$, então $|p, p'| > 0$, portanto existe $r \in \mathbb{R}$ tal que $0 < r \leq \frac{1}{2}|p, p'|$. Suponhamos que existe $p'' \in \mathrm{B}(p; r) \cap \mathrm{B}(p'; r)$. Então $|p, p''| < r$ e $|p', p''| < r$. Mas, pela desigualdade triangular, segue que

$$|p, p'| \leq |p, p''| + |p'', p'| < r + r \leq |p, p'|,$$

o que é absurdo. Portanto $\mathrm{B}(p; r) \cap \mathrm{B}(p'; r) = \emptyset$. ■

$\vdash$ **Corolário 16.7.** *Toda sequência convergente em um espaço métrico* $\boldsymbol{M}$ *tem limite único.*

$\square$ *Demonstração.* Suponhamos que p, p' são limites de $(p_n)_{n\in\mathbb{N}}$. Se $p \neq p'$, então $|p, p'| > 0$. Seja $\varepsilon \in \mathbb{R}$ tal que $0 < \varepsilon \leq \frac{1}{2}|p, p'|$. Então existe $N_1 \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N_1$, então $p_n \in \mathrm{B}(p; \varepsilon)$, e existe $N_2 \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N_2$, então $p_n \in \mathrm{B}(p'; \varepsilon)$. Assim, definindo $N := N_1 \vee N_2$, segue que, se $n \geq N$, então $n \geq N_1$ e $n \geq N_2$, e, portanto, que $p_n \in \mathrm{B}(p; \varepsilon)$ e $p_n \in \mathrm{B}(p'; \varepsilon)$; ou seja, $p_n \in \mathrm{B}(p; \varepsilon) \cap \mathrm{B}(p'; \varepsilon)$, mas isso é absurdo, pois $\mathrm{B}(p; \varepsilon) \cap \mathrm{B}(p'; \varepsilon) = \emptyset$. Portanto $p = p'$. ■

Essa proposição nos permite tratar o limite de uma sequência como um número único e, por isso, podemos usar a notação $\lim\limits_{n\in\mathbb{N}} p_n = p$ para quando $(p_n)_{n\in\mathbb{N}} \to p$.

$\vdash$ **Proposição 16.8.** *Uma sequência de em um espaço métrico* $\boldsymbol{M}$ *é convergente se, e somente se, todas suas subsequências são convergentes.*

$\square$ *Demonstração.* Suponhamos que $(p_n) \to p$ e seja $(p_{n_k})_{k\in\mathbb{N}}$ uma subsequência de $(p_n)_{n\in\mathbb{N}}$. Seja $\varepsilon \in \mathbb{R}$ tal que $\varepsilon > 0$. Como $(p_n) \to p$, existe $N \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N$, então $p_n \in \mathrm{B}(p; \varepsilon)$; como $(n_k)_{k\in\mathbb{N}}$ é estritamente crescente, existe $K \in \mathbb{N}$ tal que, para todo $k \in \mathbb{N}$, se $k \geq K$, então $n_k \geq N$. Mas então

$$k \geq K \Rightarrow n_k \geq N \Rightarrow p_{n_k} \in \mathrm{B}(p; \varepsilon)$$

e, portanto, $(p_{n_k}) \to p$. Reciprocamente, se toda subsequência de $(p_n)_{n\in\mathbb{N}}$ converge para p, $(p_n)_{n\in\mathbb{N}}$, em particular, é uma dessas subsequências e, portanto, $(p_n) \to p$. ■

$\vdash$ **Proposição 16.9.** *Toda sequência convergente em um espaço métrico* $\boldsymbol{M}$ *é limitada.*

□ *Demonstração.* Seja $(p_n)_{n\in\mathbb{N}}$ uma sequência de pontos em M tal que $(p_n) \to p$. Então, para $\varepsilon = 1$, existe $N \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N$, então $p_n \in \mathrm{B}\,(p; 1)$. Assim, seja $l \in \mathbb{R}$ tal que

$$l > \max(\{1\} \cup \{|p, p_n| : n \in [N]\}),$$

seque que, para todo $n \in \mathbb{N}$, $p_n \in \mathrm{B}\,(p; l)$ pois, se $0 \leq n \leq N$, $|p, p_n| < l$ pela definição de l e, se $n \geq N$, então $p_n \in \mathrm{B}\,(p; 1) \subseteq B_l(p)$, pois $1 < l$. Logo $(p_n)_{n\in\mathbb{N}}$ é limitada. ■

⊢ **Proposição 16.10.** *Sejam* $\boldsymbol{M}$ *um espaço métrico,* $C \subseteq M$ *um conjunto e* $p \in M$*. Então existe uma sequência de pontos em* C *que converge para* p *se, e somente se, para todo número real* $\varepsilon > 0$*,* $C \cap \mathrm{B}\,(p; \varepsilon) \neq \emptyset$*.*

□ *Demonstração.* Suponhamos que exista uma sequência $(p_n)_{n\in\mathbb{N}}$ de pontos em C tal que $(p_n) \to p$. Então, para todo número real $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N$, então $p_n \in \mathrm{B}\,(p; \varepsilon)$. Mas isso implica que $p_n \in C \cap \mathrm{B}\,(p; \varepsilon)$. Reciprocamente, suponhamos que, para todo número real $\varepsilon > 0$, $C \cap \mathrm{B}\,(p; \varepsilon) \neq \emptyset$. Então, em particular, para todo $n \in \mathbb{N}$, escolhamos $p_n \in C \cap \mathrm{B}\left(p; \frac{1}{n}\right)$. Assim, temos a sequência $(p_n)_{n\in\mathbb{N}}$. Para mostrar que $(p_n) \to p$, seja $\varepsilon \in \mathbb{R}$ tal que $\varepsilon > 0$. Então existe $N \in \mathbb{N}$ tal que $\frac{1}{N} \leq \varepsilon$. Mas isso implica que, para todo número natural $n \geq N$, $\frac{1}{n} \leq \frac{1}{N}$, e segue que

$$|p, p_n| < \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon$$

e, portanto, $(p_n) \to p$. ■

⊢ **Proposição 16.11.** *Sejam* $\boldsymbol{M}$ *um espaço métrico,* $p, q \in M$ *e* $(p_n)_{n\in\mathbb{N}}$ *e* $(q_n)_{n\in\mathbb{N}}$ *sequências em* M *que convergem para* p *e* q *respectivamente. Então a sequência* $(|p_n, q_n|)_{n\in\mathbb{N}}$ *em* $\mathbb{R}$ *converge para* $|p, q|$*.*

□ *Demonstração.* Para todo $n \in \mathbb{N}$, segue da desigualdade triangular que

$$|p_n, q_n| \leq |p_n, p| + |p, q| + |q, q_n|\,.$$

Seja $\varepsilon > 0$ um número real. Então existem $N_1, N_2 \in \mathbb{N}$ tais que

$$\forall n \in \mathbb{N} \qquad n \geq N_1 \Rightarrow |p, p_n| < \frac{\varepsilon}{2}$$

e

$$\forall n \in \mathbb{N} \qquad n \geq N_2 \Rightarrow |q, q_n| < \frac{\varepsilon}{2}.$$

Fazendo $N_3 := N_1 \vee N_2$, segue que

$$\forall n \in \mathbb{N} \qquad n \geq N_3 \Rightarrow |p_n, q_n| \leq |p_n, p| + |p, q| + |q, q_n| < |p, q| + \varepsilon;$$

ou seja, $|p_n, q_n| - |p, q| < \varepsilon$. Analogamente, achamos $N_6 \in \mathbb{N}$ tal que

$$\forall n \in \mathbb{N} \qquad n \geq N_3 \Rightarrow |p_n, q_n| \leq |p_n, p| + |p, q| + |q, q_n| < |p, q| + \varepsilon$$

e fazendo $N := N_3 \vee N_6$, segue que

$$\forall n \in \mathbb{N} \qquad n \geq N \Rightarrow |\,|p, q| - |p_n, q_n|\,| < \varepsilon,$$

o que mostra que $(|p_n, q_n|) \to |p, q|$ em $\mathbb{R}$.

■

16.2.3 Fecho e pontos aderentes

:⊢ **Definição 16.9.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Um *ponto aderente* a C é um ponto $p \in M$ para o qual existe uma sequência $(p_n)_{n\in\mathbb{N}}$ de pontos de C que converge para p. O *fecho* de C é o conjunto $\overline{C}$ de todos os pontos aderentes a C. Um *conjunto fechado* de $\boldsymbol{M}$ é um conjunto $F \subseteq M$ tal que $F = \overline{F}$.

⊢ **Proposição 16.12.** *Sejam $\boldsymbol{M}$ um espaço métrico e $F \subseteq M$. Então F é um conjunto fechado se, e somente se, $F^{\complement}$ é um conjunto aberto.*

□ *Demonstração.* Suponhamos que F é um conjunto fechado. Se $F^{\complement} = \emptyset$, Mas $\emptyset$ é aberto pois, caso contrário, existe $p \in \emptyset$ para o qual não há número real $\varepsilon > 0$ tal que $\mathrm{B}(p; \varepsilon) \subseteq \emptyset$, mas isso é absurdo. Se $F^{\complement} \neq \emptyset$, seja $p \in F^{\complement}$. Se não existe número real $\varepsilon > 0$ tal que $\mathrm{B}(p; \varepsilon) \subseteq F^{\complement}$, então, para todo número real $\varepsilon > 0$, $F \cap \mathrm{B}(p; \varepsilon) \neq \emptyset$. Mas isso implica que existe uma sequência $(p_n)_{n\in\mathbb{N}}$ de pontos em F tal que $(p_n) \to p$. Como F é fechado, isso implica $p \in F$, o que é uma contradição. Então existe número real $\varepsilon > 0$ tal que $\mathrm{B}(p; \varepsilon) \subseteq F^{\complement}$, e isso mostra que $F^{\complement}$ é aberto.

Reciprocamente, suponhamos que $F^{\complement}$ é aberto. Se $F = \emptyset$, então F é fechado. Se $F \neq \emptyset$, seja $(p_n)_{n\in\mathbb{N}}$ uma sequência em F que converge para $p \in M$. Suponhamos que $p \notin F$. Então $p \in F^{\complement}$ e, como $F^{\complement}$ é aberto, existe um número real $\varepsilon > 0$ tal que $\mathrm{B}(p; \varepsilon) \subseteq F^{\complement}$. Como $(p_n) \to p$, existe $N \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, se $n \geq N$, então $p_n \in \mathrm{B}(p; \varepsilon)$. Mas isso implica que $p_N \in \mathrm{B}(p; \varepsilon) \subseteq F^{\complement}$, o que é absurdo, pois $p_n \in F$. Portanto $p \in F$ e isso mostra que F é fechado. ■

⊢ **Proposição 16.13.** *Seja $\boldsymbol{M}$ um espaço métrico. Então, para todo $c \in M$ e para todo número real $r > 0$, a bola fechada $\bar{\mathrm{B}}(c; r)$ é um conjunto fechado.*

□ *Demonstração.* Basta notar que $\bar{\mathrm{B}}(c; r)^{\complement}$ é aberto. ■

16.2.4 Conjuntos densos

:⊢ **Definição 16.10.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Um conjunto *denso em* C é um conjunto $D \subseteq M$ tal que $C \subseteq \overline{D}$.

Isso que dizer que, para todo ponto de C, existe uma sequência em D que converge para esse ponto.

⊢ **Proposição 16.14.** *Sejam* $\boldsymbol{M} = (M, |\cdot,\cdot|)$ *um espaço métrico e* $C, D \subseteq M$ *conjuntos. Então* D *é denso em* C *se, e somente se, para todo conjunto aberto* A *de* $\boldsymbol{M}$, $A \cap C \neq \emptyset$ *implica* $A \cap D \neq \emptyset$.

□ *Demonstração.* Suponhamos que D é denso em C. Sejam A um conjunto aberto de $\boldsymbol{M}$ tal que $A \cap C \neq \emptyset$ e seja $p \in A \cap C$. Como D é denso em C e $p \in C$, existe uma sequência $(p_n)_{n\in\mathbb{N}}$ em D que converge para p. Como A é aberto e $p \in A$, existe um número real $\varepsilon > 0$ tal que $\mathrm{B}(p;\varepsilon) \subseteq A$. Então, como $(p_n) \to p$, existe um número natural N tal que, para todo natural $n \geq N$, $p_n \in \mathrm{B}(p;\varepsilon)$. Mas isso implica que $p_n \in A \cap D$.

Reciprocamente, suponhamos que, para todo conjunto aberto A de $\boldsymbol{M}$, $A \cap C \neq \emptyset$ implica $A \cap D \neq \emptyset$. Se $C = \emptyset$, então $C \subseteq \overline{D}$. Se $C \neq \emptyset$, seja $p \in C$. Para todo $n \in \mathbb{N}$, o conjunto $\mathrm{B}\left(p; \frac{1}{n}\right)$ é um conjunto aberto que contém p. Mas então $\mathrm{B}\left(p; \frac{1}{n}\right) \cap C \neq \emptyset$, o que implica $\mathrm{B}\left(p; \frac{1}{n}\right) \cap D \neq \emptyset$. Para cada $n \in \mathbb{N}$, escolhamos $p_n \in B_{\frac{1}{n}}(p) \cap D$. Assim, temos uma sequência $(p_n)_{n\in\mathbb{N}}$ de pontos em D que converge para p, pois, para todo número real $\varepsilon > 0$, existe um natural N tal que $\frac{1}{N} \leq \varepsilon$ e, então

$$\forall n \in \mathbb{N} \qquad n \geq N \Rightarrow \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon \Rightarrow p_n \in \mathrm{B}\left(p; \frac{1}{n}\right) \subseteq \mathrm{B}\left(p; \frac{1}{N}\right) \subseteq \mathrm{B}(p;\varepsilon).$$

Isso mostra que $p \in \overline{D}$ e, portanto, que $C \subseteq \overline{D}$. ■

⊢ **Proposição 16.15.** *Sejam* $\boldsymbol{M}_1$ *e* $\boldsymbol{M}_2$ *espaços métricos e* $f, g \colon M_1 \to M_2$ *funções contínuas. Então o conjunto*

$$F := \{p \in M_1 \mid f(p) = g(p)\}$$

é um conjunto fechado.

□ *Demonstração.* Se $F = \emptyset$, então F é fechado. Se $F \neq \emptyset$, seja $(p_n)_{n\in\mathbb{N}}$ uma sequência em F que converge para $p \in M_1$. Mostraremos que $p \in F$. Como f e g são contínuas em p, segue que

$$(f(p_n)) \to f(p) \ \text{ e } \ (g(p_n)) \to g(p).$$

Como $(p_n)_{n\in\mathbb{N}}$ é uma sequência em F, as sequências $(f(p_n))_{n\in\mathbb{N}}$ e $(g(p_n))_{n\in\mathbb{N}}$ são a mesma sequência e segue da unicidade do limite que $f(p) = g(p)$, o que mostra que $p \in F$ e que, portanto, F é um conjunto fechado. ■

⊢ **Proposição 16.16.** *Sejam $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ espaços métricos, $f, g : M_1 \to M_2$ funções contínuas e $C, D \subseteq M_1$ conjuntos tais que D é denso em C. Se $f|_D = g|_D$, então $f|_C = g|_C$.*

□ *Demonstração.* Pela proposição anterior, sabemos que $F := \{p \in M_1 : f(p) = g(p)\}$ é um conjunto fechado. Como $f|_D = g|_D$, então $D \subseteq F$. Mas isso significa que $\overline{D} \subseteq \overline{F} = F$ e, como D é denso em C, segue que $C \subseteq \overline{D} \subseteq F$ e, portanto, que $f|_C = g|_C$. ■

16.2.5 Conjuntos compactos

⊢ **Proposição 16.17.** *Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$. Se C é compacto, então é limitado.*

□ *Demonstração.* Seja $p \in M$ e consideremos a cobertura $\{\mathrm{B}(p; r) \mid r \in]0, \infty[\}$ de C. Pela compacidade, existe subcobertura finita $\{\mathrm{B}(p; r_0), \ldots, \mathrm{B}(p; r_{n-1})\}$ de C. Tomando $r := \max\{r_i \mid i \in [n]\}$, segue que $\mathrm{B}(p; r_i) \subseteq \mathrm{B}(p; r)$ para todo $i \in [n]$, logo $C \subseteq \mathrm{B}(p; r)$, o que implica que $\Theta(C) \leq 2r < \infty$. ■

A recíproca nem sempre é verdade. Nos espaços $\mathbb{R}^d$, $d \in \mathbb{N}$, vale que um conjunto é compacto se, e somente se, é fechado e limitado. Esse resultado é conhecido como Teorema de Heine-Borel. No entanto, isso não vale em qualquer espaço métrico[1].

▷ **Exercício 16.8.** *Seja $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$. Então C é compacto se, e somente se, é sequencialmente compacto.*

16.2.6 Continuidade

:⊢ **Definição 16.11.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ espaços métricos e $p \in M$. Uma função *contínua em p* é uma função $f\colon M \longrightarrow M'$ que satisfaz: para todo $\varepsilon \in]0, \infty[$, existe $\delta\,]0, \infty[$ tal que, para todo $x \in M$

$$x \in \mathrm{B}(p; \delta) \Rightarrow f(x) \in \mathrm{B}(f(p); \varepsilon).$$

Uma função *descontínua* em p é uma função que não é contínua em p.

Denotamos as bolas abertas em $\boldsymbol{M}$ e em $\boldsymbol{M}'$ por $\mathrm{B}(\cdot; \cdot)$, mas deve-se perceber que elas são relativas a métricas possivelmente diferentes.

[1] Para mais detalhes, conferir https://math.stackexchange.com/questions/674982/difference-between-closed-bounded-and-compact-sets.

$\vdash$ **Proposição 16.18.** *Sejam $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ espaços métricos, $f : M_1 \to M_2$ uma função e $p \in M_1$. Então f é contínua em p se, e somente se, para toda sequência $(p_n)_{n\in\mathbb{N}}$ de pontos em M_1 que converge para p, a sequência $(f(p_n))_{n\in\mathbb{N}}$ de pontos em M_2 converge para $f(p)$; ou seja*

$$\lim f(p_n) = f(\lim p_n).$$

$\square$ *Demonstração.* Suponhamos que f é contínua em p. Seja $(p_n)_{n\in\mathbb{N}}$ uma sequência de pontos em M_1 que converge para p. Seja um número real $\varepsilon > 0$. Como f é contínua, existe um número real $\delta > 0$ tal que $p_n \in \mathrm{B}(p;\delta)$ implica $f(p_n) \in \mathrm{B}(f(p);\varepsilon)$. Mas, como $(p_n) \to p$, existe $N \in \mathbb{N}$ tal que

$$\forall n \in \mathbb{N} \qquad n \geq N \Rightarrow p_n \in \mathrm{B}(p;\delta) \Rightarrow f(p_n) \in \mathrm{B}(f(p);\varepsilon)$$

o que mostra que $(f(p_n)) \to f(p)$.

Reciprocamente, suponhamos que, para toda sequência $(p_n)_{n\in\mathbb{N}}$ em M_1 que converge para p, a sequência $(f(p_n))_{n\in\mathbb{N}}$ converge para $f(p)$. Suponhamos, por absurdo, que f não é contínua em p. Então existe um número real $\varepsilon > 0$ tal que, para todo número real $\delta > 0$, existe $x \in M_1$ tal que $x \in \mathrm{B}(p;\delta)$, mas $f(x) \notin \mathrm{B}(f(p);\varepsilon)$. Vamos mostrar que isso implica que existe uma sequência $(p_n)_{n\in\mathbb{N}}$ em M_1 que converge para p, mas que a sequência $(f(p_n))_{n\in\mathbb{N}}$ não converge para $f(p)$; ou seja, que existe um número real $\varepsilon > 0$ tal que, para todo número natural N, existe $n \in \mathbb{N}$ tal que $n \geq N$, mas $f(p_n) \notin \mathrm{B}(f(p);\varepsilon)$. Seja $n \in \mathbb{N}$ e tomemos $\delta = \frac{1}{n}$. Então existe $x \in M_1$ tal que $x \in \mathrm{B}\left(p;\frac{1}{n}\right)$, mas $f(x) \notin \mathrm{B}(f(p);\varepsilon)$. Nomeando esse $x \in M_1$ de p_n, obtemos uma sequência $(p_n)_{n\in\mathbb{N}}$ que converge para p pois, para todo número real $\varepsilon' > 0$, existe um número natural $N \in \mathbb{N}$ tal que $\frac{1}{N} \leq \varepsilon'$ e isso implica que

$$\forall n \in \mathbb{N} \qquad n \geq N \Rightarrow \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon' \Rightarrow p_n \in \mathrm{B}\left(p;\frac{1}{n}\right) \subseteq \mathrm{B}\left(p;\frac{1}{N}\right) \subseteq \mathrm{B}(p;\varepsilon').$$

No entanto, $(f(p_n))_{n\in\mathbb{N}}$ é uma sequência que não converge para $f(p)$ pois, considerando o ε original tomado da descontinuidade de f, para todo número natural N, $f(p_N) \notin \mathrm{B}(f(p);\varepsilon)$ e isso contradiz a hipótese de que, para toda sequência $(p_n)_{n\in\mathbb{N}}$ em M_1 que converge para p, a sequência $(f(p_n))_{n\in\mathbb{N}}$ converge para $f(p)$. Portanto f é contínua. $\blacksquare$

$\vdash$ **Definição 16.12.** Sejam $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ espaços métricos, $D \subseteq M_1$ e $f : D \to M_2$ uma função. A função f é *contínua* em D se ela é contínua em todo ponto de D. Caso contrário, a função f é *descontínua* em D. Para $D = M_1$, dizemos simplesmente que f é contínua ou descontínua.

16.2.7 Ponto limite e conjunto derivado

:⊢ **Definição 16.13.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Um *ponto limite* (ou *ponto de acumulação*) de C é um ponto $p \in M$ para o qual existe uma sequência $(p_n)_{n\in\mathbb{N}}$ de pontos de $C \setminus \{p\}$ que converge para p. O *derivado* de C é o conjunto de todos os pontos limites de C.

Da definição, segue que $C' \subseteq \overline{C}$. A inclusão contrária caracteriza a seção a seguir.

:⊢ **Definição 16.14.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ um conjunto. Um *ponto isolado* de C é um ponto $p \in M$ que é um ponto aderente a C mas que não é um ponto limite de C.

Um ponto isolado de C é um ponto $p \in \overline{C} \setminus C'$.

16.2.8 Distância e bolas de conjuntos e separação métrica

:⊢ **Definição 16.15.** Sejam $\boldsymbol{M}$ um espaço métrico e $C, C' \subseteq M$. A *distância* entre C e C' é

$$|C, C'| := \inf \{|c, c'| \mid c \in C, c' \in C'\}.$$

Para todo $p \in M$, denotam-se $|C, p| := |C, \{p\}|$ e $|p, C| := |\{p\}, C|$.

Essa definição, em particular, estabelece a distância de pontos para conjuntos também. Existem outras definições de distâncias entre conjuntos, mas elas não serão tratadas aqui.

:⊢ **Definição 16.16.** Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ e $r \in \,]0, \infty[$. A *r-vizinhança* de C é o conjunto

$$\mathrm{B}(C; r) := \{p \in M \mid |C, p| < r\}.$$

A *r-vizinhança fechada* de C é o conjunto

$$\bar{\mathrm{B}}(C; r) := \{p \in M \mid |C, p| \leq r\}.$$

Note que excluímos $r = 0$ da definição pois basicamente teríamos $\mathrm{B}(C; 0) = \emptyset$ e $\bar{\mathrm{B}}(C; 0) = \overline{C}$.

▷ **Exercício 16.9.** *Sejam $\boldsymbol{M}$ um espaço métrico e $C \subseteq M$ e $r \in \,]0, \infty[$.*

1. *A vizinhança $\mathrm{B}(C; r)$ é um conjunto aberto e*

$$\mathrm{B}(C; r) = \bigcup_{c \in C} \mathrm{B}(c; r);$$

2. *Para todo* $r' \in]0, \infty[$ *tais que* $r \leq r'$,

$$C \subseteq \mathrm{B}(C; r) \subseteq \mathrm{B}(C; r');$$

3. *A vizinhança fechada* $\bar{\mathrm{B}}(C; r)$ *é um conjunto fechado e*

$$\bar{\mathrm{B}}(C; r) \supseteq \bigcup_{c \in C} \bar{\mathrm{B}}(c; r);$$

4. *Para todo* $r' \in]0, \infty[$ *tais que* $r \leq r'$,

$$\overline{C} \subseteq \bar{\mathrm{B}}(C; r) \subseteq \bar{\mathrm{B}}(C; r');$$

5. *Para* $r \in]0, \infty[$,

$$\bar{\mathrm{B}}(C; r) = \overline{\mathrm{B}(C; r)}$$

□ *Demonstração.* 1. Primeiro, mostramos a igualdade dos conjuntos. ($\supseteq$) Seja $p \in \bigcup_{c \in C} \mathrm{B}(c; r)$. Então existe $c \in C$ tal que $|c, p| < r$, o que implica que

$$|C, p| = \inf \{|c, p| \mid c \in C\} \leq |c, p| < r,$$

logo $p \in \mathrm{B}(C; r)$.
($\subseteq$) Seja $p \in \mathrm{B}(C; r)$. Então

$$|C, p| = \inf \{|c, p| \mid c \in C\} < r.$$

Isso significa que existe sequência $(c_n)_{n \in \mathbb{N}}$ em C tal que $\lim_{n \to \infty} |c, p| < r$, o que implica existe $n \in \mathbb{N}$ tal que $|c_n, p| < r$, logo $p \in \bigcup_{c \in C} \mathrm{B}(c; r)$.
Como as bolas $\mathrm{B}(c; r)$ são abertas, segue que $\mathrm{B}(C; r) = \bigcup_{c \in C} \mathrm{B}(c; r)$ é aberto.

■

:⊢ **Definição 16.17.** Seja $\boldsymbol{M}$ um espaço métrico. Conjuntos *metricamente separados* são conjuntos $C, C' \subseteq M$ tais que

$$|C, C'| > 0.$$

⊢ **Proposição 16.19.** *Sejam* $\boldsymbol{M}$ *um espaço métrico e* $C, C' \subseteq M$ *conjuntos metricamente separados. Então* C *e* C' *são separados por vizinhanças.*

□ *Demonstração.* Seja $\delta := |C, C'|$. Pela separação métrica, $\delta > 0$. Então $C \subseteq \mathrm{B}\left(C; \frac{\delta}{2}\right)$ e $C' \subseteq \mathrm{B}\left(C'; \frac{\delta}{2}\right)$ e $\mathrm{B}\left(C; \frac{\delta}{2}\right) \cap \mathrm{B}\left(C'; \frac{\delta}{2}\right) = \emptyset$, já que, se existe $p \in \mathrm{B}\left(C; \frac{\delta}{2}\right) \cap \mathrm{B}\left(C'; \frac{\delta}{2}\right)$, então existem $c \in C$ e $c' \in C'$ tais que $|c, p| < \frac{\delta}{2}$ e $|c', p| < \frac{\delta}{2}$, portanto

$$|c, c'| \leq |c, p| + |c', p| < \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

o que implica que $|C, C'| < \delta$, contradição. ■

Isso implica, em particular, que conjuntos metricamente separados são, além de separados por vizinhanças, separados, disjuntos e, claro, distintos.

⊢ **Proposição 16.20.** *Todo espaço métrico* $\boldsymbol{M}$ *é um espaço topológico normal.*

□ *Demonstração.* Sejam $F, F' \subseteq M$ fechados disjuntos. Mostraremos que F e F' são separados por vizinhanças, o que mostrará que o espaço é normal.

Para todo $f \in F$, existe $\delta_f \in]0, \infty[$ tal que

$$\mathrm{B}\left(f; \delta_f\right) \cap F' = \emptyset,$$

pois F' é fechado e $f \notin F'$. Definimos

$$V := \bigcup_{f \in F} \mathrm{B}\left(f; \frac{\delta_f}{2}\right).$$

Esse conjunto é uma vizinhança aberta de F. Analogamente, definimos V' uma vizinhança aberta de F'. Claramente $V \cap V' =$, pois caso contrário, se existe $p \in V \cap V'$, então existem $f \in F$e $f' \in F'$ tais que

$$p \in \mathrm{B}\left(f; \frac{\delta_f}{2}\right) \cap \mathrm{B}\left(f'; \frac{\delta_{f'}}{2}\right),$$

portanto se $\delta_f \leq \delta_{f'}$, então $f \in \mathrm{B}\left(f'; \frac{\delta_{f'}}{2}\right)$ e, se $\delta_{f'} \leq \delta_f$, então $f' \in \mathrm{B}\left(f; \frac{\delta_f}{2}\right)$, ambos contradições. ■

16.3 Estrutura uniforme

16.3.1 Sequências aproximantes

:⊢ **Definição 16.18.** Seja $\boldsymbol{M}$ um espaço métrico. Uma sequência *aproximante* em $\boldsymbol{M}$ é uma sequência $(p_n)_{n \in \mathbb{N}}$ de pontos em M tal que, para todo número real $\varepsilon > 0$, existe um número natural N satisfazendo

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m| < \varepsilon.$$

Essa sequências são conhecidas como *sequências de Cauchy.* O nome aproximante se dá pelo fato de que os termos da sequência ficam cada vez mais próximos entre si, e será adotado por ser mais intuitivo, embora não seja a nomenclatura padrão.

⊢ **Proposição 16.21.** *Toda sequência convergente em um espaço métrico* $\boldsymbol{M}$ *é aproximante.*

□ *Demonstração.* Seja $(p_n)_{n\in\mathbb{N}}$ uma sequência em M que converge para p. Seja $\varepsilon \in \mathbb{R}$ tal que $\varepsilon > 0$. Então $\frac{1}{2}\varepsilon > 0$ é um número real e segue que existe $N \in \mathbb{N}$ tal que, para todo número natural $n \geq N$, $p_n \in \mathrm{B}\left(p; \frac{1}{2}\varepsilon\right)$. Assim, segue que

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m| \leq |p_n, p| + |p, p_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

o que mostra que $(p_n)_{n\in\mathbb{N}}$ é uma sequência aproximante. ■

⊢ **Proposição 16.22.** *Toda sequência aproximante em um espaço métrico* $\boldsymbol{M}$ *que tem uma subsequência convergente é convergente.*

□ *Demonstração.* Seja $(p_{n_k})_{k\in\mathbb{N}}$ uma subsequência de $(p_n)_{n\in\mathbb{N}}$ que converge para p. Seja $\varepsilon > 0$ um número real. Como $(p_n)_{n\in\mathbb{N}}$ é uma sequência de Cauchy e $\frac{1}{2}\varepsilon > 0$ é um número real, existe um número natural N tal que

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m| < \frac{\varepsilon}{2}.$$

Como $(p_{n_k})_{k\in\mathbb{N}}$ é uma subsequência convergente, existe $K_1 \in \mathbb{N}$ tal que

$$\forall k \in \mathbb{N} \qquad k \geq K_1 \Rightarrow |p, p_{n_k}| < \frac{\varepsilon}{2}.$$

Como $(n_k)_{k\in\mathbb{N}}$ é uma sequência estritamente crescente, existe $K_2 \in \mathbb{N}$ tal que, para todo número natural $k \geq K_2$, $n_k \geq N$. Assim, tomando $K := \max\{K_1, K_2\}$, segue que, para todo número natural $n \in \mathbb{N}$, existe $k \in \mathbb{N}$ tal que $n_k \geq N$ e, pela desigualdade triangular, que

$$\forall n \in \mathbb{N} \qquad n \geq N \Rightarrow |p_n, p| \leq |p_n, p_{n_k}| + |p_{n_k}, p| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■

⊢ **Proposição 16.23.** *Toda sequência aproximante em um espaço métrico* $\boldsymbol{M}$ *é limitada.*

□ *Demonstração.* Seja $(p_n)_{n\in\mathbb{N}}$ uma sequência aproximante em $\boldsymbol{M}$. Então, para $\varepsilon = 1$, existe $N \in \mathbb{N}$ tal que

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m| < 1.$$

Definamos $P := \{p_n : n \in \mathbb{N}\}$. Então segue que

$$\begin{aligned}\Theta(P) &= \sup\{|p_n, p_m| \mid n, m \in \mathbb{N}\} \\ &= \max\{1 \cup \{|p_n, p_m| \mid 0 \leq n, m \leq N\}\} \in \mathbb{R},\end{aligned}$$

o que mostra que $(p_n)_{n\in\mathbb{N}}$ é limitada. ■

16.3.2 Continuidade uniforme

:⊢ **Definição 16.19.** Sejam $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ espaços métricos. Uma função *uniformemente contínua* é uma função $f : M_1 \to M_2$ tal que, para todo número real $\varepsilon > 0$, existe um número real $\delta > 0$ tal que

$$\forall p_1, p_2 \in M_1 \qquad |p_1, p_2|_1 < \delta \Rightarrow |f(p_1), f(p_2)|_2 < \varepsilon.$$

⊢ **Proposição 16.24.** *Sejam $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ espaços métricos, $f : M_1 \to M_2$ uma função uniformemente contínua e $(p_n)_{n\in\mathbb{N}}$ uma sequência aproximante em M_1. Então a sequência $(f(p_n))_{n\in\mathbb{N}}$ em M_2 é aproximante.*

□ *Demonstração.* Seja $\varepsilon > 0$ um número real. Da continuidade uniforme de f, existe um número real $\delta > 0$ tal que, para todo $p, p' \in M_1$, $|p, p'|_1 < \delta$ implica $|f(p), f(p')|_2 < \varepsilon$. Como $(p_n)_{n\in\mathbb{N}}$ é sequência aproximante, existe $N \in \mathbb{N}$ tal que

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m|_1 < \delta.$$

Mas, da continuidade uniforme de f, isso implica que

$$\forall n, m \in \mathbb{N} \qquad n, m \geq N \Rightarrow |p_n, p_m|_1 < \delta \Rightarrow |f(p_n), f(p_m)|_2 < \varepsilon,$$

e isso mostra que $(f(p_n))_{n\in\mathbb{N}}$ é uma sequência aproximante. ■

16.3.3 Espaços métricos completos

:⊢ **Definição 16.20.** Um espaço métrico *completo* é um espaço métrico em que todas sequências aproximantes convergem.

⊢ **Proposição 16.25.** *Seja $\boldsymbol{M}$ um espaço métrico. Todo subespaço completo de $\boldsymbol{M}$ é um conjunto fechado em $\boldsymbol{M}$.*

□ *Demonstração.* Sejam $\boldsymbol{C} \subseteq \boldsymbol{M}$ subespaço métrico completo e $(p_n)_{n\in\mathbb{N}}$ uma sequência convergente em C. Então $(p_n)_{n\in\mathbb{N}}$ é aproximante e, como C é completo, converge para um ponto em C, o que significa que C é fechado. ■

⊢ **Proposição 16.26.** *Sejam $\boldsymbol{M}$ um espaço métrico, $\boldsymbol{C} \subseteq \boldsymbol{M}$ um subespaço completo e $F \subseteq C$ um conjunto fechado em $\boldsymbol{M}$. Então $\boldsymbol{F}$ é completo.*

□ *Demonstração.* Seja $(p_n)_{n\in\mathbb{N}}$ uma sequência aproximante em F. Então $(p_n)_{n\in\mathbb{N}}$ é uma sequência aproximante em C e, como C é completo, $(p_n)_{n\in\mathbb{N}}$ converge. Porém, como F é fechado, então $(p_n)_{n\in\mathbb{N}}$ converge para um ponto em F, o que mostra que F é completo. ■

$\vdash$ **Teorema 16.27.** *Seja $\boldsymbol{M}$ um espaço métrico. Então $\boldsymbol{M}$ é completo se, e somente se, para toda sequência decrescente $(F_n)_{n\in\mathbb{N}}$ de conjuntos não vazios e fechados em $\boldsymbol{M}$ tais que $(\Theta(F_n))_{n\in\mathbb{N}} \to 0$ em $\mathbb{R}$, vale que*

$$\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset.$$

$\vdash$ **Teorema 16.28.** *Sejam $\boldsymbol{M}_1$ um espaço métrico, $\boldsymbol{M}_2$ espaço métrico completo, $D \subseteq M_1$ um conjunto denso em M_1 e $f : D \to M_2$ uma função uniformemente contínua. Então f tem uma única extensão para uma função uniformemente contínua $f^* : M_1 \to M_2$. Ainda, se f é uma isometria, então f^* é uma isometria.*

$\square$ *Demonstração.* Seja $p \in M_1$. Como D é denso em M_1, existe uma sequência $(p_n)_{n\in\mathbb{N}}$ em D que converge para p. Como $(p_n)_{n\in\mathbb{N}}$ é convergente, é uma sequência de Cauchy e, como f é uniformemente contínua em D, segue que $(f(p_n))_{n\in\mathbb{N}}$ é uma sequência de Cauchy em M_2. Mas M_2 é completo, o que implica que $(f(p_n))_{n\in\mathbb{N}}$ converge para um ponto $p' \in M_2$. Definimos, portanto, a função f^* em p como $f^*(p) = p'$. Precisamos mostrar que f^* independe da escolha da sequência em D que converge para p. Se $(q_n)_{n\in\mathbb{N}}$ é uma sequência em D que converge para p, definamos a sequência $(r_n)_{n\in\mathbb{N}}$ em D por

$$r_n := \begin{cases} p_n & \text{se } n = 2k \\ q_n & \text{se } n = 2k+1. \end{cases}$$

A sequência $(r_n)_{n\in\mathbb{N}}$ converge para p e, portanto, é uma sequência de Cauchy. A continuidade uniforme de f implica que a sequência $(f(r_n))_{n\in\mathbb{N}}$ é de Cauchy e, portanto, como $(f(p_n))_{n\in\mathbb{N}} = (f(r_{2k}))_{k\in\mathbb{N}}$ é uma subsequência que converge para p', a sequência $(f(r_n))_{n\in\mathbb{N}}$ converge para p', o que implica que a subsequência $(f(q_n))_{n\in\mathbb{N}} = (f(r_{2k+1}))_{k\in\mathbb{N}}$ converge para p'. Assim, mostramos que f^* está bem definida. Claramente, se $p \in D$, então $f(p) = f^*(p)$, pois, como D é denso em M_1, se $(p_n)_{n\in\mathbb{N}}$ é uma sequência em D que converge para p, então, como f é contínua, segue que $f(p_n) \to f(p)$, o que mostra que $f^*(p) = f(p)$.

Agora, devemos mostrar que f^* é uniformemente contínua. Seja $\varepsilon > 0$ um número real, então $\frac{1}{2}\varepsilon > 0$ é um número real e, como f é uniformemente contínua, existe número real $\delta > 0$ tal que

$$\forall p, p' \in M_1 \qquad |p, p'|_1 < \delta \Rightarrow |f(p), f(p')|_2 < \frac{\varepsilon}{2}.$$

Assim, sejam $p, q \in M_1$ tais que $|p, q|_1 < \delta$. Queremos mostrar que $|f(p), f(p')|_2 < \varepsilon$. Sejam $(p_n)_{n\in\mathbb{N}}$ e $(q_n)_{n\in\mathbb{N}}$ sequências que convergem para p e q, respectivamente. Então $|p_n, q_n|_1 \to |p, q|_1$ em $\mathbb{R}$.

...

A unicidade de f^* ocorre pois, se existem f^* e f'^* uniformemente contínuas que estendem f, como D é denso em M_1 e $f^*|_D = f'^*|_D$, segue que $f^* = f'^*$.

Por fim, mostramos que a isometria se preserva... ■

:⊢ **Definição 16.21.** Seja $\boldsymbol{M}_1$ um espaço métrico. Um *completamento* de $\boldsymbol{M}$ é um espaço métrico $\boldsymbol{M}_2$ completo tal que M_1 é denso em M_2.

▷ **Exercício 16.10.** *Seja $\boldsymbol{M}$ um espaço métrico e $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ completamentos de $\boldsymbol{M}$. Então existe uma isometria entre $\boldsymbol{M}_1$ e $\boldsymbol{M}_2$ que é a função identidade quando restrita a M.*

⊢ **Proposição 16.29.** *Sejam $K \subseteq M$ compacto e $f : M \to \bar{M}$ contínua. Então f é uniformemente contínua.*

□ *Demonstração.* Suponhamos, por absurdo, que f não é uniformemente contínua. Então existem $\varepsilon > 0$ e $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ sequências em K tais que

$$|x_n, y_n| < \frac{1}{n} \quad \text{e} \quad |f(x_n), f(y_n)| \geq \varepsilon.$$

Como K é compacto, existem subsequências $(x_{n_k})_{k\in\mathbb{N}}$ e $(y_{n_k})_{k\in\mathbb{N}}$ convergindo a $x \in K$ com $|f(x_{n_k}), f(y_{n_k})| \geq \varepsilon$. Por continuidade de f, existe $\delta > 0$ tal que, se $x_{n_k}, y_{n_k} \in B(x, \delta)$, então $|f(x_{n_k}), f(x)| < \frac{\varepsilon}{2}$ e $|f(y_{n_k}), f(x)| < \frac{\varepsilon}{2}$. Pela desigualdade triangular, temos um absurdo. ■

16.3.4 Limitação uniforme (ou total)

Sejam $(M, |\cdot,\cdot|)$ um espaço métrico e $\mathcal{T}$ a topologia induzida por $|\cdot,\cdot|$. A métrica limitada $|\cdot,\cdot| \wedge 1$ induz a mesma topologia $\mathcal{T}$ que $|\cdot,\cdot|$ sobre M, mas com respeito a $|\cdot,\cdot| \wedge 1$ todos conjuntos são limitados, enquanto que com respeito a $|\cdot,\cdot|$ isso nem sempre é verdade (somente nos casos em que o espaço inteiro é limitado). Isso mostra que o conceito de limitação não está sempre relacionado à topologia do espaço.

A convergência de sequências em $(M, \mathcal{T})$ independe da métrica que a gera e, portanto, concluímos que não pode existir um análogo ao teorema de Bolzano - Weierstrass usando o conceito de limitação, pois todas sequências em $(M, d \wedge 1)$ são limitadas, mas não necessariamente têm subsequência convergente. Definiremos a seguir um conceito distinto de limitação em espaços métricos que está mais relacionado à estrutura uniforme do espaço.

:⊢ **Definição 16.22.** Um espaço métrico *uniformemente limitado* (ou *totalmente limitado*) é um espaço métrico $\boldsymbol{M}$ em que, para todo $\varepsilon \in]0, \infty[$, existe uma ε-cobertura finita de M. Um conjunto *uniformemente limitado* é um conjunto que é uniformemente limitado com a estrutura métrica induzida.

⊢ **Proposição 16.30.** *Seja* $\boldsymbol{M}$ *um espaço métrico.*

1. $\boldsymbol{M}$ *é uniformemente limitado se, e somente se, para todo* $\varepsilon \in]0, \infty[$*, existem* $p_0, \cdots, p_{n-1}$ *tais que*
$$M \subseteq \bigcup_{i \in [n]} \mathrm{B}\left(p; \varepsilon\right);$$
2. *Se* $\boldsymbol{M}$ *é uniformemente limitado, então é limitado.*

□ *Demonstração.* 1. ($\Rightarrow$) Seja $\varepsilon \in]0, \infty[$. Como M é uniformemente limitado, existe uma $\frac{\varepsilon}{2}$-cobertura finita $\{C_i\}_{i \in [n]}$ de M. Tome, para cada $i \in [n]$, $c_i \in C_i$. Então $\Theta(\mathrm{B}\left(c_i; \frac{\varepsilon}{2}\right)) \leq \varepsilon$ e $C_i \subseteq \mathrm{B}\left(c_i; \frac{\varepsilon}{2}\right)$, portanto
$$M \subseteq \bigcup_{i \in [n]} C_i \subseteq \bigcup_{i \in [n]} \mathrm{B}\left(c_i; \frac{\varepsilon}{2}\right).$$
($\Leftarrow$) Reciprocamente, seja $\varepsilon \in]0, \infty[$. Existem existem $p_0, \cdots, p_{n-1}$ tais que
$$M \subseteq \bigcup_{i \in [n]} \mathrm{B}\left(p_i; \frac{\varepsilon}{2}\right);$$
Como, para todo $i \in [n]$, $\Theta(\mathrm{B}\left(p; \frac{\varepsilon}{2}\right)) \leq \varepsilon$, segue que $(\mathrm{B}\left(p_i; \frac{\varepsilon}{2}\right))_{i \in [n]}$ é uma ε-cobertura de M, portanto M é uniformemente limitado.

2. Seja $(C_i)_{i \in [n]}$ uma cobertura 1-precisa de M. Então, como para todo $i \in [n]$, $\Theta(C_i) < \infty$, e a cobertura é finita, $\Theta(M) < \infty$.

■

▷ **Exercício 16.11.** *Seja* $d \in \mathbb{N}$ *e consideremos* $\mathbb{R}^d$ *com a métrica reta usual. Um subconjunto de* $\mathbb{R}^d$ *é limitado se, e somente se, é uniformemente limitado.*

⊢ **Lema 16.31.** *Sejam* $\boldsymbol{M}$ *um espaço métrico e* $(x_n)_{n \in \mathbb{N}}$ *uma sequência em* M*.*

1. *Se* $(x_n)_{n \in \mathbb{N}}$ *é aproximante, então sua imagem* $x(\mathbb{N}) \subseteq M$ *é uniformemente limitada;*
2. *Se* $x(\mathbb{N})$ *é uniformemente limitada,* $(x_n)_{n \in \mathbb{N}}$ *tem subsequência aproximante.*

□ *Demonstração.* 1. Seja $\varepsilon \in]0, \infty[$. Como $(x_n)_{n \in \mathbb{N}}$ é aproximante, existe $N \in \mathbb{N}$ tal que, para todos $n, n' \in \mathbb{N}$, se $n \geq N$ e $n' \geq N$, então $|x_n, x_{n'}| < \varepsilon$, portanto
$$\Theta\left\{x_n \mid n \in \mathbb{N}, n \geq N\right\} \leq \varepsilon,$$
portanto
$$(\{x_0\}, \cdots, \{x_{N-1}\}, \{x_n \mid n \in \mathbb{N}, n \geq N\})$$
é uma cobertura ε-precisa finita de $x(\mathbb{N})$.

2. Se $x(\mathbb{N})$ for finito, então $(x_n)_{n\in\mathbb{N}}$ tem subsequência constante, logo aproximante. Suponhamos o caso em que $x(\mathbb{N})$ é infinito.

■

⊢ **Proposição 16.32.** *Seja $\boldsymbol{M}$ um espaço métrico. O espaço $\boldsymbol{M}$ é uniformemente limitado se, e somente se, toda sequência tem subsequência aproximante.*

⊢ **Proposição 16.33.** *Seja $\boldsymbol{M}$ um espaço métrico. O espaço $\boldsymbol{M}$ é compacto se, e somente se, é completo e uniformemente limitado.*

16.4 Funções métricas

16.4.1 Funções controladas

Neste seção, definiremos funções de um espaço métrico para outro que de certa forma preserva alguma informação da estrutura métrica do espaço. Essas funções não preservam o valor da distância entre dois pontos, mas são limitadas de em um sentido específico, o que garante sua continuidade (uniforme) e traduz alguma propriedades de um espaço para o outro. No que segue, não distinguiremos a notação das distâncias em espaços métricos diferentes.

:⊢ **Definição 16.23.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ espaços métricos. Uma *função controlada*[2] de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é uma função $f\colon M \longrightarrow M'$ tal que, para algum $c \in [0,\infty[$, todos $p, p' \in M$ satisfazem

$$|f(p), f(p')| \leq c\,|p, p'|\,.$$

A *distorção* (ou *norma de distorção*) de f é

$$\langle\!\langle f \rangle\!\rangle := \inf\left\{c \in [0,\infty[\mid \forall_{p,p'\in M}\, |f(p), f(p')| \leq c\,|p, p'|\right\}.$$

A distorção de uma função controlada f é portanto o menor valor tal que vale a desigualdade entre as distâncias da definição. Vale que

$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle\,|p, p'|\,.$$

Além disso, a distorção de f satisfaz a seguinte definição em termos do supremo:

$$\langle\!\langle f \rangle\!\rangle = \sup\left\{\frac{|f(p), f(p')|}{|p, p'|}\,\middle|\, p, p' \in M \wedge p \neq p'\right\}.$$

Note que, caso $\#\,(M) \leq 1$, o conjunto nesse supremos é vazio, e nesse caso devemos ressaltar que o supremo é tomado sobre $[0,\infty[$, portanto o supremo de $\emptyset$ é 0.

[2] Essas funções são conhecidas na literatura como funções 'Lipschitz contínuas'.

▷ **Exercício 16.12.** *Seja* $\boldsymbol{M}$ *um espaço métrico. A identidade* $\mathrm{I}\colon M \longrightarrow M$ *é função controlada e* $\langle\!\langle \mathrm{I} \rangle\!\rangle = 1$.

⊢ **Proposição 16.34.** *Sejam* $\boldsymbol{M}, \boldsymbol{M}'$ *e* $\boldsymbol{M}''$ *espaços métricos e* $f\colon M \longrightarrow M'$ *e* $f'\colon M' \longrightarrow M''$ *funções controladas. Então* $f' \circ f\colon M \longrightarrow M''$ *é uma função controlada e*

$$\langle\!\langle f' \circ f \rangle\!\rangle \leq \langle\!\langle f' \rangle\!\rangle \langle\!\langle f \rangle\!\rangle.$$

□ *Demonstração.* Para todos $p, p' \in M$,

$$|f' \circ f(p), f' \circ f(p')| \leq \langle\!\langle f' \rangle\!\rangle \, |f(p), f(p')| \leq \langle\!\langle f' \rangle\!\rangle \langle\!\langle f \rangle\!\rangle \, |p, p'| \,,$$

o que mostra que $f' \circ f$ é controlada e $\langle\!\langle f' \circ f \rangle\!\rangle \leq \langle\!\langle f' \rangle\!\rangle \langle\!\langle f \rangle\!\rangle$. ■

⊢ **Proposição 16.35.** *Sejam* $\boldsymbol{M}$ *e* $\boldsymbol{M}'$ *espaços métricos e* $f\colon M \longrightarrow M'$ *uma função controlada. Então* f *é uniformemente contínua.*

□ *Demonstração.* Se $\langle\!\langle f \rangle\!\rangle = 0$, a demonstração é óbvia. Se $\langle\!\langle f \rangle\!\rangle \neq 0$, seja $\varepsilon > 0$. Tomando $\delta = \frac{\varepsilon}{\langle\!\langle f \rangle\!\rangle}$, segue que, para todos $p, p' \in M$, se $|p, p'|_0 \leq \delta$, então

$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle \, |p, p'| \leq \langle\!\langle f \rangle\!\rangle \frac{\varepsilon}{\langle\!\langle f \rangle\!\rangle} = \varepsilon.$$ ■

⊢ **Proposição 16.36.** *Sejam* $\boldsymbol{M}$ *e* $\boldsymbol{M}'$ *espaços métricos e* $f\colon M \longrightarrow M'$ *uma função controlada.*

1. *f é constante se, e somente se,* $\langle\!\langle f \rangle\!\rangle = 0$*;*
2. *Se* $\#(M) > 1$ *e f é injetiva,* $\langle\!\langle f \rangle\!\rangle > 0$*;*
3. *Se* $\#(M') > 1$ *e f é sobrejetiva,* $\langle\!\langle f \rangle\!\rangle > 0$.

□ *Demonstração.*

1. Se f é constante, então para todos $p, p' \in M$ vale $f(p) = f(p')$, portanto
$$|f(p), f(p')| = 0 \leq 0 \, |p, p'| \,,$$
o que mostra que $\langle\!\langle f \rangle\!\rangle = 0$.
Reciprocamente, se $\langle\!\langle f \rangle\!\rangle = 0$, então para todos $p, p' \in M$ vale
$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle \, |p, p'| = 0,$$
portanto $f(p) = f(p')$.
2. Como $\#(M) > 1$, existem $p, p' \in M$ tais que $p \neq p'$; como f é injetiva, segue que $f(p) \neq f(p')$, portanto f não é constante; do primeiro item segue que $\langle\!\langle f \rangle\!\rangle > 0$.

3. Como $\#(M') > 1$, existem $q, q' \in M'$ tais que $q \neq q'$; como f é sobrejetiva, $f(M) = M'$, logo $\#(f(M)) > 1$, o que significa que f não é constante; do primeiro item segue que $\langle\langle f \rangle\rangle > 0$.

■

⊢ **Proposição 16.37.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M$ função controlada.*

1. *Se $\#(M) > 1$ e f é injetiva com inversa à esquerda $f^{-1}\colon f(M) \longrightarrow M$ controlada, então*
$$1 \leq \langle\langle f \rangle\rangle \langle\langle f^{-1} \rangle\rangle.$$
2. *Se $\#(M') > 1$ e f é sobrejetiva com inversa à direita $f^{-1}\colon M' \longrightarrow M$ controlada, então*
$$1 \leq \langle\langle f \rangle\rangle \langle\langle f^{-1} \rangle\rangle.$$

□ *Demonstração.* 1. Sejam $p, p' \in M$ tais que $p \neq p'$ e $q := f(p)$ e $q' := f(p')$. Como f e f^{-1} são controladas e $f^{-1}(q) = p$ e $f^{-1}(q') = p'$, segue que

$$\begin{aligned} |p, p'| &= \left| f^{-1}(q), f^{-1}(q') \right| \\ &\leq \langle\langle f^{-1} \rangle\rangle \, |q, q'| \\ &= \langle\langle f^{-1} \rangle\rangle \, |f(p), f(p')| \\ &\leq \langle\langle f^{-1} \rangle\rangle \langle\langle f \rangle\rangle \, |p, p'| \, . \end{aligned}$$

Como $|p, p'| \neq 0$, segue que $1 \leq \langle\langle f^{-1} \rangle\rangle \langle\langle f \rangle\rangle$.

2. Segue do item anterior, pois f^{-1} é injetiva com inversa à esquerda f. ■

:⊢ **Definição 16.24.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos. Um *mergulho métrico* de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é uma função $f\colon M \longrightarrow M'$ tal que, para algum $c \in \,]0, \infty[$, f satisfaz que, para todos $p, p' \in M$,

$$c^{-1} \, |p, p'| \leq |f(p), f(p')| \leq c \, |p, p'| \, .$$

A proposição a seguir dá uma caracterização alternativa de mergulho métrico.

⊢ **Proposição 16.38.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$. Então f é mergulho métrico se, e somente se, f é função controlada injetiva e sua inversa à esquerda $f^{-1}\colon f(M) \longrightarrow M$ é controlada.*

□ *Demonstração.* Notemos que se $\#(M) \leq 1$, a demonstração é trivial, portanto supomos $\#(M) > 1$.

- ($\Rightarrow$) Seja $c \in]0, \infty[$ tal que valem as desigualdades na definição de mergulho métrico. Segue direto que f é controlada. Para todos $p, p' \in M$ tais que $p \neq p'$, vale $|p, p'| > 0$, logo de $0 < c^{-1} |p, p'| \leq |f(p), f(p')|$ segue que $|f(p), f(p')| > 0$, o que implica $f(p) \neq f(p')$, portanto f é injetiva. Ainda, temos que $|p, p'| \leq c\,|f(p), f(p')|$, logo para todos $q, q' \in f(M)$, existem $p, p' \in M$ tais que $q = f(p)$ e $q' = f(p')$, portanto
$$\begin{aligned}\left|f^{-1}(q), f^{-1}(q')\right| &= \left|f^{-1}(f(p)), f^{-1}(f(p'))\right| \\ &= |p, p'| \\ &\leq c\,|f(p), f(p')| \\ &= c\,|q, q'|\,,\end{aligned}$$
o que mostra que f^{-1} é controlada.
- ($\Leftarrow$) Como f é função controlada injetiva e $\#(M) > 1$, segue de 16.36 que $\langle\!\langle f \rangle\!\rangle > 0$. A injetividade de f implica que f^{-1} é sobrejetiva e que $\#(f(M)) > 0$, segue de 16.36 que $\langle\!\langle f^{-1} \rangle\!\rangle > 0$. Definindo $c := \langle\!\langle f \rangle\!\rangle \vee \langle\!\langle f^{-1} \rangle\!\rangle$, segue que $c > 0$, $c \geq \langle\!\langle f \rangle\!\rangle$ e $c^{-1} \leq \langle\!\langle f^{-1} \rangle\!\rangle^{-1}$.
Como f é função controlada, então, para todos $p, p' \in M$,
$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle\, |p, p'| \leq c\,|p, p'|\,.$$
Como a inversa à esquerda f^{-1} é função controlada, então, para todos $p, p' \in M$,
$$\begin{aligned}c^{-1}\,|p, p'| &= c^{-1}\left|f^{-1}(f(p)), f^{-1}(f(p'))\right| \\ &\leq c^{-1}\langle\!\langle f^{-1} \rangle\!\rangle\, |f(p), f(p')| \\ &\leq c^{-1} c\,|f(p), f(p')| \\ &= |f(p), f(p')|\,.\end{aligned}$$
Portanto
$$c^{-1}\,|p, p'| \leq |f(p), f(p')| \leq c\,|p, p'|\,. \qquad \blacksquare$$

$\vdash$ **Proposição 16.39.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ um mergulho métrico. Então f é mergulho topológico.*

$\square$ *Demonstração.* Segue direto de 16.38 que f é contínua, injetiva e sua inversa à esquerda é contínua, portanto $f\colon M \longrightarrow f(M)$ é homeomorfismo. $\blacksquare$

16.4.2 Homometrias e isometrias

$\dashv$ **Definição 16.25.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos. Uma *homometria*[3] de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é uma função controlada $f\colon M \longrightarrow M'$ tal que $\langle\!\langle f \rangle\!\rangle \leq 1$. O conjunto das

[3]Essas funções são também conhecidas como funções métricas, funções não expansoras, entre outros. O nome homometria faz referência a uma relação que essas funções têm com as isometrias,

homometrias de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é denotado $\mathscr{C}_{|\cdot|}(\boldsymbol{M}, \boldsymbol{M}')$.

Uma homometria f satisfaz portanto

$$|f(p), f(p')| \leq |p, p'|\,.$$

⊢ **Proposição 16.40.** *Sejam $\boldsymbol{M}, \boldsymbol{M}'$ e $\boldsymbol{M}''$ espaços métricos e $f\colon M \longrightarrow M'$ e $f'\colon M' \longrightarrow M''$ homometrias. Então $f' \circ f\colon M \longrightarrow M''$ é uma homometria.*

□ *Demonstração.* Segue de 16.34 que se $\langle\!\langle f \rangle\!\rangle \leq 1$ e $\langle\!\langle f' \rangle\!\rangle \leq 1$, então $\langle\!\langle f' \circ f \rangle\!\rangle \leq 1$, portanto $f' \circ f$ é uma homometria. ■

:⊢ **Definição 16.26.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos. Uma *isometria local* (ou *imersão isométrica*) de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é uma função $f\colon M \longrightarrow M'$ que satisfaz, para todos $p, p' \in M$,

$$|f(p), f(p')| = |p, p'|\,.$$

Uma *isometria* é isometria local bijetiva.

⊢ **Proposição 16.41.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ uma isometria local. Então f é injetiva.*

□ *Demonstração.* Sejam $p, p' \in M$ tais que $p \neq p'$. Então $|p, p'|_0 \neq 0$, logo

$$|f(p), f(p')| = |p, p'| \neq 0,$$

o que implica $f(p) \neq f(p')$. ■

⊢ **Proposição 16.42.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ uma homometria injetiva cuja inversa à esquerda é homometria. Então f é uma isometria local.*

□ *Demonstração.* Sejam $p, p' \in M$. Então, como f é homometria,

$$|f(p), f(p')| \leq |p, p'|$$

e, como f^{-1} é homometria,

$$|p, p'| = \left|f^{-1} \circ f(p), f^{-1} \circ f(p')\right| \leq |f(p), f(p')|\,;$$

portanto $|p, p'| = |f(p), f(p')|$. ■

⊢ **Proposição 16.43.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ homometria. A função f é isometria se, e somente se, é invertível e sua inversa é homometria.*

---

pois elas representam morfismos de espaços métricos cujos respectivos isomorfismos são as isometrias.

□ *Demonstração.* Suponhamos que f é isometria. Então f é bijetiva e, portanto, invertível. Sua inversa satisfaz, para todos $p, p' \in M'$,

$$\left|f^{-1}(p), f^{-1}(p')\right| = \left|f(f^{-1}(p)), f(f^{-1}(p'))\right| = |p, p'|\,.$$

Portanto f^{-1} é isometria local, logo homometria.

Reciprocamente, suponhamos que f é invertível e sua inversa é homometria. Segue da proposição anterior que f é isometria local e, como é bijetiva, é isometria. ■

16.4.3 Contrações

:⊢ **Definição 16.27.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ espaços métricos. Uma *contração* de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é uma função controlada $f\colon M \longrightarrow M'$ tal que $\langle\!\langle f \rangle\!\rangle < 1$.

Uma contração f satisfaz portanto

$$|f(p), f(p')| < |p, p'|\,.$$

Em particular, uma contração é uma homometria.

⊢ **Proposição 16.44.** *Sejam $\boldsymbol{M}, \boldsymbol{M}'$ e $\boldsymbol{M}''$ espaços métricos e $f\colon M \longrightarrow M'$ e $f'\colon M' \longrightarrow M''$ contrações. Então $f' \circ f\colon M \longrightarrow M''$ é uma contração.*

□ *Demonstração.* Segue de 16.34 que se $\langle\!\langle f \rangle\!\rangle < 1$ e $\langle\!\langle f' \rangle\!\rangle < 1$, então $\langle\!\langle f' \circ f \rangle\!\rangle < 1$, portanto $f' \circ f$ é uma contração. ■

⊢ **Proposição 16.45** (Ponto fixo para contrações)**.** *Sejam $\boldsymbol{M}$ um espaço métrico completo não vazio e $f\colon M \longrightarrow M$ uma contração. Existe único ponto fixo $\bar{p} \in M$ para f e, para todo $p \in M$,*

$$\lim_{n\to\infty} f^n(p) = \bar{p}.$$

□ *Demonstração.* Mostremos por indução que, para todos $p \in M$ e $n \in \mathbb{N}$,

$$\left|f^n(p), f^{n+1}(p)\right| \leq \langle\!\langle f \rangle\!\rangle^n \, |p, f(p)|\,.$$

Claramente, para $n = 0$ isso claramente vale. Agora, suponhamos que a desigualdade valha para $n = k$ e mostremos que ela vale para $n = k + 1$. Como f é contração,

$$\left|f^k(p), f^{k+1}(p)\right| \leq \langle\!\langle f \rangle\!\rangle \left|f^{k-1}(p), f^k(p)\right| \leq \langle\!\langle f \rangle\!\rangle \langle\!\langle f \rangle\!\rangle^{k-1} \, |p, f(p)| = \langle\!\langle f \rangle\!\rangle^k \, |p, f(p)|\,.$$

Agora, notemos que, para todos $n, p \in \mathbb{N}$, segue da desigualdade triangular generalizada que

$$\begin{aligned}
\left|f^n(p), f^{n+p}(p)\right| &\leq \mathop{+}_{i=0}^{p-1} \left|f^{n+i}(p), f^{n+i+1}(p)\right| \\
&\leq \mathop{+}_{i=0}^{p-1} \langle\!\langle f \rangle\!\rangle^{n+i} \left|p, f(p)\right| \\
&= \langle\!\langle f \rangle\!\rangle^{n} \frac{1 - \langle\!\langle f \rangle\!\rangle^{p}}{1 - \langle\!\langle f \rangle\!\rangle} \left|p, f(p)\right| \\
&\leq \frac{\langle\!\langle f \rangle\!\rangle^{n}}{1 - \langle\!\langle f \rangle\!\rangle} \left|p, f(p)\right|,
\end{aligned}$$

pois $\langle\!\langle f \rangle\!\rangle \geq 0$ implica $1 - \langle\!\langle f \rangle\!\rangle^p < 1$. Como $\langle\!\langle f \rangle\!\rangle < 1$, então $\lim_{n\to\infty} \frac{\langle\!\langle f \rangle\!\rangle^n}{1-\langle\!\langle f \rangle\!\rangle} = 0$, portanto, para todos $n, p \in \mathbb{N}$,

$$\lim_{n\to\infty} \left|f^n(p), f^{n+p}(p)\right| = 0,$$

o que mostra que $(f^n(p))_{n\in\mathbb{N}}$ é uma sequência aproximante e, como $\boldsymbol{M}$ é completo, converge para $\bar{p} \in M$. Como f é contínua,

$$f(\bar{p}) = f\left(\lim_{n\to\infty} f^n(p)\right) = \lim_{n\to\infty} f^{n+1}(p) = \bar{p},$$

esse ponto $\bar{p}$ é um ponto fixo. Para mostrarmos que $\bar{p}$ é único, suponhamos que p é ponto fixo de f. Então

$$|\bar{p}, p| = |f(\bar{p}), f(p)| \leq \langle\!\langle f \rangle\!\rangle \, |\bar{p}, p|$$

e como $\langle\!\langle f \rangle\!\rangle < 1$ isso implica $|\bar{p}, p| = 0$, logo $\bar{p} = p$. ■

Esse resultado pode parecer trivial por sua demonstração, mas suas consequências na matemática são enormes. Como consequência dessa proposição seguem por exemplo o teorema da função inversa do cálculo diferencial e o teorema de existência e unicidade de equações diferenciais.

Como corolário direto da proposição 16.45, temos a seguinte proposição.

⊢ **Proposição 16.46.** *Sejam $\boldsymbol{M}$ um espaço métrico completo não vazio e $f\colon M \longrightarrow M$ uma função contínua tal que, para algum $m \in \mathbb{N} \smallsetminus \{0\}$, f^m é contração. Existe único ponto fixo $\bar{p} \in M$ para f e, para todo $p \in M$,*

$$\lim_{n\to\infty} f^n(p) = \bar{p}.$$

□ *Demonstração.* Exercício. ■

16.4.4 Semelhanças

:⊢ **Definição 16.28.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos. Uma *semelhança local* ou (*imersão semelhante*) de $\boldsymbol{M}$ para $\boldsymbol{M'}$ é uma função $f\colon M \longrightarrow M'$ tal que, para algum $c \in [0, \infty[$, f satisfaz que, para todos $p, p' \in M$,

$$|f(p), f(p')| = c\,|p, p'|\,.$$

Uma *semelhança* é semelhança local bijetiva.

⊢ **Proposição 16.47.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ uma semelhança local. Então, para todos $p, p' \in M$,*

$$|f(p), f(p')| = \langle\!\langle f \rangle\!\rangle\,|p, p'|\,.$$

□ *Demonstração.* Para algum $c \in [0, \infty[$, f satisfaz que, para todos $p, p' \in M$,

$$|f(p), f(p')| = c\,|p, p'|\,,$$

o que implica $\langle\!\langle f \rangle\!\rangle \leq c$. Se valesse $\langle\!\langle f \rangle\!\rangle < c$, então, para todos $p, p' \in M$ tais que $p \neq p'$, seguiria que

$$c\,|p, p'| = |f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle\,|p, p'| < c\,|p, p'|\,,$$

o que seria uma contradição. Assim $\langle\!\langle f \rangle\!\rangle \geq c$, portanto $\langle\!\langle f \rangle\!\rangle = c$ e concluímos que

$$|f(p), f(p')| = \langle\!\langle f \rangle\!\rangle\,|p, p'|\,. \qquad \blacksquare$$

⊢ **Proposição 16.48.** *Sejam $\boldsymbol{M}$ e $\boldsymbol{M'}$ espaços métricos e $f\colon M \longrightarrow M'$ uma função controlada com $\langle\!\langle f \rangle\!\rangle > 0$. Então f é semelhança local se, e somente se, f é mergulho métrico e $\langle\!\langle f^{-1} \rangle\!\rangle = \langle\!\langle f \rangle\!\rangle^{-1}$.*

□ *Demonstração.*

- ($\Rightarrow$) Suponhamos que f é semelhança local. Mostremos que f é injetiva, sua inversa à esquerda $f^{-1}\colon f(M) \longrightarrow M$ é semelhança e $\langle\!\langle f^{-1} \rangle\!\rangle = \langle\!\langle f \rangle\!\rangle^{-1}$, e seguirá de 16.38 que f é mergulho métrico.
  Sejam $p, p' \in M$ tais que $p \neq p'$. Então

  $$|f(p), f(p')| = \langle\!\langle f \rangle\!\rangle\,|p, p'| > 0$$

  portanto $f(p) \neq f(p')$, o que mostra que f é injetiva.
  Para todos $q, q' \in f(M)$, sejam $p, p' \in M$ tais que $f(p) = q$ e $f(p') = q'$. Então

  $$\left|f^{-1}(q), f^{-1}(q')\right| = |p, p'| = \langle\!\langle f \rangle\!\rangle^{-1}\,|f(p), f(p')| = \langle\!\langle f \rangle\!\rangle^{-1}\,|q, q'|\,,$$

  o que mostra que f^{-1} é semelhança e $\langle\!\langle f^{-1} \rangle\!\rangle = \langle\!\langle f \rangle\!\rangle^{-1}$.

- ($\Leftarrow$) Se f é mergulho métrico, segue de 16.38 que f é função controlada injetiva e sua inversa à esquerda $f^{-1}\colon f(M) \longrightarrow M$ é controlada. Para todos $p, p' \in M$, vale
$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle \, |p, p'| \, .$$
Como $\langle\!\langle f^{-1} \rangle\!\rangle = \langle\!\langle f \rangle\!\rangle^{-1}$, vale
$$\langle\!\langle f \rangle\!\rangle \left| f^{-1}(f(p)), f^{-1}(f(p')) \right| \leq \langle\!\langle f \rangle\!\rangle \langle\!\langle f \rangle\!\rangle^{-1} \, |f(p), f(p')| =^{-1} |f(p), f(p')| \, .$$
Disso concluímos que $|f(p), f(p')| = \langle\!\langle f \rangle\!\rangle \, |p, p'|$. ■

16.4.5 Funções α-controladas

Nesta seção, expandimos o conceito de uma função controlada usando um parâmetro.

:⊢ **Definição 16.29.** Sejam $\boldsymbol{M}$ e $\boldsymbol{M}'$ espaços métricos e $\alpha \in \mathbb{R}_{>0}$. Uma *função α-controlada*[4] de $\boldsymbol{M}$ para $\boldsymbol{M}'$ é uma função $f\colon M \longrightarrow M'$ tal que, para algum $c \in [0, \infty[$, todos $p, p' \in M$ satisfazem
$$|f(p), f(p')| \leq c \, |p, p'|^{\alpha} \, .$$
O conjunto dessas funções é denotado $\mathscr{C}^{0,\alpha}(\boldsymbol{M}, \boldsymbol{M}')$. A *$\alpha$-distorção* (ou *norma de α-distorção*) de f é
$$\langle\!\langle f \rangle\!\rangle_{\mathscr{C}^{0,\alpha}} := \inf \left\{ c \in [0, \infty[\;\middle|\; \forall_{p,p' \in M} \, |f(p), f(p')| \leq c \, |p, p'|^{\alpha} \right\} .$$

A α-distorção de uma função α-controlada f é portanto o menor valor tal que vale a desigualdade entre as distâncias da definição. Vale que
$$|f(p), f(p')| \leq \langle\!\langle f \rangle\!\rangle_{\mathscr{C}^{0,\alpha}} \, |p, p'|^{\alpha} \, .$$
Além disso, a α-distorção de f satisfaz a seguinte definição em termos do supremo:
$$\langle\!\langle f \rangle\!\rangle_{\mathscr{C}^{0,\alpha}} = \sup \left\{ \frac{|f(p), f(p')|}{|p, p'|^{\alpha}} \;\middle|\; p, p' \in M \wedge p \neq p' \right\} .$$
Note que, caso $\#(M) \leq 1$, o conjunto nesse supremos é vazio, e nesse caso devemos ressaltar que o supremo é tomado sobre $[0, \infty[$, portanto o supremo de $\emptyset$ é 0.

O parâmetro α é geralmente tomado em $]0, 1[$ quando o domínio da função é um intervalo real, pois os casos restantes são triviais, como mostra a proposição a seguir.

[4] Essas funções são conhecidas na literatura como funções 'Hölder contínuas', em homenagem ao matemático alemão *Otto Ludwig Hölder* (22/12/1859 – 29/08/1937) https://en.wikipedia.org/wiki/Otto_H%C3%B6lder.

⊢ **Proposição 16.49.** *Sejam $I \subseteq \mathbb{R}$ um intervalo, $\boldsymbol{M}$ um espaço métrico, $\alpha \in \mathbb{R}_{>0}$ e $f \in \mathscr{C}^{0,\alpha}(I, \boldsymbol{M})$. Se $\alpha > 1$, então f é constante.*

□ *Demonstração.* Sejam $x, x' \in I$. Para todo $n \in \mathbb{N}_{\setminus\{0\}}$, definimos $\delta_n := \frac{x'-x}{n}$ e, para todo $k \in [n]$, definimos os pontos $x_{n,k} := x + k\delta_n$, $x_{n,n} := x'$, e os intervalos

$$I_{n,k} :=]x_{n,k}, x_{n,k+1}[\,.$$

Disso segue que

$$|f(x_{n,k}), f(x_{n,k+1})| \leq c\,|x_{n,k+1} - x_{n,k}|^{\alpha} = c\,|x' - x|^{\alpha}\, n^{-\alpha}.$$

e então, da desigualdade triangular, segue que

$$|f(x), f(x')| \leq \mathop{+}_{k\in[n]} |f(x_{n,k}), f(x_{n,k+1})| \leq c\,|x' - x|^{\alpha}\, n^{1-\alpha}.$$

Como $\alpha > 1$, então $\lim_{n\to\infty} n^{1-\alpha} = 0$, portanto

$$|f(x), f(x')| \leq \lim_{n\to\infty} c\,|x' - x|^{\alpha}\, n^{1-\alpha} = 0,$$

o que mostra que $f(x) = f(x')$, logo f é constante. ■

16.5 Medida em espaços métricos

16.5.1 Medidas exteriores métricas

:⊢ **Definição 16.30.** Seja $\boldsymbol{M}$ um espaço métrico. Uma medida exterior *métrica* em $\boldsymbol{M}$ é uma medida exterior $m\colon \wp(M) \to [0, \infty]$ sobre M tal que, para todos $C, C' \subseteq M$ metricamente separados,

$$m(C \cup C') = m(C) + m(C').$$

⊢ **Proposição 16.50.** *Sejam $\boldsymbol{M}$ um espaço métrico, com σ-álgebra topológica $\mathcal{M}_{\mathcal{T}}$ e m uma medida exterior métrica em $\boldsymbol{M}$. Então todo $M \in \mathcal{M}_{\mathcal{T}}$ é m-mensurável.*

□ *Demonstração.* Para mostrar isso, mostraremos que todo conjunto fechado é m-mensurável. Basta mostrar que, para todo $C \subseteq M$ com $m(C) < \infty$ e todo fechado $F \subseteq M$,

$$m(C) \geq m(C \cap F) + m(C \cap F^{\complement}),$$

pois a desigualdade contrária sempre vale por subaditividade e a igualdade vale trivialmente se $m(C) = \infty$. Consideremos as vizinhanças fechadas

$$F_j := \bar{\mathrm{B}}\left(F; \frac{1}{j}\right) = \left\{ p \in M \;\middle|\; |F, p| \leq \frac{1}{j} \right\}.$$

Vale que $\left|C \cap F, C \cap F_j^{\complement}\right| > 0$, portanto

$$m(C) \geq m\left((C \cap F) \cup (C \cap F_j^{\complement})\right) = m(C \cap F) + m(C \cap F_j^{\complement}).$$

Resta mostrar agora que $\lim_{j\to\infty} m(C \cap F_j^{\complement}) = m(C \cap F^{\complement})$. Como F é fechado, podemos escrever, para todo $j \in \mathbb{N}^*$,

$$C \cap F^{\complement} = \{p \in C \mid |F, p| > 0\} = (C \cap F_j^{\complement}) \cup \bigcup_{k=j}^{\infty} R_k,$$

em que $R_k := C \cap \bar{\mathrm{B}}\left(F; \frac{1}{k}\right) \smallsetminus \bar{\mathrm{B}}\left(F; \frac{1}{k+1}\right) = \left\{p \in C \;\middle|\; \frac{1}{k+1} < |F, p| \leq \frac{1}{k}\right\}$. Pela subaditividade de m, segue que

$$m(C \cap F_j^{\complement}) \leq m(C \cap F^{\complement}) \leq m(C \cap F_j^{\complement}) + \mathop{\mathbin{+}}_{k=j}^{\infty} m(R_k).$$

Mas note que $+_{k=1}^{\infty} m(R_k) < \infty$. Isso ocorre pois, para todo $j \geq i+2$, $|F_i, F_j| > 0$, portanto por indução em N segue que

$$\mathop{\mathbin{+}}_{k=1}^{N} m(R_{2k}) = m\left(\bigcup_{k=1}^{N} R_{2k}\right) \leq m(C) < \infty$$

e

$$\mathop{\mathbin{+}}_{k=1}^{N} m(R_{2k-1}) = m\left(\bigcup_{k=1}^{N} R_{2k-1}\right) \leq m(C) < \infty.$$

Portanto $+_{k=1}^{\infty} m(R_k) < \infty$, o que implica que $\lim_{j\to\infty} m(C \cap F_j^{\complement}) = m(C \cap F^{\complement})$, e concluímos que F é m-mensurável, resultando que todo $M \in \mathcal{M}_{\mathcal{T}}$ é m-mensurável. ■

16.5.2 Medidas por coberturas métricas

Nesta seção, definiremos uma família de medidas exteriores em um espaço métrico e usaremos essas medidas para definir a dimensão do espaço métrico e de seus subconjuntos mensuráveis. No entanto, é importante ressaltar que existem diferentes definições de medida e de dimensão em espaços métricos e aqui abordaremos somente uma delas, a medida por coberturas métricas. Uma outra abordagem considera, em vez de coberturas métricas, empacotamentos, e essa abordagem chega a resultados semelhantes, mas às vezes distintos dos que chegaremos aqui. Essa abordagem por empacotamentos é, de certa forma, a noção dual da abordagem que estudaremos usando coberturas. No entanto, um paradigma que é em geral seguido é que as

dimensões definidas coincidam com as dimensões de espaços lineares como a reta, o plano, e de variedades.

Consideremos a função diâmetro em $\boldsymbol{M}$

$$\begin{aligned}\Theta\colon \wp(M) &\longrightarrow [0,\infty]\\ C &\longmapsto \Theta(C).\end{aligned}$$

Essa função Θ não é uma medida exterior em M. Ela satisfaz (1) $\Theta(\emptyset) = 0$ e (2) Para todos $C, D \subseteq M$ tais que $C \subseteq D$, então $C \times C \subseteq D \times D$, portanto $d(C \times C) \subseteq d(D \times D)$, o que implica

$$\Theta(C) \leq \Theta(D);$$

No entanto, não satisfaz (3) Para todos $(C_i)_{i\in\mathbb{N}}$ subconjuntos de M,

$$\Theta\left(\bigcup_{i\in\mathbb{N}} C_i\right) \leq \mathop{+}_{i\in\mathbb{N}} \Theta\left(C_i\right).$$

Para ver isso, considere o intervalo $[0,1]$ e os conjuntos C_i como os intervalos de tamanho $\frac{1}{2^{i+2}}$ e que tocam pela direita nos pontos $\frac{1}{2^i}$ do intervalo $[0,1]$. Facilmente nota-se que

$$\Theta\left(\bigcup_{i\in\mathbb{N}} C_i\right) = 1 > \frac{1}{2} = \mathop{+}_{i\in\mathbb{N}} \Theta\left(C_i\right).$$

Isso ocorre porque a distância entre os conjuntos C_i não é considerada na soma dos diâmetros individuais, mas é considerada na união, e essa distância resulta em $\frac{1}{2}$ nesse caso. Pode-se fazer com que essa diferença seja tão grande quanto se queira.

Definiremos uma família de medidas exteriores em $\boldsymbol{M}$ com um parâmetro d, que representa a 'dimensão' da medida utilizando a função diâmetro Θ. Mas antes brevemente comentamos a definição de cobertura que usaremos.

:⊢ **Definição 16.31.** Sejam $\boldsymbol{M}$ um espaço métrico, $C \subseteq M$ e $\delta \in \,]0,\infty[$. Uma *δ-cobertura* de C é uma cobertura $(C_i)_{i\in I}$ de C tal que, para todo $i \in I$, $\Theta(C_i) \leq \delta$. O conjunto de δ-coberturas de C é $\mathcal{C}_\delta(C)$.

:⊢ **Definição 16.32.** Sejam $\boldsymbol{M}$ um espaço métrico, $C \subseteq M$, $d \in [0,\infty[$ e $\delta \in \,]0,\infty[$. A *medida d-dimensional δ-precisa* de C em $\boldsymbol{M}$ é

$$H^d_\delta(C) := \inf\left\{ \mathop{+}_{i\in\mathbb{N}} \Theta(U_i)^d \;\middle|\; (U_i)_{i\in\mathbb{N}} \in \mathcal{C}_\delta(C) \right\}.$$

A *medida d-dimensional δ-precisa* em $\boldsymbol{M}$ é a função

$$\begin{aligned}H^d_\delta\colon \wp(M) &\longrightarrow [0,\infty]\\ C &\longmapsto H^d_\delta(C).\end{aligned}$$

A sequência $(U_i)_{i\in\mathbb{N}}$ é uma *δ-cobertura* de C. Mostremos que H^d_δ é uma medida exterior em M.

$\vdash$ **Proposição 16.51.** *Sejam $\boldsymbol{M}$ um espaço métrico, $d \in [0,\infty[$ e $\delta \in]0,\infty[$. A função $H^d_\delta \colon \wp(M) \to [0,\infty]$ é uma medida exterior sobre M.*

$\square$ *Demonstração.* Segue direto de 15.20, pois $(\mathcal{C}_\delta, \Theta)$ é um pré-sistema de medida. ■

Definimos agora a medida H^d que independe de δ. Notemos que, se $\delta \leq \delta'$, então $H^d_{\delta'}(C) \leq H^d_\delta(C)$, pois toda cobertura de C com diâmetro δ é uma cobertura com diâmetro δ'. Isso implica que existe em $[0,\infty]$ o limite

$$\lim_{\delta\to 0} H^d_\delta(C) = \sup_{\delta\in]0,\infty[} H^d_\delta(C).$$

$\dashv$ **Definição 16.33.** Sejam $\boldsymbol{M}$ um espaço métrico e $d \in [0,\infty[$. A *medida d-dimensional* em $\boldsymbol{M}$ é a função

$$\begin{aligned} H^d \colon \wp(M) &\longrightarrow [0,\infty] \\ C &\longmapsto H^d(C) := \sup_{\delta\in]0,\infty[} H^d_\delta(C). \end{aligned}$$

$\vdash$ **Proposição 16.52.** *Sejam $\boldsymbol{M}$ um espaço métrico e $d \in [0,\infty[$. A função*

$$H^d \colon \wp(M) \to [0,\infty]$$

é uma medida exterior métrica em M.

$\square$ *Demonstração.* (Conjunto vazio) $H^d(\emptyset) = 0$, pois $H^d_\delta(\emptyset) = 0$ para todo $\delta \in]0,\infty[$; (Monotonicidade) Para todos $C, C' \subseteq M$ tais que $C \subseteq C'$, temos que $H^d_\delta(C) \leq H^d_\delta(C')$ para todo $\delta \in]0,\infty[$, logo $H^d(C) \leq H^d(C')$; (Subaditividade contável) Seja $(C_i)_{i\in\mathbb{N}}$ uma sequência de subconjuntos de M. Como para todo $i \in \mathbb{N}$ e $\delta \in]0,\infty[$ vale por definição que $H^d_\delta(C_i) \leq H^d(C_i)$, segue que

$$H^d_\delta\left(\bigcup_{i\in\mathbb{N}} C_i\right) \leq \mathop{+}_{i\in\mathbb{N}} H^d_\delta(C_i) \leq \mathop{+}_{i\in\mathbb{N}} H^d(C_i).$$

Como isso vale para todo δ, segue que

$$H^d\left(\bigcup_{i\in\mathbb{N}} C_i\right) \leq \mathop{+}_{i\in\mathbb{N}} H^d\left(C_i\right).$$

Por fim, pode-se mostrar que, para todos conjuntos $C, C' \in M$ e para todo $\delta < d(C, C')$, tem-se

$$H^d_\delta(C \cup C') = H^d_\delta(C) + H^d_\delta(C'),$$

portanto, se $d(C, C') > 0$, tem-se

$$H^d(C \cup C') = H^d(C) + H^d(C').$$ ■

Essa medida é comumente chamada de *medida de Hausdorff* d-dimensional. Definem-se os conjuntos mensuráveis como usual para medidas exteriores. Um conjunto $E \subseteq M$ é mensurável se, e somente se, para todo conjunto $C \in M$,

$$H^d(C) = H^d(C \cap E) + H^d(C \cap E^{\complement}).$$

A proposição mostra que H^d é uma medida exterior métrica, todos os conjuntos da σ-álgebra topológica (conjuntos de Borel) são mensuráveis pela medida exterior H^d (16.50) e H^d pode ser restringida para uma medida em M. Em geral, não se pode garantir o mesmo para as medidas exteriores[5] H^d_δ. Pode-se ainda mostrar a seguinte proposição.

$\vdash$ **Proposição 16.53.** *Seja $n \in \mathbb{N}$ e $\mathbb{R}^n$ o espaço métrico real n-dimensional. A medida H^n é um múltiplo da medida de volume* vol^n *em $\mathbb{R}^n$ (Lebesgue):*

$$H^n = \frac{2^n}{\mathrm{vol}^n(\mathbb{B}^n)}\,\mathrm{vol}^n\,.$$

Lembrando que

$$\mathrm{vol}^n(\mathbb{B}^n) = \frac{(\tau/2)^{\frac{n}{2}}}{(n/2)!}$$

em que $\tau = 6,28\ldots$ é a constante do círculo (razão da circunferência pelo raio) e a função fatorial ! é entendida como a extensão dada pela função Γ, definida $x! = \Gamma(x+1)$, ou mais explicitamente por

$$\begin{aligned} !\colon [0,\infty] &\longrightarrow [0,\infty] \\ x &\longmapsto \int_0^\infty t^x \mathrm{e}^{-t}\mathrm{d}t. \end{aligned}$$

Alguns casos particulares são

$$H^0 = \mathrm{vol}^0 = \# \qquad H^1 = \mathrm{vol}^1$$
$$H^2 = \frac{8}{\tau}\,\mathrm{vol}^2 \qquad H^3 = \frac{12}{\tau}\,\mathrm{vol}^3$$

O fator $\mathrm{vol}^n(\mathbb{B}^n)$ poderia ser evitado multiplicando-o na definição de H^n_δ, e o fator 2^n poderia ser evitado avaliando a soma de $\left(\frac{\Theta(U_i)}{2}\right)^n$ em vez de somente $\Theta(U_i)^n$. O número $\frac{\Theta(C)}{2}$ pode ser naturalmente entendido como o *raio* do conjunto C.

[5] https://web.stanford.edu/class/math285/ts-gmt.pdf

16.5.3 Dimensão métrica e fractais

Seja $C \subseteq M$ um conjunto. A função

$$\begin{aligned} H^{(\cdot)}(C)\colon [0,\infty[&\longrightarrow [0,\infty] \\ d &\longmapsto H^d(C) \end{aligned}$$

é uma função com uma propriedade interessante. Ela admite no máximo três valores. Pode-se notar que, se $d \leq d'$, então $H^{d'}(C) \leq H^d(C)$. Além disso, existe $d \in [0,\infty[$ tal que $H^d(C) = 0$. Portanto podemos definir a *dimensão métrica* de C como

$$\dim(C) := \inf \left\{ d \in [0,\infty[\;\middle|\; H^d(C) = 0 \right\}.$$

Nesse caso, pode-se mostrar que, para todo $d > \dim(C)$, $H^d(C) = 0$ e, para todo $d < \dim(C)$, $H^d(C) = \infty$. No entanto, o valor $H^{\dim(C)}(C)$ pode ser qualquer número em $[0,\infty]$. O valor de $H^{\dim(C)}(C)$ pode ser qualquer valor na linha tracejada do gráfico da figura 16.2. Exitem subconjuntos de $\mathbb{R}^d$ que têm dimensões não inteiras. Esses conjuntos são conhecidos como *fractais*.

FIGURA 16.2: Gráfico de $H^d(C)$ em função de d.

16.5.4 Comprimento de caminhos

Seja $\boldsymbol{M}$ um espaço métrico. Lembremos que um caminho em $\boldsymbol{M}$ é uma função contínua $c\colon [a,a'] \longrightarrow M$, em que $[a,a'] \subseteq \mathbb{R}$ é um intervalo compacto não degenerado, e o ponto inicial de c é $c(a)$ e o ponto final de c é $c(a')$. Uma reparametrização de c é um caminho $c'\colon [b,b'] \longrightarrow X$ tal que, para algum homeomorfismo $h\colon [b,b'] \longrightarrow [a,a']$,

$$c' = c \circ h.$$

Além disso, a concatenação de um caminho $c\colon [a,a'] \longrightarrow M$ com um caminho $c'\colon [a',a''] \longrightarrow M$ é o caminho

$$\begin{aligned} c \smile c' \colon [a,a''] &\longrightarrow M \\ t &\longmapsto \begin{cases} c(t), & t \in [a,a'] \\ c'(t), & t \in [a',a''] \end{cases}. \end{aligned}$$

Para a definição do comprimento de um caminho, usamos o conceito de uma partição por intervalos de $[a,a']$.

:⊢ **Definição 16.34.** Sejam $[a,a'] \subseteq \mathbb{R}$ um intervalo compacto não degenerado e $n \in \mathbb{N} \smallsetminus \{0\}$. Uma *$n$-partição por intervalos* (ou *partição por n intervalos*) de $[a,a']$ é uma lista $(t_i)_{i\in[n+1]} \in \mathbb{R}^{n+1}$ tal que

$$a = t_0 < \cdots < t_n = a'.$$

O conjunto dessas partições é denotado $\mathsf{P}_{[,]}([a,a'])$.

Claramente, o conjunto $\{[t_0,t_1[\,,\ldots,[t_{n-2},t_{n-1}[\,,[t_{n-1},t_n]\}$ é uma partição de $[a,a']$ e qualquer partição desse tipo define uma n-partição por intervalos de $[a,a']$.

:⊢ **Definição 16.35.** Sejam $\boldsymbol{M}$ um espaço métrico e $c\colon [a,a'] \longrightarrow M$ um caminho em $\boldsymbol{M}$. O *comprimento* de c é

$$L(c) := \sup \left\{ \mathop{+}_{i\in[n]} |c(t_i), c(t_{i+1})| \;\middle|\; (t_i)_{i\in[n+1]} \in \mathsf{P}_{[,]}([a,a']) \right\}.$$

Um caminho *retificável* é um caminho cujo comprimento é finito.

O comprimento de uma curva é um número real positivo ou ∞.

⊢ **Proposição 16.54.** *Sejam $\boldsymbol{M}$ um espaço métrico e $c\colon [a,a'] \longrightarrow M$, $c'\colon [a',a''] \longrightarrow M$ caminhos em $\boldsymbol{M}$ tais que $c(a') = c'(a')$. Então*

$$L(c \smile c') = L(c) + L(c').$$

□ *Demonstração.* • ($\leq$) Seja $(t''_i)_{i\in[n''+1]} \in \mathsf{P}_{[,]}([a,a''])$ uma partição por intervalos. Defina $n := \min\{i \in [n''] \mid a' \leq t_i\}$, de modo que valha $t_{n-1} < a' \leq t_n$. Defina a partição por intervalos $(t_i)_{i\in[n+1]}$ de $[a,a']$ por

$$t_i := \begin{cases} t''_i, & 0 \leq i \leq n-1 \\ a', & i = n. \end{cases}$$

Se $a' = t_n$, defina $n' := n'' - n$ e a partição por intervalos $(t'_i)_{i\in[n'+1]} := (t''_{i-n})_{i\in[n'+1]}$ de $[a', a'']$; se $a' < t_n$, defina $n' := n'' + 1 - n$ e a partição por intervalos $(t'_i)_{i\in[n'+1]}$ de $[a', a'']$ por

$$t_i := \begin{cases} a', & i = 0 \\ t''_{i+1-n}, & 1 \leq i \leq n'. \end{cases}$$

Em ambos os casos, vale

$$\begin{aligned} \left|c \smile c'(t''_{n-1}), c \smile c'(t''_n)\right| &\leq \left|c \smile c'(t''_{n-1}), c \smile c'(a')\right| + \left|c \smile c'(a'), c \smile c'(t''_n)\right| \\ &= \left|c(t''_{n-1}), c(a')\right| + \left|c'(a'), c'(t''_n)\right|, \end{aligned}$$

portanto

$$\begin{aligned} \mathop{\mathbin{+}}_{i\in[n'']} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right| &= \mathop{\mathbin{+}}_{i=0}^{n-2} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right| \\ &\quad + \left|c \smile c'(t''_{n-1}), c \smile c'(t''_n)\right| \\ &\quad + \mathop{\mathbin{+}}_{i=n}^{n''-1} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right| \\ &\leq \mathop{\mathbin{+}}_{i=0}^{n-2} \left|c(t''_i), c(t''_{i+1})\right| \\ &\quad + \left|c(t_{n-1}), c(a')\right| + \left|c'(a'), c'(t_n)\right| \\ &\quad + \mathop{\mathbin{+}}_{i=n}^{n''-1} \left|c'(t''_i), c'(t''_{i+1})\right| \\ &= \mathop{\mathbin{+}}_{i=0}^{n-1} \left|c(t_i), c(t_{i+1})\right| + \mathop{\mathbin{+}}_{i=0}^{n'-1} \left|c'(t'_i), c'(t'_{i+1})\right|. \end{aligned}$$

Disso concluímos que $L(c \smile c') \leq L(c) + L(c')$.

- ($\geq$) Sejam $(t_i)_{i\in[n+1]} \in \mathsf{P}_{[,]}([a, a'])$ e $(t'_i)_{i\in[n'+1]} \in \mathsf{P}_{[,]}([a', a''])$ partições por intervalos. Defina $n'' := n + n'$ e a partição por intervalos $(t''_i)_{i\in[n''+1]}$ de $[a, a'']$ por

$$t''_i := \begin{cases} t_i, & 0 \leq i \leq n \\ t'_i, & 1 \leq i \leq n'. \end{cases}$$

Segue então que

$$\begin{aligned}
&\mathop{+}_{i=0}^{n-1} |c(t_i), c(t_{i+1})| + \mathop{+}_{i=0}^{n'-1} \left|c'(t'_i), c'(t'_{i+1})\right| \\
&\quad = \mathop{+}_{i=0}^{n-1} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right| + \mathop{+}_{i=n}^{n''-1} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right| \\
&\quad = \mathop{+}_{i\in[n'']} \left|c \smile c'(t''_i), c \smile c'(t''_{i+1})\right|.
\end{aligned}$$

Disso concluímos que $L(c) + L(c') \leq L(c \smile c')$.

■

⊢ **Proposição 16.55.** *Sejam* $\boldsymbol{M}$ *um espaço métrico,* $c\colon [a, a'] \longrightarrow M$ *um caminho e* $h\colon [b, b'] \longrightarrow [a, a']$ *um homeomorfismo. Então*

$$L(c \circ h) = L(c).$$

□ *Demonstração.* Basta mostrar que $L(c \circ h) \leq L(c)$ pois, como h é homeomorfismo, sua inversa $h^{-1}\colon [a, a'] \longrightarrow [b, b']$ é homeomorfismo e $(c \circ h) \circ h^{-1} = c$, portanto seguirá que

$$L(c) = L((c \circ h) \circ h^{-1}) \leq L(c \circ h),$$

logo $L(c \circ h) = L(c)$.

Seja $(t_i)_{i\in[n+1]} \in \mathsf{P}_{[,]}([b, b'])$ uma partição por intervalos. Como h é homeomorfismo, existem duas possibilidades: $h(b) = a$ ou $h(b) = a'$. No primeiro caso, em que $h(b) = a$, vale

$$a = h(t_0) < \cdots < h(t_n) = a',$$

portanto, definindo a partição por intervalos $(t'_i)_{i\in[n+1]} := (h(t_i))_{i\in[n+1]}$ de $[a, a']$, segue que

$$\mathop{+}_{i=0}^{n-1} |c \circ h(t_i), c \circ h(t_{i+1})| = \mathop{+}_{i=0}^{n-1} \left|c(t'_i), c(t'_{i+1})\right|.$$

Disso concluímos que $L(c \circ h) \leq L(c)$. No segundo caso, em que $h(b) = a'$, vale

$$a = h(t_n) < \cdots < h(t_0) = a',$$

portanto, definindo a partição por intervalos $(t'_i)_{i\in[n+1]} := (h(t_{n-i}))_{i\in[n+1]}$ de $[a, a']$, segue que

$$\mathop{+}_{i=0}^{n-1} |c \circ h(t_i), c \circ h(t_{i+1})| = \mathop{+}_{i=0}^{n-1} \left|c(t'_{n-i}), c(t'_{n-i+1})\right| = \mathop{+}_{i=0}^{n-1} \left|c(t'_i), c(t'_{i+1})\right|.$$

Disso concluímos que $L(c \circ h) \leq L(c)$. ■

Capítulo 17

Grupos topológicos

17.1 Grupo topológico

:⊢ **Definição 17.1.** Um *grupo topológico* é uma quádrupla $\boldsymbol{G} = ((G, \mathcal{T}), \times, \diagup, 1)$ em que $(G, \mathcal{T})$ é um espaço topológico, $(G, \times, \diagup, 1)$ é um grupo e as operações

$$\times \colon G^2 \to G \qquad \text{e} \qquad \diagup \colon G \to G$$

são funções contínuas.

Por simplicidade, denotamos $gg' := g \times g'$ e $\diagup g := g^{-1}$. Note que não é necessária nenhuma condição sobre a identidade $1 \in G$, pois se vista como uma operação zerária $1 \colon \{0\} \to G$, ela é contínua em qualquer topologia de G, pois $\{0\}$ só admite a topologia trivial.

Lembremos que para um grupo $\boldsymbol{G}$ e $g \in G$, a conjugação por g é a função

$$\begin{aligned} C_g \colon G &\longrightarrow G \\ h &\longmapsto ghg^{-1}, \end{aligned}$$

a translação à direita por g é a função

$$\begin{aligned} D_g \colon G &\longrightarrow G \\ h &\longmapsto hg \end{aligned}$$

e a translação à esquerda por g é a função

$$\begin{aligned} E_g \colon G &\longrightarrow G \\ h &\longmapsto gh. \end{aligned}$$

⊢ **Proposição 17.1.** *Seja $\boldsymbol{G}$ um grupo topológico. As funções C_g, D_g e E_g são homeomorfismos para todo $h \in G$.*

□ *Demonstração.* Para ver que as funções são bijeções, basta notar que $(C_g)^{-1} = C_{g^{-1}}$, $(D_g)^{-1} = D_{g^{-1}}$ e $(E_g)^{-1} = E_{g^{-1}}$. Para mostrar a continuidade, basta notar que da continuidade de $\times$ e das relações $E_g = \times \circ (\cdot, g)$ e $D_g = \times \circ (g, \cdot)$, segue que E_g e D_g são contínuas, e que da continuidade de $\diagup$ e da relação $C_g = D_g \circ E_{g^{-1}}$, segue que C_g é contínua. ■

► **Exemplo 17.1.** *Sejam* $\boldsymbol{X}$ *um espaço topológico e* $\boldsymbol{G}$ *um grupo topológico. Consideremos o conjunto* $\mathscr{C}(X, G)$ *das funções contínuas de* $\boldsymbol{X}$ *para* $\boldsymbol{G}$ *e as operações induzidas pontualmente de* G *em* $\mathscr{C}(X, G)$ *por* $(f \times g)(x) := f(x) \times g(x)$, $(f^{-1})(x) := (f(x))^{-1}$ *e* $1(x) := 1$. *A quádrupla* $(\mathscr{C}(X, G), \times, \diagup, 1)$ *é um grupo. Consideramos agora a topologia* compacto-aberto $\mathcal{T}$ *gerada pelos conjuntos*

$$\mathscr{A}_{K,A} := \{f \in \mathscr{C}(X, G) \mid f(K) \subseteq A\}$$

em que $K \subseteq X$ *é um compacto e* $A \subseteq G$ *é um aberto. Então* $((\mathscr{C}(X, G), \mathcal{T}), \times, \diagup, 1)$ *é um grupo topológico.*

⊢ **Proposição 17.2.** *Seja* $\boldsymbol{G}$ *um grupo topológico.*

1. *Para todos* $g \in G$ *e* $A \subseteq G$ *aberto,* gA *e* Ag *são abertos;*
2. *Para todos* $g \in G$ *e* $A \subseteq G$ *fechado,* gA *e* Ag *são fechados;*
3. *Para todos* $A \subseteq G$ *aberto e* $C \subseteq G$, AC *e* CA *são abertos;*
4. *Para todos* $F \subseteq G$ *fechado e* $K \subseteq G$ *compacto,* FK *e* KF *são fechados.*

□ *Demonstração.*

1. Segue do fato de que E_g e D_g são homeomorfismos e $gA = E_g(A)$, $Ag = D_g(A)$.
2. Segue do fato de que E_g e D_g são homeomorfismos e $gA = E_g(A)$, $Ag = D_g(A)$.
3. Segue do fato de que

$$AC = \bigcup_{c \in C} Ac \quad \text{e} \quad CA = \bigcup_{c \in C} cA$$

   e de que os conjuntos cA e Ac são abertos para todo $c \in C$.
4. Seja $x \in \overline{FK}$. Então existe uma rede $(x_\lambda)_{\lambda \in \Lambda} = (f_\lambda k_\lambda)$ tal que $x = \lim_{\lambda \in \Lambda} x_\lambda$. Como K é compacto, existe uma sub-rede k_{λ_μ} tal que $k := \lim_{\mu \in M} k_{\lambda_\mu} \in K$, o que implica que $k^{-1} = \lim_{\mu \in M} {k_{\lambda_\mu}}^{-1}$ pela continuidade da inversa. Pela continuidade do produto, segue que a sub-rede $f_{\lambda_\mu} = (f_{\lambda_\mu} k_{\lambda_\mu}){k_{\lambda_\mu}}^{-1}$ converge para $f = xk^{-1}$, e $f \in F$ pois F é fechado. Assim segue que $x = fk$, portanto FK é fechado. A demonstração é análoga para KF.

■

Para o exemplo a seguir, seja $m \in \mathbb{N} \setminus \{0\}$ e lembremos que um m-multi-índice é um elemento $i \in I$ e sua norma é norma 1 dada por

$$|i| = |i_0| + \cdots + |i_{m-1}| \in \mathbb{N}.$$

Além disso, para um grupo $\boldsymbol{G} = (G, +, -, 0)$, o suporte de uma função $s\colon I \longrightarrow G$, que é um elemento de $s \in G^I$, é o conjunto de multi-índices $i \in I$ tais que $s_i \neq 0$, ou seja,

$$\operatorname{supp}(s) = \{i \in I \mid s_i \neq 0\}.$$

A ordem de uma função $s \in I \smallsetminus \{0\}$ é

$$|s|_\circ := \min\{|i| \mid i \in \operatorname{supp}(s)\},$$

a menor norma de um índice de uma entrada não 0. A ordem de $0 \in I$ pode ser definida como ∞, mas não usaremos essa definição explicitamente, embora ela seja compatível com a definição de distância a seguir. As mesmas definições valem analogamente para qualquer $I \subseteq I$.

$\vdash$ **Proposição 17.3.** *Sejam $\boldsymbol{G} = (G, +, -, 0)$ um grupo comutativo discreto, $m \in \mathbb{Z} \smallsetminus \{0\}$, $I \subseteq \mathbb{Z}^m$ e $\alpha \in]0, 1[$. A topologia produto de $\boldsymbol{G}^I$ é dada pela distância*

$$\begin{aligned} |\cdot,\cdot|_\alpha \colon G^I \times G^I &\longrightarrow \mathbb{R} \\ (s, s') &\longmapsto |s, s'|_\alpha := \begin{cases} 0, & s = s' \\ \alpha^{|s'-s|_\circ}, & s \neq s'. \end{cases} \end{aligned}$$

$\square$ *Demonstração.* Mostremos que $|\cdot,\cdot|_\alpha$ é uma distância.

- (Separação) Sejam $s, s' \in G^I$. Se $s = s'$, por definição $|s, s'|_\alpha = 0$. Se $s \neq s'$, então $|s, s'|_\alpha = \alpha^{|s'-s|_\circ}$ e, como $\alpha > 0$, $|s, s'|_\alpha > 0$.
- (Simetria) Sejam $s, s' \in G^I$. Se $s = s'$, então $|s, s'|_\alpha = 0 = |s', s|_\alpha$. Se $s \neq s'$, então $|s' - s|_\circ = |-(s' - s)|_\circ = |s - s'|_\circ$, portanto $|s, s'|_\alpha = |s', s|_\alpha$.
- (Desigualdade triangular) Sejam $s, s', s'' \in G^I$. Se $s = s'$, então $|s, s'|_\alpha = 0$, portanto

  $$|s, s''|_\alpha = |s', s''|_\alpha \leq 0 + |s', s''|_\alpha = |s, s'|_\alpha + |s', s''|_\alpha.$$

  Se $s' = s''$, então $|s', s''|_\alpha = 0$, portanto

  $$|s, s''|_\alpha = |s, s'|_\alpha \leq |s, s'|_\alpha + 0 = |s, s'|_\alpha + |s', s''|_\alpha.$$

  Se $s = s''$, então $|s, s''|_\alpha = 0$, portanto

  $$|s, s''|_\alpha = 0 \leq |s, s'|_\alpha + |s', s''|_\alpha.$$

  Se $s \neq s'$, $s' \neq s''$ e $s \neq s''$, então $|s, s'|_\alpha = \alpha^{|s'-s|_\circ}$, $|s', s''|_\alpha = \alpha^{|s''-s'|_\circ}$ e $|s, s''|_\alpha = \alpha^{|s''-s|_\circ}$. Como $0 \neq s'' - s = (s'' - s') + (s' - s)$, vale que

  $$|s'' - s|_\circ = |(s'' - s') + (s' - s)|_\circ \geq |s'' - s'|_\circ \wedge |s' - s|_\circ,$$

portanto, como $\alpha < 1$,

$$\begin{aligned}
|s, s''|_\alpha &= \alpha^{|s''-s|_\circ} \\
&= \alpha^{|(s''-s')+(s'-s)|_\circ} \\
&\leq \alpha^{|s''-s'|_\circ \wedge |s'-s|_\circ} \\
&\leq \alpha^{|s''-s'|_\circ} + \alpha^{|s'-s|_\circ} \\
&= |s, s'|_\alpha + |s', s''|_\alpha .
\end{aligned}$$

Mostraremos que as topologias são iguais.

- (Topologia métrica é mais fina que topologia produto) Não usaremos que a topologia de G é discreta. Seja $U \subseteq G^I$ um aberto básico. Por definição da topologia produto, U é da forma

$$U = \prod_{i \in I} U_i,$$

e existe um conjunto finito $F \subseteq I$ tal que, para todo $i \notin F$, $U_i = G$. Se $U = \emptyset$, definindo $c := 0 \in G^I$ e $r := 0 \in \mathbb{R}$, segue que $\mathrm{B}(c; r) = \emptyset \subseteq U$. Caso contrário, seja $u \in U$. Se $F = \emptyset$, então $U = G^I$; definindo $r := 1 \in \mathbb{R}$, segue que $u \in \mathrm{B}(u; r) \subseteq G^I = U$. Se $F \neq \emptyset$, defina

$$r := \alpha^{\max|F|} = \alpha^{\max\{|i| \mid i \in F\}}.$$

Mostremos que $u \in \mathrm{B}(u; r) \subseteq U$. Note que $r > 0$, pois $\alpha \in]0, 1[$, logo $u \in \mathrm{B}(u; r)$. Seja $x \in \mathrm{B}(u; r)$. Se $x = u$, então $x \in U$; caso contrário, $x - u \neq 0$, logo $|u, x|_\alpha = \alpha^{|x-u|_\circ}$. Como $x \in \mathrm{B}(u; r)$, segue que

$$\alpha^{|x-u|_\circ} = |u, x|_\alpha < r = \alpha^{\max|F|},$$

o que implica que $\max|F| < |x - u|_\circ$, pois $\alpha \in]0, 1[$. Por sua vez, isso implica que, para todo $i \in I$ tal que $|i| \leq \max|F|$, $x_i - u_i = 0$, ou seja, $x_i = u_i$. Como todo $i \in F$ satisfaz $|i| \leq \max|F|$, segue que $x_i = u_i \in U_i$, o que mostra que $x \in U$. Portanto $\mathrm{B}(u; r) \subseteq U$.

- (Topologia produto é mais fina que topologia métrica) Usaremos que a topologia de G é discreta. Seja $\mathrm{B}(c; r) \subseteq G^I$ uma bola aberta. Se $r = 0$, $\mathrm{B}(c; r) = \emptyset$; defina $U := \emptyset$, e segue que $U \subseteq \mathrm{B}(c; r)$. Se $r > 0$, $\mathrm{B}(c; r) = \emptyset$; defina $k := \lceil \log_\alpha(r) \rceil$ e $F := \{i \in I \mid |i| \leq k\}$, que é finito. Como a topologia de G é discreta, para todo $i \in F$, $\{c_i\} \subseteq G$ é aberto; defina assim o aberto básico

$$U := \prod_{i \in I} U_i,$$

em que $U_i := \{c_i\}$ se $i \in F$ e $U_i = G$ se $i \notin F$.
Mostremos que $c \in U \subseteq \mathrm{B}(c;r)$. Seja $x \in U$. Então, para todo $i \in F$, $x_i = c_i$, ou seja, $x_i - c_i = 0$. Se $x \neq c$, isso implica que $|x-c|_\circ > k$ pela definição de F; como $\alpha < 1$, segue que $|c,x|_\alpha = \alpha^{|x-c|_\circ} < \alpha^k$. Se $x = c$, então $|c,x|_\alpha = 0 < \alpha^k$, pois $\alpha > 0$. Em ambos os casos, $|c,x|_\alpha < \alpha^k$. Como $\alpha < 1$ e $k = \lceil \log_\alpha(r) \rceil \geq \log_\alpha(r)$, segue que $\alpha^k \leq \alpha^{\log_\alpha(r)} = r$, portanto que $|c,x|_\alpha < \alpha^k \leq r$, logo $x \in \mathrm{B}(c;r)$. Isso mostra que $U \subseteq \mathrm{B}(c;r)$. Além disso, $c \in U$, pois, para todo $i \in F$, $c_i \in \{c_i\}$, logo $c \in U$.
Assim, dado $u \in \mathrm{B}(c;r)$, tomamos $r' > 0$ tal que $\mathrm{B}(u;r') \subseteq \mathrm{B}(c;r)$, e pela demonstração anterior existe aberto básico $U' \subseteq G^I$ tal que $u \in U' \subseteq \mathrm{B}(u;r') \subseteq \mathrm{B}(c;r)$.

■

17.2 Topologia em grupos

Uma peculiaridade dos grupos topológico é que, como temos homeomorfismo do grupo para ele mesmo que levam qualquer elemento para a identidade, segue que basta estudar as vizinhanças abertas da identidade para entender a topologia do grupo. Isso fica mais claro com os resultados a seguir.

:⊢ **Definição 17.2.** Seja $\boldsymbol{G}$ um grupo. Um *sistema de vizinhanças da identidade* em $\boldsymbol{G}$ é um conjunto $\mathcal{V}$ de subconjuntos de G tal que

1. Para todo $A \in \mathcal{V}$, $1 \in A$;
2. Para todos $A, A' \in \mathcal{V}$, $A \cap A' \in \mathcal{V}$;
3. Para todo $A \in \mathcal{V}$, existe $B \in \mathcal{V}$ tal que $B^2 \subseteq A$;
4. Para todo $A \in \mathcal{V}$, $A^{-1} \in \mathcal{V}$;
5. Para todos $A \in \mathcal{V}$, $g \in G$, $gAg^{-1} \in \mathcal{V}$.

Denotaremos por $\mathcal{V} := \mathcal{V}_1^\circ$ as vizinhanças abertas da identidade 1.

⊢ **Proposição 17.4.** *Seja $\boldsymbol{G}$ um grupo topológico. O conjunto $\mathcal{V}$ de vizinhanças abertas de 1 é um sistema de vizinhanças da identidade.*

:⊢ **Definição 17.3.** Seja $\boldsymbol{G}$ um grupo. Uma topologia *invariante por translação à esquerda* em $\boldsymbol{G}$ é uma topologia $\mathcal{T}$ em G tal que, para todos $A \in \mathcal{T}$ e $g \in G$, $gA \in \mathcal{T}$. Uma topologia *invariante por translação à direita* em $\boldsymbol{G}$ é uma topologia $\mathcal{T}$ em G tal que, para todos $A \in \mathcal{T}$ e $g \in G$, $Ag \in \mathcal{T}$.

⊢ **Proposição 17.5.** *Sejam $\boldsymbol{G}$ um grupo e $\mathcal{T}$ uma topologia em G.*

1. *Se $\mathcal{T}$ é invariante por translação à esquerda em $\boldsymbol{G}$ então a topologia produto $\mathcal{T}^2$ em $G \times G$ é invariante por translação à esquerda em $\boldsymbol{G} \times \boldsymbol{G}$;*

2. *Se $\mathcal{T}$ é invariante por translação à direita em $\boldsymbol{G}$ então a topologia produto $\mathcal{T}^2$ em $G \times G$ é invariante por translação à direita em $\boldsymbol{G} \times \boldsymbol{G}$.*

$\vdash$ **Proposição 17.6.** *Seja $\boldsymbol{G}$ um grupo e $\mathcal{T}$ uma topologia invariante por translação à esquerda e à direita em $\boldsymbol{G}$. Então $((G, \mathcal{T}), \times, /, 1)$ é um grupo topológico se, e somente se,*

1. *$\times$ é contínua em $(1, 1)$;*
2. *$/$ é contínua em 1.*

$\vdash$ **Proposição 17.7.** *Sejam $\boldsymbol{G}$ um grupo e $\mathcal{V}$ um sistema de vizinhanças da identidade em $\boldsymbol{G}$. Existe uma única topologia $\mathcal{T}$ em G tal que $((G, \mathcal{T}), \times, /, 1)$ é um grupo topológico e $\mathcal{V}$ é um sistema fundamental de vizinhanças de 1 em $(G, \mathcal{T})$.*

$\vdash$ **Proposição 17.8.** *Seja $\boldsymbol{G}$ um grupo topológico. São equivalentes:*

1. *$(G, \mathcal{T})$ é separado;*
2. *$\{1\}$ é fechado;*
3. *$\bigcap_{A \in \mathcal{V}} A = \{1\}$.*

17.3 Distância em grupos

:$\vdash$ **Definição 17.4.** Seja $\boldsymbol{G}$ um grupo. Uma distância *invariante por translação à esquerda* em $\boldsymbol{G}$ é uma distância $|\cdot, \cdot| : G \times G \to [0, \infty[$ em G tal que, para todo $g \in G$, a translação à esquerda E_g é uma isometria local em $(G, |\cdot, \cdot|)$. Uma distância *invariante por translação à direita* em $\boldsymbol{G}$ é uma distância em G tal que, para todo $g \in G$, a translação à direita D_g é uma isometria local. Uma distância *invariante por translação* em $\boldsymbol{G}$ é uma distância invariante à esquerda e à direita em $\boldsymbol{G}$.

No caso da invariância à esquerda, isso quer dizer que, para todos $g, g', g'' \in G$,

$$|gg', gg''| = |g', g''| .$$

O caso à direita é análogo.

$\vdash$ **Proposição 17.9.** *Seja $\boldsymbol{G}$ um grupo topológico e $\mathcal{V}$ um sistema de vizinhanças da identidade enumerável. Existem distâncias $|\cdot, \cdot|_E$ e $|\cdot, \cdot|_D$ em G invariantes por translação à esquerda e à direita em $\boldsymbol{G}$, respectivamente, e compatíveis com a topologia de $\boldsymbol{G}$.*

17.4 Homomorfismos contínuos

⊢ **Proposição 17.10.** *Sejam* $\boldsymbol{G}$ *e* $\boldsymbol{H}$ *grupos topológicos e* $\phi\colon G \to H$ *um homomorfismo de grupos. Então* ϕ *é contínuo se, e somente se,* ϕ *é contínuo na identidade* $1 \in G$.

□ *Demonstração.* A ida é evidente; basta mostrar a volta. Como ϕ é homomorfismo, $\phi \circ E_g = E_{\phi(g)} \circ \phi$ para todo $g \in G$. Mas $E_{\phi(g)} \circ \phi$ é contínua em 1, o que implica que $\phi \circ E_g$ é contínuo em 1 e, como E_g é um homeomorfismo, segue que ϕ é contínuo em $g = E_g(1)$. ■

⊢ **Proposição 17.11.** *Sejam* $\boldsymbol{G}$ *e* $\boldsymbol{G}'$ *grupos topológicos,* $\boldsymbol{G}'$ *separado. Uma função* $\phi\colon G \to G'$ *é homomorfismo contínuo se, e somente se,* $\operatorname{graf}(\phi)$ *é subgrupo de* $\boldsymbol{G} \times \boldsymbol{G}'$ *homeomorfo a* G *pela projeção* $\mathrm{p}_G\colon G \times G' \to G$.

17.5 Subgrupos topológicos

:⊢ **Definição 17.5.** Seja $\boldsymbol{G} = ((G, \mathcal{T}), \times, \diagup, 1)$ um grupo topológico. Um *subgrupo topológico* de $\boldsymbol{G}$ é uma quádrupla $\boldsymbol{S} = ((S, \mathcal{T}_S), \times|_S, \diagup|_S, 1)$ em que $(S, \mathcal{T}|_S)$ é um subespaço topológico de $(G, \mathcal{T})$ e $(S, \times|_S, \diagup|_S, 1)$ é um subgrupo de $(G, \times, \diagup, 1)$.

⊢ **Proposição 17.12.** *Seja* $\boldsymbol{G}$ *um grupo topológico e* $\boldsymbol{S}$ *um subgrupo de* $\boldsymbol{G}$. *Então* $\boldsymbol{S}$ *é um grupo topológico.*

□ *Demonstração.* Temos que mostrar que as operações $\times|_S$ e $\diagup|_S$ são contínuas.

- ($\times|_S$ é contínua) Seja $C \subseteq G$. Como
$$\times|_S^{-1}(C \cap S) = \times^{-1}(C) \cap (S \times S),$$
portanto para todo aberto $A \cap S$ de S, com $A \subseteq G$ aberto de G, segue que $\times|_S^{-1}(A \cap S) = \times^{-1}(A) \cap (S \times S)$ é aberto em $S \times S$, logo $\times|_S$ é contínua.
- ($\diagup|_S$ é contínua) Análogo ao caso do produto. ■

⊢ **Proposição 17.13.** *Seja* $\boldsymbol{G}$ *um grupo topológico.*

1. *Se* $\boldsymbol{S}$ *é subgrupo topológico de* $\boldsymbol{G}$, *então* $\overline{\boldsymbol{S}}$ *é um subgrupo topológico de* $\boldsymbol{G}$.
2. *Se* $\boldsymbol{N}$ *é subgrupo topológico normal de* $\boldsymbol{G}$, *então* $\overline{\boldsymbol{N}}$ *é um subgrupo topológico normal de* $\boldsymbol{G}$.

⊢ **Proposição 17.14.** *Sejam* $\boldsymbol{G}$ *um grupo topológico e* $\boldsymbol{S}$ *é subgrupo topológico de* $\boldsymbol{G}$.

1. *Se* $S^\circ \neq \emptyset$, *então* S *é aberto;*

2. *Se S é aberto, então S é fechado.*

□ *Demonstração.* 1. Seja $g \in S$. Para todo $g' \in S$, o conjunto $g'g^{-1}S$ é aberto e $g' \in g'g^{-1}S \subseteq S$,

2.

■

17.6 Conexidade em grupos topológicos

:⊢ **Definição 17.6.** Seja $\boldsymbol{G}$ um grupo topológico. A componente conexa de 1 é denotada G_1.

⊢ **Proposição 17.15.** *Seja $\boldsymbol{G}$ um grupo topológico. Então $\boldsymbol{G_1}$ é um subgrupo topológico normal e fechado de $\boldsymbol{G}$ e existe bijeção entre as componentes conexas de $\boldsymbol{G}$ e as classes laterais de G_1.*

⊢ **Proposição 17.16.** *Seja $\boldsymbol{G}$ um grupo topológico.*

1. *Se $\boldsymbol{G}$ é localmente conexo, então G_1 é aberta;*
2. *Se $\boldsymbol{G}$ é conexo, para toda vizinhança A de 1,*

$$G = \bigcup_{n \in \mathbb{N}} A^n.$$

17.7 Grupo de homeomorfismos

Seja $\boldsymbol{X}$ um espaço topológico. O espaço $\overleftrightarrow{\mathscr{C}}(X) = \overleftrightarrow{\mathscr{C}}(X, X)$ dos homeomorfismos de $\boldsymbol{X}$ para $\boldsymbol{X}$ é um grupo com produto de composição $\circ$, a inversa de função $^{-1}$ e a função identidade I. Uma questão interessante então é definir uma topologia para esse espaço.

Num primeiro momento, podemos notar que podemos munir esse espaço da topologia produto no seguinte sentido. O conjunto $\overleftrightarrow{\mathcal{F}}(X)$ de bijeções de X para X é subconjunto do conjunto $\mathcal{F}(X)$ de funções de X para X. Esse conjunto é também denotado X^X, o que deixa mais explícito sua estrutura como um espaço produto. Como $\boldsymbol{X}$ é um espaço topológico, podemos explicitar os abertos sub-básicos dessa topologia. Eles consistem em um produto do tipo $\mathcal{A}_{x_0} := \prod_{x \in X} A_x$, em que $x_0 \in X$, $A_{x_0} \subseteq X$ é aberto e, para todo $x \in X \smallsetminus \{x_0\}$, $A_x = X$. O conjunto de pontos desse aberto de X^X corresponde ao conjunto de funções $f \in \mathcal{F}(X)$ tais que $f(x_0) \in A_{x_0}$. Os abertos básicos dessa topologia são interseções finitas desses abertos sub-básicos. Pode-se mostrar que essa é a topologia pontual ou finito-aberto, a topologia gerada pelos abertos sub-básicos

$$\mathcal{A}_{C,A} := \{f \in \mathcal{F}(X) \mid f(C) \subseteq A\},$$

em que $C \subseteq X$ é um conjunto finito e $A \subseteq X$ é um conjunto aberto. Finalmente, tendo essa topologia produto em $\mathcal{F}(X)$, podemos munir $\overset{\leftrightarrow}{\mathscr{C}}(X)$ com a topologia induzida de subespaço.

No entanto, notemos que a topologia de X só foi utilizado quando X tinha o papel de contradomínio das funções $\mathcal{F}(X) = \mathcal{F}(X, X)$; a mesma construção poderia ser feita para o conjunto de funções $\mathcal{F}(I, X)$, em que I é um conjunto qualquer, e ainda teríamos uma topologia nesse espaço de funções. Se queremos levar em consideração a topologia de X como domínio, devemos construir uma outra topologia. A topologia geralmente adotada nesse caso é a topologia compacto-aberto, a topologia gerada pelos abertos sub-básicos

$$\mathscr{A}_{K,A} := \{f \in \mathscr{C}(X) \mid f(K) \subseteq A\},$$

em que $K \subseteq X$ é um conjunto compacto e $A \subseteq X$ é um conjunto aberto. Note que agora nos restringimos a $\mathscr{C}(X)$ em vez de $\mathcal{F}(X)$. Essa topologia é maior (ou mais fina) que a topologia finito-aberto, pois todo conjunto finito é compacto, e leva em consideração a topologia de X tanto como domínio quanto como contradomínio.

A topologia compacto-aberto não garante, porém, que o grupo de homeomorfismos seja um grupo topológico. Tanto a composição como a inversão de funções pode não ser contínua para alguns espaços topológicos. Existe uma outra topologia que garante isso, a topologia compacto-cocompacto.

⊢ **Proposição 17.17.** *Seja $\boldsymbol{X}$ um espaço topológico. O espaço do homeomorfismos*

$$\left(\left(\overset{\leftrightarrow}{\mathscr{C}}(X), \mathcal{T}_{\mathscr{C}}\right), \circ, ^{-1}, \mathrm{I}\right)$$

é um grupo topológico.

□ *Demonstração.* Para qualquer conjunto X, o espaço de bijeções $(\overset{\leftrightarrow}{\mathcal{F}}(X), \circ, ^{-1}, \mathrm{I})$ é um grupo. No caso de $\boldsymbol{X}$ ser um espaço topológico, temos que $\overset{\leftrightarrow}{\mathscr{C}}(\boldsymbol{X}) \subseteq \overset{\leftrightarrow}{\mathcal{F}}(X)$, pois todo homeomorfismo é uma bijeção. Devemos mostrar que $\overset{\leftrightarrow}{\mathscr{C}}(\boldsymbol{X})$ é um subgrupo. Para isso, notemos que, para todas $f, f' \in \overset{\leftrightarrow}{\mathscr{C}}(X)$, $f' \circ f$ é contínua, pois é composta de contínuas, e tem inversa contínua, pois

$$(f' \circ f)^{-1} = f^{-1} \circ (f')^{-1},$$

que é contínua porque é composta das contínuas f^{-1} e $(f')^{-1}$. Ainda, para toda $f \in \overset{\leftrightarrow}{\mathscr{C}}(X)$, por definição temos $f^{-1} \in \overset{\leftrightarrow}{\mathscr{C}}(X)$. Por fim, claramente I é contínua, o que mostra que $\overset{\leftrightarrow}{\mathscr{C}}(X)$ é um subgrupo de $\overset{\leftrightarrow}{\mathcal{F}}(X)$.

Devemos agora mostrar que a composição e a inversão de funções são contínuas com respeito à topologia compacto-aberto. ■

17.8 Ação contínua

:⊢ **Definição 17.7.** Sejam $\boldsymbol{G}$ um grupo topológico e $\boldsymbol{X}$ um espaço topológico. Uma *ação contínua* de $\boldsymbol{G}$ em $\boldsymbol{X}$ é uma ação

$$\begin{aligned} \cdot\colon G \times X &\longrightarrow X \\ (g, x) &\longmapsto g \cdot x \end{aligned}$$

que é uma função contínua[1]. Denota-se $\cdot\colon \boldsymbol{G} \curvearrowright \boldsymbol{X}$. O grupo $\boldsymbol{G}$ *age continuamente* em $\boldsymbol{X}$.

Lembremos que uma ação é uma função que satisfaz

1. (Identidade) Para todo $x \in X$,

$$1 \cdot x = x;$$

2. (Compatibilidade) Para todos $g, g' \in G$ e $x \in X$,

$$(g'g) \cdot x = g' \cdot (g \cdot x).$$

Para cada $g \in G$, ação $\cdot$ induz uma função contínua

$$\begin{aligned} g\cdot\colon X &\longrightarrow X \\ x &\longmapsto g \cdot x. \end{aligned}$$

Nesse caso, temos um homomorfismo de grupos topológicos (homomorfismo contínuo de grupo)

$$\begin{aligned} \cdot\colon G &\longrightarrow \overleftrightarrow{\mathscr{C}}(X) \\ g &\longmapsto g\cdot\colon X \longrightarrow X \\ &\qquad\qquad x \longmapsto g \cdot x \end{aligned}$$

17.9 Medidas em grupos topológicos

:⊢ **Definição 17.8.** Sejam $\boldsymbol{G}$ um grupo topológico. Uma medida *invariante por translação à direita* sobre $(G, \mathcal{M}_{\mathcal{T}})$ é uma medida m sobre $(G, \mathcal{M}_{\mathcal{T}})$ tal que, para todo $g \in G$ e todo mensurável $M \subseteq G$,

$$m(gM) = m(M).$$

Uma medida *invariante por translação à esquerda* sobre $(G, \mathcal{M}_{\mathcal{T}})$ é uma medida m sobre $(G, \mathcal{M}_{\mathcal{T}})$ tal que, para todo $g \in G$ e todo mensurável $M \subseteq G$,

$$m(Mg) = m(M).$$

[1] Com respeito à topologia produto de $G \times X$.

$\vdash$ **Proposição 17.18** ([2]). *Seja* $\boldsymbol{G}$ *um grupo topológico localmente compacto. Existe uma única (a menos de constante multiplicativa) medida não trivial sobre* $(G, \mathcal{M}_{\mathcal{T}})$ *que é*

1. *Invariante por translação à esquerda;*
2. *Finita em conjuntos compactos;*
3. *Exteriormente regular (em conjuntos mensuráveis);*
4. *Interiormente regular em conjuntos abertos.*

Se $\boldsymbol{G}$ *é compacto, existe única medida como acima tal que* $m(G) = 1$.

[2]Este teorema é conhecido como 'Teorema de Haar', em homenagem ao matemático húngaro *Alfréd Haar* (11/10/1885 – 16/03/1933). https://en.wikipedia.org/wiki/Haar_measure.

Capítulo 18

Espaços lineares topológicos

18.1 Anéis e corpos topológicos

:⊢ **Definição 18.1.** Um *anel topológico* é uma lista $\boldsymbol{A} = ((A, \mathcal{T}), +, -, 0, \times, 1)$ em que $(A, \mathcal{T})$ é um espaço topológico, $(A, +, -, 0, \times, 1)$ é um anel e as operações $+$, $-$ e $\times$ são contínuas.

Um *corpo topológico* é uma lista $\boldsymbol{C} = ((C, \mathcal{T}), +, -, 0, \times, ^{-1}, 1)$ em que $(C, \mathcal{T})$ é um espaço topológico, $(C, +, -, 0, \times, ^{-1}, 1)$ é um corpo e as operações $+$, $-$, $\times$ e $^{-1}$ são contínuas.

Note que, como no caso dos grupos topológicos, não é necessária nenhuma condição sobre 0 e 1, pois se vistos como operações zerárias, são contínuas em qualquer topologia.

Nesta seção deixamos somente as definições de anel e corpo topológico para referência, não exploraremos a teoria de anéis e corpos topológicos.

18.2 Espaço linear topológico

:⊢ **Definição 18.2.** Um *espaço linear topológico* é um espaço linear $\boldsymbol{L} = (L, \cdot)$ sobre um corpo topológico $\boldsymbol{C}$ em que L é um grupo topológico e $\cdot : \boldsymbol{C} \curvearrowright \boldsymbol{L}$ é uma ação contínua.

18.3 Espaço de funções a valores vetoriais

Primeiro relembremos que, se X é um conjunto e $(L, +, 0, -, \cdot)$ é um espaço linear sobre um corpo C, o espaço L^X de funções de X para L é um espaço linear com a

adição pontual

$$\begin{aligned} +\colon L^X \times L^X &\longrightarrow L^X \\ (f, f') &\longmapsto f + f'\colon X \longrightarrow L \\ &\qquad\qquad x \longmapsto f(x) + f'(x), \end{aligned}$$

a função nula

$$\begin{aligned} 0\colon X &\longrightarrow L \\ x &\longmapsto 0, \end{aligned}$$

a inversa pontual da adição

$$\begin{aligned} -\colon L^X &\longrightarrow L^X \\ f &\longmapsto -f\colon X \longrightarrow L \\ &\qquad\qquad x \longmapsto -(f(x)) \end{aligned}$$

e a multiplicação pontual por escalar

$$\begin{aligned} \cdot\colon C \times L^X &\longrightarrow L^X \\ (c, f') &\longmapsto cf\colon X \longrightarrow L \\ &\qquad\qquad x \longmapsto c(f(x)). \end{aligned}$$

⊢ **Proposição 18.1.** *Sejam X um conjunto e $\boldsymbol{L}$ um espaço linear topológico sobre um corpo topológico $\boldsymbol{C}$. O espaço de funções L^X é um espaço linear topológico com a topologia produto.*

□ *Demonstração.* Temos que mostrar que as operações algébricas em L^X são contínuas. As projeções da topologia produto de $L^X = \prod_{x \in X} L$ são as funções avaliação para cada $x \in X$

$$\begin{aligned} \mathrm{p}_x\colon L^X &\longrightarrow L \\ f &\longmapsto f(x). \end{aligned}$$

Notemos que

$$\begin{aligned} \mathrm{p}_x \circ + &= + \circ (\mathrm{p}_x, \mathrm{p}_x), \\ \mathrm{p}_x \circ - &= - \circ \mathrm{p}_x. \\ \mathrm{p}_x \circ \cdot &= \cdot \circ (\mathrm{I}, \mathrm{p}_x), \end{aligned}$$

em que as operações à esquerda são as operações em L^X e as à direita, as em L. Nesse caso, segue que as operações algébricas em L^X são contínuas pela propriedade

universal do produto de espaços topológicos. No caso da adição $+\colon L^X \times L^X \to L^X$, o seguinte diagrama comuta:

$$\begin{tikzcd}
 & L^X \arrow[d, "\mathrm{p}_x"] \\
L^X \times L^X \arrow[ur, dashed, "+"] \arrow[r, "+ \circ (\mathrm{p}_x, \mathrm{p}_x)"'] & L.
\end{tikzcd}$$

Como p_x é contínua para cada $x \in X$ pela definição da topologia produto, segue que $(\mathrm{p}_x, \mathrm{p}_x)\colon L^X \times L^X \to L \times L$ também é contínua, logo da continuidade de $+\colon L \times L \to L$ segue que $+ \circ (\mathrm{p}_x, \mathrm{p}_x)\colon L^X \times L^X \to L$ é contínua. Pela propriedade universal, $+\colon L^X \times L^X \to L^X$ é a única função contínua tal que o diagrama comuta. A demonstração da continuidade das outras operações algébricas é análoga. Os seguintes diagramas comutam:

$$\begin{tikzcd}
 & L^X \arrow[d, "\mathrm{p}_x"] \\
L^X \arrow[ur, dashed, "-"] \arrow[r, "- \circ \mathrm{p}_x"'] & L
\end{tikzcd}
\qquad
\begin{tikzcd}
 & L^X \arrow[d, "\mathrm{p}_x"] \\
C \times L^X \arrow[ur, dashed, "\cdot"] \arrow[r, "\cdot \circ (\mathrm{I}, \mathrm{p}_x)"'] & L.
\end{tikzcd}$$

■

⊢ **Proposição 18.2.** *Sejam $\boldsymbol{X}$ um espaço topológico e $\boldsymbol{L}$ um espaço linear topológico sobre um corpo topológico $\boldsymbol{C}$. O espaço de funções contínuas $\mathscr{C}(\boldsymbol{X}, \boldsymbol{L})$ é espaço linear topológico.*

□ *Demonstração.* Para mostrar isso, basta mostrar que $\mathscr{C}(X, L)$ é um subespaço linear de L^X. (Fechado pela adição) Para todas $f, f' \in \mathscr{C}(X, L)$, $(f, f')\colon X \times X \to L \times L$ é contínua. Como $+\colon L^X \times L^X \to L^X$ é contínua, segue que

$$f + f' = + \circ (f, f')$$

é contínua, pois é composição de contínuas.

(Fechado pela multiplicação por escalar) Para todos $c \in C$, $f \in \mathscr{C}(X, L)$, $(c, f)\colon \{0\} \times X \to C \times L$ é contínua, em que c é interpretado como uma função $c\colon \{0\} \to C$. Como $\cdot$ é contínua, segue que

$$cf = \cdot \circ (c, f)$$

é contínua, pois é composição de contínuas. Isso mostra que $\mathscr{C}(X, L)$ é subespaço linear de L^X, portanto é um espaço linear. ■

18.4 Funções lineares contínuas

:⊢ **Definição 18.3.** Sejam $\boldsymbol{L}, \boldsymbol{L}'$ espaços lineares topológicos. Uma *função linear contínua* de $\boldsymbol{L}$ para $\boldsymbol{L}'$ é uma função $f\colon L \to L'$ que é linear com respeito à estrutura de espaço linear de $\boldsymbol{L}$ e $\boldsymbol{L}'$ e contínua com respeito à estrutura de espaço topológico de $\boldsymbol{L}$ e $\boldsymbol{L}'$. O conjunto dessas funções é denotado $\mathscr{L}(\boldsymbol{L}, \boldsymbol{L}')$.

Isso é o mesmo que dizer que

$$\mathscr{L}(\boldsymbol{L}, \boldsymbol{L}') = \mathcal{L}(\boldsymbol{L}, \boldsymbol{L}') \cap \mathscr{C}(\boldsymbol{L}, \boldsymbol{L}').$$

18.5 Espaço dual contínuo

No estudo de espaços lineares, um objeto importante é o espaço linear dual, o espaço dos funcionais lineares em um espaço linear. No caso de espaços lineares topológicos, um subespaço do espaço linear dual mostra-se mais interessante, o espaço dos funcionais lineares contínuos. No caso de espaço lineares topológico de dimensão finita, todo funcional linear é contínuo, mas no caso geral devemos nos restringir aos contínuos para que a teoria tenha propriedades mais desejáveis e interessantes.

:⊢ **Definição 18.4.** Seja L um espaço linear topológico sobre um corpo topológico C. O *espaço linear dual contínuo* de L é o espaço linear

$$L^{\circledast} := \mathscr{L}(L, C).$$

Como o espaço $L^* = \mathcal{L}(L, C)$ dos funcionais lineares em L e o espaço $\mathscr{C}(L, C)$ dos funcionais contínuos em L são espaços lineares, o espaço

$$L^{\circledast} = L^* \cap \mathscr{C}(L, C),$$

dos funcionais lineares contínuos em L é um espaço linear.

:⊢ **Definição 18.5.** Sejam L um espaço linear topológico sobre um corpo topológico C e $v \in L$. A *avaliação* em v é o funcional linear

$$\begin{aligned} e_v \colon L^{\circledast} &\longrightarrow C \\ f &\longmapsto f(v). \end{aligned}$$

A *topologia*[1] *de* $L^{\circledast}$ é a topologia inicial com respeito à família $\{e_v\}_{v \in L}$

A topologia definida é a menor topologia em $L^{\circledast}$ tal que, para todo $v \in L$, a avaliação $e_v \colon L^{\circledast} \to C$ é contínua; ou seja, é a topologia

$$\mathcal{T} := \left\langle \bigcup_{v \in L} {e_v}^{\star}(\mathcal{T}_C) \right\rangle,$$

em que $\mathcal{T}_C$ é a topologia de C.

Em $\mathbb{R}$, essa é a mesma topologia que a gerada pelos abertos

$$B_{\varepsilon}^{v}(f) := \{f' \in L^* \mid |f'(v) - f(v)| < \varepsilon\},$$

para cada $v \in L$, $\varepsilon \in \,]0, \infty[$ e $f \in L^*$. Basta notar que, para cada $v \in V$,

$$|e_v(f') - e_v(f)| = |f'(v) - f(v)| < \varepsilon,$$

ou seja, essa é a condição para e_v ser contínua.

18.6 Teoremas de representação

18.6.1 Representação linear de produto interno

⊢ **Proposição 18.3.** *Seja* $\boldsymbol{L}$ *um espaço linear com produto interno completo. A função*

$$\begin{aligned} I \colon L &\longrightarrow L^{\circledast} \\ v &\longmapsto I_v \colon V \longrightarrow \mathbb{R} \\ &\qquad\quad\; v' \longmapsto \langle v, v' \rangle \end{aligned}$$

é um isomorfismo de espaços lineares com produto interno.

[1] Chamada topologia fraca ou topologia fraca $*$ (lido 'estrela') de $L^{\circledast}$.

18.6.2 Representação contínua de medida

Lembremos que, para todo espaço topológico $\boldsymbol{X} = (X, \mathcal{T})$, o conjunto $\mathcal{M}_{\mathcal{T}}$ é a álgebra de mensuráveis em X gerada pelos abertos de $\mathcal{T}$. O conjunto $\mathfrak{M}_r(X, \mathcal{M}_{\mathcal{T}})$ é o conjunto das medidas (positivas ou não) completas regulares, o que nesse contexto quer dizer que são finitas em compactos, interiormente regulares em abertos e mensuráveis com medida finita, e exteriormente regulares em mensuráveis.

$\vdash$ **Proposição 18.4.** *Seja $\boldsymbol{X}$ um espaço topológico separado e localmente compacto. A função*

$$\begin{aligned} I \colon \mathfrak{M}_{\mathcal{T}}(X) &\longrightarrow \mathscr{C}_c(X, \mathbb{R})^{\circledast} \\ m &\longmapsto I_m \colon \mathscr{C}_c(X, \mathbb{R}) \longrightarrow \mathbb{R} \\ & \qquad\qquad\qquad\quad f \longmapsto \int_X f \mathrm{d}m \end{aligned}$$

é um isomorfismo de espaços lineares normados. Se restrito às medidas positivas, I é um isomorfismo entre as medidas positivas e os funcionais positivos.

Capítulo 19

Espaços normados

19.1 Normas

19.1.1 Norma em corpos (valor absoluto)

Consideremos um corpo $\boldsymbol{C}$. Queremos definir uma função $|\cdot| : C \to [0, \infty[$ que satisfaça a propriedade de multiplicatividade para todos $c, c' \in C$,

$$|cc'| = |c|\,|c'|\,.$$

Nesse caso, temos

$$|0| = |0c| = |0|\,|c|\,.$$

Se $|0| \neq 0$, então $|c| = 1$ para todo $c \in C$ (inclusive $|0| = 1$), o que mostra que $|\cdot|$ é uma função trivial e, portanto, não tem comportamento tão interessante. É natural então assumir que $|0| = 0$. Nesse caso, não concluímos da multiplicatividade que $|c| = 1$ para todo $c \in C$. No entanto, temos

$$|1| = |1 \times 1| = |1|\,|1|\,.$$

Então $|1| = 0$ ou $|1| = 1$. No primeiro caso, segue que, para todo $c \in C$, $|c| = |1c| = |1|\,|c| = 0\,|c| = 0$, logo $|c| = 0$ para todo $c \in C$, o que mostra que $|\cdot|$ novamente é uma função trivial. É natural, assim, assumir que $|1| \neq 0$, e portanto $|1| = 1$. Assumiremos, no entanto, que para todo $c \in C$, se $|c| = 0$ então $c = 0$, o que implica que $|\cdot|$ é um homomorfismo de grupos entre o grupo multiplicativo de $\boldsymbol{C}$ e o grupo multiplicativo $]0, \infty[$. Notemos ainda que

$$1 = |1| = |(-1)(-1)| = |-1|\,|-1|\,,$$

portanto $|-1| = 1$, já que $|-1| \in [0, \infty[$. Disso segue que

$$|-c| = |-1|\,|c| = |c|\,.$$

Ainda, para todo $c \in C$ segue que

$$\left|c^{-1}\right| = \left|c^{-1}\right| |c| \, |c|^{-1} = \left|c^{-1}c\right| |c|^{-1} = |1| \, |c|^{-1} = |c|^{-1} .$$

Além da multiplicatividade, uma segunda propriedade desejável é a subaditividade: para todos $c, c' \in C$,

$$|c + c'| \leq |c| + |c'| .$$

Note que não adotamos a propriedade mais forte de aditividade. Isso ocorre porque queremos que $|\cdot|$ se comporte de fato como o valor absoluto em $\mathbb{R}$, e também porque queremos que $|\cdot|$ tenha valores positivos, o que a aditividade não permitiria pois de $|0| = 0$ e $|1| = 1$ seguiria que $|-1| = -1$.

:⊢ **Definição 19.1.** Seja $\boldsymbol{C}$ um corpo. Uma *norma* (ou *valor absoluto*) em $\boldsymbol{C}$ é uma função $|\cdot| : C \to [0, \infty[$ que satisfaz

1. (Separação) Para todo $c \in C$, $|c| = 0$ se, e somente se, $c = 0$;
2. (Multiplicatividade) Para todos $c, c' \in C$,

$$|cc'| = |c| \, |c'| ;$$

3. (Subaditividade) Para todos $c, c' \in C$,

$$|c + c'| \leq |c| + |c'| .$$

⊢ **Proposição 19.1** (Propriedades da norma)**.** *Sejam $\boldsymbol{C}$ um corpo e $|\cdot|$ uma norma em $\boldsymbol{C}$.*

1. $|1| = |-1| = 1$;
2. *Para todo $c \in C$, $|-c| = |c|$;*
3. *Para todo $c \in C$, $|c| \geq 0$;*
4. *(Subaditividade generalizada) Para todos $c_0, \ldots, c_{n-1} \in C$,*

$$\left| \mathop{+}_{i \in [n]} c_i \right| \leq \mathop{+}_{i \in [n]} |c_i| ;$$

5. *Para todos $c, c' \in C$, $||c'| - |c|| \leq |c' - c|$.*

□ *Demonstração.* 1. Notemos que

$$|1| = \left|1^2\right| = |1|^2 ,$$

logo $|1| = 0$ ou $|1| = 1$. Da separação de $|\cdot|$, segue que $|1| = 1$.

2. Para todo $c \in C$,

$$|-c| = |-1| \, |c| = |c| .$$

3. Seja $c \in C$. Temos que

$$0 = |0| = |c - c| \leq |c| + |-c| = 2\,|c|\,,$$

logo $|c| \geq 0$
4. Segue por indução da subaditividade de $|\cdot|$.
5. Da subaditividade, segue que

$$|c'| = |(c' - c) + c| \leq |c' - c| + |c|\,,$$

portanto

$$|c'| - |c| \leq |c' - c|\,.$$

Simetricamente obtém-se $|c| - |c'| \leq |c' - c|$, e segue que

$$||c'| - |c|| \leq |c' - c|\,.$$

■

:⊢ **Definição 19.2.** Um *corpo normado* é um par $(\boldsymbol{C}, |\cdot|)$ em que $\boldsymbol{C}$ é um corpo e $|\cdot|$ é uma norma em $\boldsymbol{C}$.

Uma norma em $\boldsymbol{C}$ induz uma métrica e essa métrica, por sua vez, induz uma topologia em C.

:⊢ **Definição 19.3.** Seja $(\boldsymbol{C}, |\cdot|)$ um corpo normado. A *métrica* (induzida pela norma) de $\boldsymbol{C}$ é a função

$$\begin{aligned} |\cdot,\cdot| \colon C \times C &\longrightarrow [0, \infty[\\ (c, c') &\longmapsto |c' - c|\,. \end{aligned}$$

⊢ **Proposição 19.2.** *Seja $(\boldsymbol{C}, |\cdot|)$ um corpo normado. A métrica $|\cdot,\cdot|$ induzida pela norma de $\boldsymbol{C}$ é uma métrica em C.*

⊢ **Proposição 19.3.** *Seja $(\boldsymbol{C}, |\cdot|)$ um corpo normado.*

1. *A norma $|\cdot| : C \to [0, \infty[$ é uma função contínua;*
2. *$\boldsymbol{C}$ é um corpo topológico.*

□ *Demonstração.* 1. Segue direta da propriedade de que, para todos $c, c' \in C$, $||c'| - |c|| \leq |c' - c|$, pois dado $\varepsilon > 0$, tomando $\delta = \varepsilon$ temos que, se $|c, c'| = |c' - c| \leq \delta$, então

$$||c|, |c'|| = ||c'| - |c|| \leq |c' - c| \leq \delta = \varepsilon.$$

2. Para mostrar a continuidade de $+$, podemos usar qualquer norma em $C \times C$; escolhemos a norma

$$\begin{aligned} |\cdot|_{C\times C} \colon C \times C &\longrightarrow [0, \infty[\\ (c, c') &\longmapsto |c| + |c'|\,. \end{aligned}$$

Agora, basta notarmos que, dados $(c_0, c_1), (c_0', c_1') \in C \times C$,

$$\begin{aligned} |(c_0' + c_1') - (c_0 + c_1)| &= |(c_0' - c_0) + (c_1' - c_1)| \\ &\leq |c_0' - c_0| + |c_1' - c_1| \\ &= |(c_0' - c_0, c_1' - c_1)|_{C\times C} \\ &= |(c_0', c_1') - (c_0, c_1)|_{C\times C}\,, \end{aligned}$$

portanto, dado $\varepsilon > 0$, basta tomarmos $\delta = \varepsilon$ e segue que, se

$$|(c_0', c_1') - (c_0, c_1)|_{C\times C} < \delta,$$

então

$$|(c_0' + c_1') - (c_0 + c_1)| \leq |(c_0', c_1') - (c_0, c_1)|_{C\times C} < \delta = \varepsilon.$$

A continuidade das outras operações é análoga.

■

19.1.2 Seminormas

:⊢ **Definição 19.4.** Seja $\boldsymbol{L}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Uma *seminorma* em $\boldsymbol{L}$ é uma função $p \colon L \to \mathbb{R}$ que satisfaz

1. (Homogeneidade absoluta) Para todos $c \in C$ e $v \in L$,

$$p(cv) = |c|\, p(v);$$

2. (Subaditividade) Para todos $v, v' \in L$,

$$p(v + v') \leq p(v) + p(v').$$

⊢ **Proposição 19.4.** *Sejam $\boldsymbol{L}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$ e $p \colon L \to C$ uma seminorma em $\boldsymbol{L}$.*

1. *$p(0) = 0$;*
2. *Para todo $v \in L$, $p(v) \geq 0$;*
3. *Para todo $v, v' \in L$, $|p(v) - p(v')| \leq p(v - v')$.*

:⊢ **Definição 19.5.** Seja $\boldsymbol{L}$ um espaço linear sobre um corpo $(\boldsymbol{C}, \leq)$ ordenado. Um conjunto *absorvedor* em L é um conjunto $A \subseteq L$ tal que, para todo $v \in L$, existe $c \in C_{>0}$ tal que $v \in cA$.

⊢ **Proposição 19.5.** *Sejam L um espaço linear sobre um corpo $(\boldsymbol{C}, \leq)$ ordenado e $A \subseteq L$ um conjunto absorvedor.*

1. $0 \in A$;
2. *Para todo $v \in L$, existe* $\inf \{c > 0 \mid v \in cA\}$.

:⊢ **Definição 19.6.** Sejam $\boldsymbol{L}$ um espaço linear real[1] e $A \subseteq L$ um conjunto absorvedor. O *calibre*[2] de A é a função

$$\begin{aligned} p_A \colon L &\longrightarrow \mathbb{R} \\ v &\longmapsto \inf \{c > 0 \mid v \in cA\}. \end{aligned}$$

A função está bem definida pois A é absorvedor (proposição anterior).

▷ **Exercício 19.1.** *Sejam $\boldsymbol{L}$ um espaço linear real e $A \subseteq L$ um conjunto absorvedor.*

1. *Para todos $v \in L$ e $c \in]0, \infty[$,*
$$p_A(cv) = cp_A(v);$$
2. *Se A é convexo, então p_A é subaditivo: para todos $v, v' \in L$,*
$$p_A(v + v') \leq p_A(v) + p_A(v');$$
3. *Se A é convexo e balanceado, p_A é uma seminorma;*
4. *Definindo $A^\circ := \{v \in L \mid p_A(v) < 1\}$ e $\overline{A} := \{v \in L \mid p_A(v) \leq 1\}$, vale $A^\circ \subseteq A \subseteq \overline{A}$ e $p_{A^\circ} = p_A = p_{\overline{A}}$.*

19.1.3 Normas, espaços normados e métricas lineares

:⊢ **Definição 19.7.** Seja $\boldsymbol{E}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Uma *norma* em $\boldsymbol{E}$ é uma função $\|\cdot\| : E \to \mathbb{R}$ que satisfaz

1. (Separação) Para todo $v \in E$, se $\|v\| = 0$, então $v = 0$.
2. (Homogeneidade absoluta[3]) Para todos $c \in C$ e $v \in E$,
$$\|cv\| = |c| \, \|v\| \, ;$$

[1] Mais geralmente, poderiam ser considerados corpos ordenados normados $(\boldsymbol{C}, \leq, |\cdot|)$, mas os detalhes não serão feitos aqui.

[2] Essas funções são conhecidas como funcionais de Minkowski

[3] Esta propriedade recebe diferentes nomes, incluindo 'Dilatação'.

3. (Subaditividade[4]) Para todos $v, v' \in E$,

$$\|v + v'\| \leq \|v\| + \|v'\| .$$

Uma norma é uma seminorma separada. Claramente $|\cdot| : \mathbb{C} \to \mathbb{R}$ é uma norma em $\mathbb{C}$.

$\vdash$ **Proposição 19.6** (Propriedades da Norma)**.** *Sejam $\boldsymbol{E}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$ e $\|\cdot\|$ uma norma em $\boldsymbol{E}$.*

1. $\|0\| = 0$;
2. *Para todo $v \in E$, $\|-v\| = \|v\|$;*
3. *Para todo $v \in E$, $\|v\| \geq 0$.*
4. *(Subaditividade generalizada) Para todos $v_0, \ldots, v_{n-1} \in E$,*

$$\left\| \mathop{+}_{i \in [n]} v_i \right\| \leq \mathop{+}_{i \in [n]} \|v_i\| .$$

5. *Para todos $v, v' \in E$, $|\|v'\| - \|v\|| \leq \|v' - v\|$.*

$:\vdash$ **Definição 19.8.** Um *espaço normado* é um par $(\boldsymbol{E}, \|\cdot\|)$ em que $\boldsymbol{E}$ é um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$ e $\|\cdot\|$ é uma norma em $\boldsymbol{E}$. A *dimensão* de $(\boldsymbol{E}, \|\cdot\|)$ é a dimensão do espaço linear $\boldsymbol{E}$.

$:\vdash$ **Definição 19.9.** Seja $\boldsymbol{E}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Uma *métrica linear* em $\boldsymbol{E}$ é uma métrica[5] $|\cdot, \cdot| : E \times E \to \mathbb{R}$ em E que satisfaz

1. (Invariância por translação) Para todos $v, v', w \in E$,

$$|v + w, v' + w| = |v, v'| .$$

2. (Homogeneidade absoluta) Para todos $v, v' \in E$ e $c \in C$,

$$|cv, cv'| = |c| \, |v, v'| ;$$

$:\vdash$ **Definição 19.10.** Seja $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado. A *métrica* (induzida pela norma) de $(\boldsymbol{E}, \|\cdot\|)$ é a função

$$\begin{aligned} |\cdot, \cdot| \colon E \times E &\longrightarrow \mathbb{R} \\ (v, v') &\longmapsto \|v - v'\| . \end{aligned}$$

A *topologia de* $(\boldsymbol{E}, \|\cdot\|)$ é a topologia de $(E, |\cdot, \cdot|)$.

[4] Esta propriedade recebe diferentes nomes, incluindo 'Desigualdade triangular'.
[5] Adotamos aqui a notação $|x, x'|$ em vez da notação mais usual $d(x, x')$.

Para que essa definição seja boa, mostramos a seguir a proposição.

⊢ **Proposição 19.7.** *Seja* $(\boldsymbol{E}, \|\cdot\|)$ *um espaço normado. A função*

$$\begin{aligned} |\cdot,\cdot| \colon E \times E &\longrightarrow \mathbb{R} \\ (v, v') &\longmapsto \|v - v'\|\,. \end{aligned}$$

é uma métrica linear em E*.*

□ *Demonstração.* Primeiro mostramos que $|\cdot,\cdot|$ é uma métrica.

1. (Separação) Sejam $v, v' \in E$. Se $v = v'$, então segue da positividade que

$$|v, v'| = |v, v| = \|v - v\| = \|v - v\| = \|0\| = 0.$$

   Reciprocamente, se $|v, v'| = 0$, então $\|v - v'\| = 0$. Segue da separação que $v - v' = 0$, logo $v = v'$.
2. (Simetria) Sejam $v, v' \in E$. Então segue da homogeneidade absoluta que

$$|v, v'| = \|v - v'\| = \|-1(v' - v)\| = |-1|\,\|v' - v\| = |v', v|\,.$$

3. (Desigualdade triangular) Sejam $v, v', v'' \in E$. Então segue da subaditividade que

$$\begin{aligned} |v, v''| &= \|v - v''\| \\ &= \|v - v' + v' - v''\| \\ &\leq \|v - v'\| + \|v' - v''\| \\ &= |v, v'| + |v', v''|\,. \end{aligned}$$

Agora, mostremos que $|\cdot,\cdot|$ é métrica linear.

1. (Invariância por translação) Sejam $v, v', w \in E$. Então

$$|v + w, v' + w| = \|(v + w) - (v' + w)\| = \|v - v'\| = |v, v'|\,.$$

2. (Homogeneidade absoluta) Sejam $v, v' \in E$ e $c \in C$. Então

$$|cv, cv'| = \|cv - cv'\| = \|c(v - v')\| = |c|\,\|v - v'\| = |c|\,|v, v'|\,. \qquad \blacksquare$$

Reciprocamente, se temos uma métrica linear em um espaço linear, essa métrica define uma norma no espaço e a métrica que essa norma define, por sua vez, é a métrica original. Isso mostra, de fato, que existe uma relação bijetiva entre normas e métricas lineares em um espaço linear.

⊢ **Proposição 19.8.** *Sejam* $\boldsymbol{E}$ *um espaço linear sobre um corpo normado* $(\boldsymbol{C}, |\cdot|)$ *e* $|\cdot,\cdot|$ *uma métrica linear em* $\boldsymbol{E}$*. A função*

$$\begin{aligned} \|\cdot\| \colon E &\longrightarrow \mathbb{R} \\ v &\longmapsto |v, 0| \end{aligned}$$

é uma norma em $\boldsymbol{E}$ *e a métrica induzida por essa norma é* $|\cdot,\cdot|$*.*

□ *Demonstração.* Mostremos primeiro que a função é uma norma.

1. (Separação) Seja $v \in E$. Então $\|v\| = |v, 0| = 0$, logo da separação de $|\cdot,\cdot|$ segue que $v = 0$.
2. (Homogeneidade absoluta) Sejam $c \in C$ e $v \in E$. Então segue da homogeneidade absoluta de $|\cdot,\cdot|$ que

$$\|cv\| = |cv, 0| = |c|\,|v, 0| = |c|\,\|v\|\,.$$

3. (Subaditividade) Sejam $v, v' \in E$. Então da invariância por translação, da simetria e da desigualdade triangular de $|\cdot,\cdot|$ que

$$\begin{aligned} \|v + v'\| &= |v + v', 0| \\ &= |v, -v'| \\ &\leq |v, 0| + |0, -v'| \\ &\leq |v, 0| + |v', 0| \\ &= \|v\| + \|v'\|\,. \end{aligned}$$

Agora, mostremos que a métrica $|\cdot,\cdot|'$ induzida por essa norma é a métrica original $|\cdot,\cdot|$. Sejam $v, v' \in E$. Então da invariância por translação de $|\cdot,\cdot|$ segue que

$$|v, v'|' = \|v - v'\| = |v - v', 0| = |v, v'|\,. \qquad \blacksquare$$

19.1.3.1 Funções limitadas e norma do supremo

⊢ **Proposição 19.9.** *Sejam* $\boldsymbol{E}$ *um espaço normado e* $C \subseteq E$ *um conjunto. O conjunto* C *é limitado se, e somente se, o conjunto* $\|C\| \subseteq \mathbb{R}$ *é limitado.*

□ *Demonstração.* 1. ($\Rightarrow$) Suponhamos que C é limitado. Por 16.4 tomando $x \in C$ e $r := \Theta(C)$ segue que $C \subseteq \mathrm{B}\,(x; r)$. Isso implica que, para todo $x' \in C$,

$$\|x'\| = \|x' - 0\| \leq \|x' - x\| + \|x - 0\| \leq \Theta(C) + \|x\|\,,$$

portanto $C \subseteq \mathrm{B}\,(0; \Theta(C) + \|x\|)$, o que implica que $\|C\|$ é limitado.

2. ($\Leftarrow$) Suponhamos que $\|C\|$. Vale então que $\Theta(\|C\|) = \sup \|C\|$ e portanto que $C \subseteq \mathrm{B}\,(0; \Theta\|C\|)$, logo segue de (16.4) que C é limitado. ■

Consideremos um conjunto X e um espaço normado $\boldsymbol{E}$. Como $\boldsymbol{E}$ tem estrutura de espaço métrico, podemos considerar o conjunto das funções limitadas de X para $\boldsymbol{E}$, denotado $\mathscr{B}(X, \boldsymbol{E})$. Como $\boldsymbol{E}$ tem estrutura de espaço linear, o conjunto $\mathcal{F}(X, E)$ das funções $f\colon X \longrightarrow E$ é um espaço linear, com a adição, a inversa aditiva, o zero e a ação por escalar dadas pontualmente: para todas $f, f' \in \mathcal{F}(X, E)$, todo $c \in \mathbb{R}$ e todo $x \in X$, definem-se $(f + f')(x) = f(x) + f'(x)$, $(-f)(x) = -f(x)$, $0(x) = 0$ e $(cf)(x) = cf(x)$. Para mostrar que o espaço das funções limitadas também é um espaço linear, basta mostrarmos que ele é subespaço de $\mathcal{F}(X, E)$.

⊢ **Proposição 19.10.** *Sejam X um conjunto e $\boldsymbol{E}$ um espaço normado. O conjunto $\mathscr{B}(X, \boldsymbol{E})$ das funções limitadas é subespaço linear de $\mathcal{F}(X, E)$.*

□ *Demonstração.* Sejam $f, f' \in \mathscr{B}(X, \boldsymbol{E})$ e $c \in \mathbb{R}$. Como f e f' são limitadas, os conjuntos imagem $f(X)$ e $f'(X)$ são limitados. Isso implica que $(cf + f')(X)$ também é limitado. ■

Além da estrutura linear, podemos definir no espaço $\mathscr{B}(X, \boldsymbol{E})$ uma norma, pois para cada função limitada, como sua imagem é limitada, podemos considerar o supremo de seus valores no espaço normado.

:⊢ **Definição 19.11.** Sejam X um conjunto e $\boldsymbol{E} = (E, \|\cdot\|)$ um espaço normado. A *norma do supremo* em $\mathscr{B}(X, E)$ é

$$\begin{aligned} \|\cdot\|_\infty \colon \mathscr{B}(X, E) &\longrightarrow [0, \infty] \\ f &\longmapsto \|f\|_\infty := \sup_{x \in X} \|f(x)\| \,. \end{aligned}$$

A norma do supremo assume somente valores finitos pois como a imagem de $f(X)$ é limitada, sua imagem em $\mathbb{R}$ pela norma do espaço normado $\boldsymbol{E}$ também é limitada e assim, como $\mathbb{R}$ é completo, esse conjunto tem supremo finito, como mostra a proposição 19.9.

▷ **Exercício 19.2.** *Sejam X um conjunto e $\boldsymbol{E}$ um espaço normado. A norma do supremo $\|\cdot\|_\infty$ é uma norma.*

⊢ **Proposição 19.11.** *Sejam $\boldsymbol{K}$ um espaço topológico compacto e $\boldsymbol{E}$ um espaço normado. Então*

$$\mathscr{C}(\boldsymbol{K}, \boldsymbol{E}) \subseteq \mathscr{B}(K, \boldsymbol{E}).$$

□ *Demonstração.* Seja $f \in \mathscr{C}(\boldsymbol{K}, \boldsymbol{E})$. Como f é contínua e K é compacto, então $f(K) \subseteq E$ é compacto. Como $\|\cdot\|$ é contínua e $f(K)$ é compacto, então $\|f(K)\| \subseteq \mathbb{R}$ é compacto. Mas, como $\mathbb{R}$ é completo, isso significa que $\|f(K)\|$ tem supremo, portanto segue de (19.9) que f é limitada. ■

19.1.4 Bolas e esferas unitárias e topologia

:⊢ **Definição 19.12.** Seja $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado. A *bola unitária* de $(\boldsymbol{E}, \|\cdot\|)$ é o conjunto

$$\mathbb{B} := \{v \in E \mid \|v\| \leq 1\}$$

e a *esfera unitária* de $(\boldsymbol{E}, \|\cdot\|)$ é o conjunto

$$\mathbb{S} := \{v \in E \mid \|v\| = 1\}.$$

A bola unitária é a bola fechada, de raio 1 e centro na origem, com respeito à métrica induzida pela norma. Isto é, $\mathbb{B} = \bar{\mathrm{B}}(0; 1)$. Com essa notação para a bola unitária, podemos representar qualquer bola de centro c e raio r como $c + r\mathbb{B}$, pois

$$\begin{aligned} c + r\mathbb{B} &= \{c + rv \mid v \in \mathbb{B}\} \\ &= \{c + rv \mid \|v\| \leq 1\} \\ &= \left\{v \,\middle|\, \left\|\frac{v-c}{r}\right\| \leq 1\right\} \\ &= \{v \mid \|v - c\| \leq r\} \\ &= \bar{\mathrm{B}}(c; r). \end{aligned}$$

⊢ **Proposição 19.12.** *Seja $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado. A bola unitária $\mathbb{B}$ é um conjunto convexo e centrossimétrico na origem.*

□ *Demonstração.* Sejam $t \in]0, 1[$ e $v, v' \in \mathbb{B}$. Então

$$\|(1-t)v + tv'\| \leq (1-t)\|v\| + t\|v'\| = (1-t) + t = 1,$$

logo $(1-t)v + tv' \in \mathbb{B}$, o que mostra que $\mathbb{B}$ é convexo. Agora, seja $v \in \mathbb{B}$. Então $1 \geq \|v\| = \|-v\|$, logo $-v \in \mathbb{B}$, o que mostra a centrossimetria. ■

A topologia de um espaço normado é dada pela sua norma, através da base de abertos formadas pelas bolas. Essa é a topologia dada pela métrica induzida pela norma.

⊢ **Proposição 19.13.** *Seja $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado.*

1. *A norma $\|\cdot\| : E \to \mathbb{R}$ é uma função contínua.*
2. *$\boldsymbol{E}$ é um espaço linear topológico.*

19.1.5 Equivalência de normas

:⊢ **Definição 19.13.** Seja $\boldsymbol{E}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Normas (*topologicamente*) *equivalentes* em $\boldsymbol{E}$ são normas $\|\cdot\|, \|\cdot\|'$ em $\boldsymbol{E}$ que induzem a mesma topologia em E.

⊢ **Proposição 19.14.** *Seja $\boldsymbol{E}$ um espaço linear sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Normas $\|\cdot\|, \|\cdot\|'$ em $\boldsymbol{E}$ são equivalentes se, e somente se, existem $c, C \in]0, \infty[$ tais que, para todo $v \in E$,*

$$c\,\|v\|' \leq \|v\| \leq C\,\|v\|'.$$

□ *Demonstração.* Demonstraremos a volta pois a ida é evidente. Basta mostrar que as bases de bolas são equivalentes. Para isso, mostramos primeiro que uma bola de uma topologia contém uma bola de outra. Sejam $\|\cdot\|, \|\cdot\|'$ normas equivalentes em $\boldsymbol{E}$. Tomemos uma bola $\mathrm{B}(v; r)$ da norma $\|\cdot\|$. Como existe $C \in]0, \infty[$ tal que, para todo $v \in E$, $\|v\| \leq C\,\|v\|'$, segue que

$$\mathrm{B}\left(v; rC^{-1}\right)' \subseteq \mathrm{B}(v; r),$$

pois se $v' \in \mathrm{B}(v; rC^{-1})'$, então $\|v' - v\| < rC^{-1}$, logo $C\,\|v' - v\| < r$, o que implica que $\|v' - v\| < r$, portanto $v' \in \mathrm{B}(v; r)$.

Sendo assim, tomemos uma bola $\mathrm{B}(v; r)$ da norma $\|\cdot\|$ e $x \in \mathrm{B}(v; r)$. Então temos que $\mathrm{B}(x; r - \|x - v\|) \subseteq \mathrm{B}(v; r)$ e portanto temos que, pelo argumento do parágrafo anterior,

$$\mathrm{B}\left(x; (r - \|x - v\|)C^{-1}\right)' \subseteq \mathrm{B}(x; r - \|x - v\|),$$

logo

$$\mathrm{B}\left(x; (r - \|x - v\|)C^{-1}\right)' \subseteq \mathrm{B}(v; r).$$

Simetricamente, mostra-se que

$$\mathrm{B}\left(x; (r - \|x - v\|')c\right) \subseteq \mathrm{B}(v; r)',$$

portanto as bases de bolas são equivalentes, o que quer dizer que as topologias são a mesma. ■

19.2 Funções lineares contínuas

Nesta seção, trataremos de funções lineares contínuas entre espaços normados. Espaços normados são ao mesmo tempo espaços topológicos e espaços métricos. Por isso, podemos considerar tanto as funções lineares que são contínuas com

respeito à topologia como as que são controladas com respeito à métrica. Vamos a seguir mostrar que para funções lineares essas duas propriedades são equivalentes. Lembremos que uma função $f\colon M \longrightarrow M'$ entre espaços métricos $\boldsymbol{M}$ e $\boldsymbol{M}'$ é controlada quando, para algum $c \in [0, \infty[$, f satisfaz que, para todos $p, p' \in M$,

$$|f(p), f(p')| \leq c\,|p, p'|\,.$$

A distorção de uma função controlada é definida como

$$\langle\!\langle f \rangle\!\rangle = \inf\left\{c \in [0, \infty[\ \mid \forall_{p,p' \in M}\, |f(p), f(p')| \leq c\,|p, p'|\right\}.$$

Essa definição em espaços normados, considerando a distância definida pela norma, é equivalente à seguinte propriedade.

▷ **Exercício 19.3.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *e* $(\boldsymbol{E}', \|\cdot\|')$ *espaços normados e* $L\colon \boldsymbol{E} \longrightarrow \boldsymbol{E}'$ *uma função linear. Então* L *é controlada se, e somente se, para algum* $c \in [0, \infty[$, L *satisfaz que, para todo* $v \in E$,

$$\|L(v)\|' \leq c\,\|v\|\,.$$

Nesse caso, a distorção de L *satisfaz*

$$\|L\| = \inf\left\{c \in [0, \infty[\ \middle|\ \forall_{v \in E}\, \|L(v)\|' \leq c\,\|v\|\right\}.$$

A distorção de uma função linear controlada L é portanto o menor valor tal que vale a desigualdade entre as normas. Vale que

$$\|L(v)\|' \leq \|L\|\,\|v\|\,.$$

Veremos a seguir que essa distorção é uma norma.

O termo usual para funções lineares entre espaços normados que satisfazem a condição anterior é 'limitada', mas esse termo também é usado com um significado diferente para funções entre espaços métricos e para subconjuntos de espaços métricos, incluindo sequências limitadas. Para evitar ambiguidade, adotamos aqui o nome alternativo que expressa igualmente — talvez melhor — a ideia por trás do conceito.

⊢ **Proposição 19.15.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *e* $(\boldsymbol{E}', \|\cdot\|')$ *espaços normados e* $L\colon \boldsymbol{E} \longrightarrow \boldsymbol{E}'$ *uma função linear. Então* L *é controlada se, e somente se, é contínua.*

□ *Demonstração.* Se L é controlada, é uniformemente contínua, portanto contínua (16.35).

Reciprocamente, suponhamos que L é contínua. Para $v = 0$, claramente vale $\|L(0)\| = 0 \leq 0 - \|0\|$. Consideremos o seguinte para $v \neq 0$: como L é contínua,

é contínua em 0; portanto existe $\delta \in]0,\infty[$ tal que, para todo $v \in E$, se $\|v\| \leq \delta$ então $\|L(v)\|' \leq 1$. Sendo assim, seja $v \in E \smallsetminus \{0\}$. Então, como $\left\|\delta \frac{v}{\|v\|}\right\| = \delta$, segue que

$$\|L(v)\|' = \left\|L\left(\frac{\|v\|}{\delta}\frac{\delta}{\|v\|}v\right)\right\|' = \left\|\frac{\|v\|}{\delta}L\left(\delta\frac{v}{\|v\|}\right)\right\|' = \frac{\|v\|}{\delta}\left\|L\left(\delta\frac{v}{\|v\|}\right)\right\|' \leq \frac{1}{\delta}\|v\|,$$

o que mostra que L é controlada. ■

▷ **Exercício 19.4.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *e* $(\boldsymbol{E}', \|\cdot\|')$ *espaços normados e* $L \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$ *uma função linear contínua.*

1. *Para todo* $v \in E$,
$$\|L(v)\|' \leq \|L\|\,\|v\|;$$
2. *Valem as seguintes igualdades:*

$$\begin{aligned}
\|L\| &= \inf\{c \in [0,\infty[\mid \forall_{v\in E}\, \|L(v)\| \leq c\,\|v\|\} \\
&= \sup\{\|L(v)\| \mid v \in \mathbb{B}\} \\
&= \sup\{\|L(v)\| \mid v \in \mathbb{B}^\circ\} \\
&= \sup\{\|L(v)\| \mid v \in \mathbb{S} \cup \{0\}\} \\
&= \sup\{\|L(v)\| \mid v \in \mathbb{S}\} \\
&= \sup\left\{\frac{\|L(v)\|}{\|v\|} \,\middle|\, v \in \mathbb{B} \smallsetminus \{0\}\right\}.
\end{aligned}$$

Note que, caso $E = \{0\}$, os conjuntos das duas últimas igualdades são vazios, e então devemos ressaltar que o supremo é tomado sobre $[0,\infty[$, portanto o supremo de $\emptyset$ é 0.

⊢ **Proposição 19.16.** *Sejam* $\boldsymbol{M}$ *um espaço métrico e* $(\boldsymbol{E}, \|\cdot\|)$ *um espaço normado sobre um corpo normado* $(\boldsymbol{C}, |\cdot|)$. *A função*

$$\begin{aligned}
\|\cdot\| \colon \mathscr{C}_{|\cdot|}(\boldsymbol{M}, \boldsymbol{E}) &\longrightarrow \mathbb{R} \\
f &\longmapsto \|f\|
\end{aligned}$$

é uma seminorma no espaço $\mathscr{C}_{|\cdot|}(\boldsymbol{M}, \boldsymbol{E})$ *das funções controladas de* $\boldsymbol{M}$ *para* $\boldsymbol{E}$.

□ *Demonstração.* 1. (Homogeneidade absoluta) Para $c = 0$ a demonstração é imediata; suponhamos $c \neq 0$. Para todos $p, p' \in M$,

$$\|cf(p) - cf(p')\| = |c|\,\|f(p) - f(p')\| \leq |c|\,\|f\|\,|p, p'|,$$

o que mostra que $\|cf\| \leq |c| \, \|f\|$. Por outro lado,

$$\begin{aligned}\|f(p) - f(p')\| &= |c|^{-1} \, |c| \, \|f(p) - f(p')\| \\ &= |c|^{-1} \, \|cf(p) - cf(p')\| \\ &\leq |c|^{-1} \, \|cf\| \, |p, p'| \, ,\end{aligned}$$

o que mostra que $\|f\| \leq |c|^{-1} \, \|cf\|$. Assim segue que $\|cf\| = |c| \, \|f\|$.

2. (Subaditividade) Para todos $p, p' \in M$,

$$\begin{aligned}\|(f + f')(p) - (f + f')(p')\| &= \|f(p) + f'(p) - f(p') - f'(p')\| \\ &\leq \|f(p) - f(p')\| + \|f'(p) - f'(p')\| \\ &\leq \|f\| \, |p, p'| + \|f'\| \, |p, p'| \\ &= (\|f\| + \|f'\|) \, |p, p'| \, ,\end{aligned}$$

o que mostra que $\|f + f'\| \leq \|f\| + \|f'\|$. ■

⊢ **Proposição 19.17.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *e* $(\boldsymbol{E}', \|\cdot\|')$ *espaços normados. A função*

$$\begin{aligned}\|\cdot\| \colon \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}') &\longrightarrow \mathbb{R} \\ L &\longmapsto \|L\|\end{aligned}$$

é uma norma em $\mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$.

□ *Demonstração.* 1. (Separação) Seja $L \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$ tal que $\|L\| = 0$. De 16.36 segue que L é constante e, como $L(0) = 0$, $L = 0$.

2. (Homogeneidade absoluta) Segue de 19.16.
3. (Subaditividade) Segue de 19.16. ■

19.3 Espaços normados completos

Espaços normados completos são espaços normados que são completos com respeito à métrica induzida pela sua norma. Esses espaços são comumente chamados de 'espaços de Banach'[6]. Aqui adotaremos simplesmente a nomenclatura de espaços normados completos. Os espaços normados completos mais comuns são os espaços lineares finitos com a norma p, sendo $p \in [1, \infty]$. Além desses espaços, um tipo de espaço mais geral (que inclui esses) são os chamados espaços de funções absolutamente p-somáveis. Esses espaços serão descritos nas seções a seguir.

⊢ **Proposição 19.18.** *Sejam* $\boldsymbol{E}$ *um espaço linear sobre um corpo normado* $(\boldsymbol{C}, |\cdot|)$ *e* $\|\cdot\|, \|\cdot\|'$ *normas equivalentes em* $\boldsymbol{E}$. *O espaço* $(\boldsymbol{E}, \|\cdot\|)$ *é completo se, e somente se,* $(\boldsymbol{E}, \|\cdot\|')$ *é completo.*

[6] Em homenagem ao matemático polonês *Stefan Banach* (30/03/1892 – 31/08/1945).

Essa proposição mostra que, além de induzirem a mesma topologia, a propriedade de completude — que faz parte da estrutura uniforme, e não topológica, do espaço — também é induzida por normas equivalentes.

▷ **Exercício 19.5.** *Sejam $\boldsymbol{X}$ um espaço topológico separado compacto e $(\boldsymbol{C}, |\cdot|)$ um corpo normado. A função*

$$\begin{aligned} \|\cdot\| \colon \mathscr{C}(X, C) &\longrightarrow [0, \infty[\\ f &\longmapsto \sup_{x \in X} |f(x)| = \max_{x \in X} |f(x)| \end{aligned}$$

é uma norma em $\mathscr{C}(\boldsymbol{X}, \boldsymbol{C})$ e, se $(\boldsymbol{C}, |\cdot|)$ é completo, o espaço $(\mathscr{C}(\boldsymbol{X}, \boldsymbol{C}), \|\cdot\|)$ é completo.

⊢ **Proposição 19.19.** *Sejam $\boldsymbol{X}$ um espaço topológico compacto e $(\boldsymbol{L}, |\cdot|)$ um espaço linear normado. A função*

$$\begin{aligned} \|\cdot\| \colon \mathscr{C}(X, L) &\longrightarrow [0, \infty[\\ f &\longmapsto \max_{x \in X} |f(x)| \,. \end{aligned}$$

é uma norma no espaço de funções contínuas $\mathscr{C}(\boldsymbol{X}, \boldsymbol{L})$ que induz a mesma topologia do espaço. Se $(\boldsymbol{L}, |\cdot|)$ é completo, $\mathscr{C}(\boldsymbol{X}, \boldsymbol{L})$ é completo.

□ *Demonstração.* Primeiro temos que mostrar que a função $\|\cdot\|$ está bem definida, ou seja, que para toda $f \in \mathscr{C}(X, L)$ existe $m \in [0, \infty[$ tal que $m = \max_{x \in X} |f(x)|$. Como X é compacto, para toda $f \in \mathscr{C}(X, L)$ a imagem $f(X) \subseteq L$ é compacta, portanto $|f(X)| \subseteq [0, \infty[$ é compacto e $\max_{x \in X} |f(x)| \in [0, \infty[$.

Mostremos agora que $\|\cdot\|$ é uma norma em $\mathscr{C}(X, L)$: (Separação) Para toda $f \in \mathscr{C}(X, L)$, se $\|f\| = 0$, então $|f(x)| \leq 0$ para todo $x \in X$, logo $|f(x)| = 0$ para todo $x \in X$, e segue pela separação de $|\cdot|$ que $f(x) = 0$, portanto $f = 0$.

(Homogeneidade absoluta) Para todos $c \in C$ e $f \in \mathscr{C}(X, L)$,

$$\|cf\| = \max_{x \in X} |cf(x)| = \max_{x \in X} |c|\,|f(x)| = |c| \max_{x \in X} |f(x)| = |c|\,\|f\| \,.$$

(Desigualdade triangular) Para todas $f, f' \in \mathscr{C}(X, L)$,

$$\begin{aligned} \|f + f'\| &= \max_{x \in X} |f(x) + f'(x)| \\ &\leq \max_{x \in X} (|f(x)| + |f'(x)|) \\ &= \max_{x \in X} |f(x)| + \max_{x \in X} |f'(x)| \\ &= \|f\| + \|f'\| \,. \end{aligned}$$

Por fim, mostremos que $\mathscr{C}(X, L)$ é completo. Seja $\{f_n\}_{n\in\mathbb{N}}$ uma sequência acumulante[7] em $\mathscr{C}(X, L)$. Notemos que, para cada $x \in X$, $\{f_n(x)\}_{n\in\mathbb{N}}$ é uma sequência acumulante em L, portanto uma sequência convergente, já que L é completo. Isso significa que existe $f \in L^X$ que é o limite pontual de $\{f_n\}_{n\in\mathbb{N}}$: para todo $x \in X$,

$$f(x) = \lim_{n\to\infty} f_n(x).$$

Basta mostrar agora que $\{f_n\}_{n\in\mathbb{N}}$ converge para f em $\mathscr{C}(X, L)$, pois isso mostra também que $f \in \mathscr{C}(X, L)$. Seja $\varepsilon \in]0, \infty[$. Então existe $N \in \mathbb{N}$ tal que, para todos $n, n' \in \mathbb{N}$ tais que $n, n' \geq N$,

$$\|f_{n'} - f_n\| \leq \varepsilon.$$

Logo, para todo $x \in X$,

$$|f(x) - f_n(x)| = \lim_{n'\to\infty} |f_{n'} - f_n| \leq \varepsilon,$$

o que implica que

$$\|f - f_n\| = \max_{x\in X} |f(x) - f_n(x)| \leq \varepsilon.$$

Isso mostra que $f_n \to f$ em $\mathscr{C}(X, L)$. ■

19.3.1 Sequências absolutamente somáveis

:⊢ **Definição 19.14.** Seja $\boldsymbol{L}$ um espaço linear topológico sobre um corpo topológico $\boldsymbol{C}$. Uma *sequência somável* em $\boldsymbol{L}$ é uma sequência $(v_n)_{n\in\mathbb{N}}$ em L tal que a sequência

$$\left(\mathop{+}_{k\in[n]} v_k \right)_{n\in\mathbb{N}}$$

é convergente.

:⊢ **Definição 19.15.** Sejam $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. Uma *sequência absolutamente somável* em $\boldsymbol{E}$ é uma sequência $(v_n)_{n\in\mathbb{N}}$ em E tal que a sequência

$$\left(\mathop{+}_{k\in[n]} \|v_k\| \right)_{n\in\mathbb{N}}$$

é convergente.

⊢ **Proposição 19.20.** *Sejam $(\boldsymbol{E}, \|\cdot\|)$ um espaço normado sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$. O espaço $(\boldsymbol{E}, \|\cdot\|)$ é completo se, e somente se, toda sequência absolutamente $(v_n)_{n\in\mathbb{N}}$ é somável.*

[7] Sequência de Cauchy.

□ *Demonstração.* Suponha que $(\boldsymbol{E}, \|\cdot\|)$ é completo. Seja $(v_n)_{n\in\mathbb{N}}$ um sequência absolutamente somável. Defina a sequência

$$s_n := \left(\mathop{+}_{k\in[n]} v_k \right)_{n\in\mathbb{N}}$$

em E. Para $m > 1$,

$$\|s_m - s_n\| = \left\| \mathop{+}_{k=n}^{m-1} v_k \right\| \leq \mathop{+}_{k=n}^{m-1} \|v_k\| .$$

Como as somas parciais de $+_{n\in\mathbb{N}} \|v_n\|$ formam uma sequência convergente (e portanto aproximante), pois $(v_n)_{n\in\mathbb{N}}$ é absolutamente somável, $+_{k=n}^{m-1} \|v_k\| \to 0$ quando $n, m \to \infty$. Isso mostra que $(s_n)_{n\in\mathbb{N}}$ é aproximante, e da completude de $(\boldsymbol{E}, \|\cdot\|)$ segue que é convergente, o que significa que $(v_n)_{n\in\mathbb{N}}$ é somável.

Reciprocamente, suponha que toda sequência absolutamente somável no espaço $(\boldsymbol{E}, \|\cdot\|)$ é somável. Seja $(v_n)_{n\in\mathbb{N}}$ uma sequência aproximante. Para mostrar que essa sequência converge, basta achar uma subsequência que converge. Escolha $(v_{n_k})_{k\in\mathbb{N}}$ tal que, para todo $k \in \mathbb{N}$, $\left\| v_{n_{k+1}} - v_{n_k} \right\| < 2^{-k}$. Então $+_{k\in\mathbb{N}} \left\| v_{n_{k+1}} - v_{n_k} \right\|$ converge, o que implica que $+_{k\in\mathbb{N}} \left(v_{n_{k+1}} - v_{n_k} \right)$ converge, já que toda sequência absolutamente somável é somável. Isso implica que a sequência

$$v_{n_m} = v_{n_0} + \mathop{+}_{k\in[m]} \left(v_{n_{k+1}} - v_{n_k} \right)$$

converge. Portanto $(v_n)_{n\in\mathbb{N}}$ converge. ■

19.3.2 Espaços normados de dimensão finita

Espaços lineares $\boldsymbol{E}$ de dimensão finita $d \in \mathbb{N}$ sobre um corpo podem ser identificados com $\boldsymbol{C^d}$. Nesses casos, a menos que seja mencionado o contrário, sempre consideraremos a base canônica

$$e_i = (0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0)$$

em $\boldsymbol{C^d}$ e todo vetor $v \in \boldsymbol{C^d}$ será representado como $v = (v_0, \ldots, v_{d-1})$.

:⊢ **Definição 19.16.** Sejam $\boldsymbol{E}$ um espaço linear finito d-dimensional sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$ e $p \in [1, \infty[$. A *norma* p em $\boldsymbol{E}$ é a função

$$\begin{aligned} \|\cdot\|_p \colon E &\longrightarrow \mathbb{R} \\ v &\longmapsto \left(\mathop{+}_{i=0}^{d-1} |v_i|^p \right)^{\frac{1}{p}} . \end{aligned}$$

A *norma* ∞ em $\boldsymbol{E}$ é a função

$$\begin{aligned} \|\cdot\|_\infty \colon E &\longrightarrow \mathbb{R} \\ v &\longmapsto \max_{i\in[d]} |v_i| \,. \end{aligned}$$

Pode-se verificar que $\lim_{p\to\infty} \|v\|_p = \|v\|_\infty$.

⊢ **Proposição 19.21.** *Sejam $\boldsymbol{E}$ um espaço linear finito d-dimensional sobre um corpo normado $(\boldsymbol{C}, |\cdot|)$ e $p \in [1, \infty[$.*

1. *A norma p em $\boldsymbol{E}$ é uma norma;*
2. *Para todo $v \in E$,*
$$\|v\|_\infty \leq \|v\|_p \leq d^{p^{-1}} \|v\|_\infty \,.$$

⊢ **Proposição 19.22.** *Sejam $\boldsymbol{E}$ um espaço vetorial sobre um corpo normado completo $(\boldsymbol{C}, |\cdot|)$.*

1. *A bola $\mathbb{B}$ é compacta se, e somente se, a dimensão de $\boldsymbol{E}$ é finita.*
2. *Se a dimensão de $\boldsymbol{E}$ é finita, todas normas em $\boldsymbol{E}$ são equivalentes;*
3. *Se a dimensão de $\boldsymbol{E}$ é finita, todas as normas fazem de $\boldsymbol{E}$ um espaço completo.*

□ *Demonstração.* 1. Será demonstrado mais adiante.

2. Vamos mostrar que toda norma em $\boldsymbol{E}$ é equivalente a $\|\cdot\|_1$ e, como equivalência de normas é uma relação de equivalência, seguirá que todas normas são equivalentes em $\boldsymbol{E}$. Seja $\|\cdot\|$ uma norma em $\boldsymbol{E}$. Para todo $v \in E$, definindo $c := \max_{i\in[d]} \|e_i\|$, em que $\{e_i\}_{i\in[d]}$ é a base canônica de $\boldsymbol{E}$, segue que

$$\|v\| = \left\| \mathop{+}_{i=0}^{d-1} v_i e_i \right\| \leq \mathop{+}_{i=0}^{d-1} |v_i| \, \|e_i\| \leq c \, \|v\|_1 \,.$$

A outra parte da equivalência segue do fato de que todo conjunto fechado e limitado em $\boldsymbol{C}$ é sequencialmente compacto.
Suponha, por absurdo, que não exista $k \in]0, \infty[$ tal que, para todo $v \in E$, $k^{-1} \|v\|_1 \leq \|v\|$. Assim, para todo $n \in \mathbb{N}$, existe $v_n \in E$ com $\|v_n\|_1 = 1$ e $1 = \|v_n\|_1 > n \|v_n\|$. Como $\mathbb{S}$ é compacta, já que a dimensão é finita, existe subsequência $(v_{n_k})_{k\in\mathbb{N}}$ convergindo a v' em $(\boldsymbol{E}, \|\cdot\|_1)$. Como a norma é contínua, segue que $\|v'\|_1 = 1$. Pela desigualdade anterior, tem-se

$$\|v'\| \leq \|v' - v_{n_k}\| + \|v_{n_k}\| \leq c \, \|v' - v_{n_k}\|_1 + \frac{1}{n_k},$$

que converge para 0 quando $k \to \infty$; ou seja, $\|v'\| = 0$ e portanto $v' = 0$, o que contradiz $\|v'\|_1 = 1$.

3. Como todas as normas são equivalentes, basta provar para $\|\cdot\|_1$. Seja $(v_n)_{n\in\mathbb{N}} = (+_{i\in[d]} v_n^i e_i)_{n\in\mathbb{N}}$ uma sequência aproximante em $(\boldsymbol{E}, \|\cdot\|_1)$. Como

$$\mathop{+}_{i\in[d]} \left|v_n^i - v_{n'}^i\right| = \|v_n - v_{n'}\|_1 ,$$

segue que, para todo $i \in [d]$, a sequência $(v_n^i)_{n\in\mathbb{N}}$ é aproximante em $\boldsymbol{C}$ e então, da completude de $\boldsymbol{C}$, converge para algum $v_\infty^i \in C$. Definindo $v_\infty := +_{i\in[d]} v_\infty^i e_i$ em $\boldsymbol{E}$, segue que

$$\lim_{n\to\infty} \|v_n - v_\infty\|_1 = \lim_{n\to\infty} \mathop{+}_{i\in[d]} \left|v_n^i - v_\infty^i\right| = 0,$$

ou seja, $(v_n)_{n\in\mathbb{N}} \to v_\infty$ e o espaço é completo.

■

Concluímos que um espaço linear normado de dimensão finita tem uma única topologia determinada por norma. Portanto todas noções topológicas relacionadas a espaços normados são independentes da norma escolhida.

19.3.3 Espaços de funções absolutamente somáveis

:⊢ **Definição 19.17.** Sejam X um conjunto, $(\boldsymbol{C}, |\cdot|)$ um corpo normado e $p \in [0, \infty[$. Uma função *absolutamente p-somável* (ou *absolutamente somável na potência p*) é uma função $f\colon X \to C$ tal que

$$\mathop{+}_{x\in X} |f(x)|^p < \infty.$$

O conjunto das funções absolutamente p-somáveis é denotado $\mathscr{S}^p(X, C)$.

Uma função *absolutamente ∞-somável* (ou *absolutamente somável na potência ∞*, ou ainda *essencialmente absolutamente somável*) é uma função $f\colon X \to C$ tal que

$$\sup_{x\in X} |f(x)| < \infty.$$

O conjunto das funções absolutamente ∞-somáveis é denotado $\mathscr{S}^\infty(X, C)$.

Note que as definições implicam que, para todo $p \in [0, \infty]$, as funções $f \in \mathscr{S}^p(X, C)$ são nulas a menos de um subconjunto contável de X; ou seja, têm suporte contável: $\#(\mathrm{supp}(f)) \leq \#(\mathbb{N})$. O conjunto $\mathscr{S}^0(X, C)$ é, de fato, o conjunto de funções com suporte finito, pois a soma será finita se, e somente se, a função tiver uma quantidade finita de valores. O conjunto $\mathscr{S}^\infty(X, C)$ é o conjunto das funções de valor absoluto (ou norma) limitado. Quando $X = \mathbb{N}$ ou $X = \mathbb{Z}$, os espaços são espaços de sequências infinitas unilaterais ou bilaterais, respectivamente. Esses espaços são todos subespaços lineares do espaço C^X de funções $f\colon X \to C$, que é espaço linear já que o corpo $\boldsymbol{C}$ é um espaço linear sobre si mesmo.

$\vdash$ **Proposição 19.23.** *Sejam X um conjunto, $(\boldsymbol{C}, |\cdot|)$ um corpo normado e $p \in [0, \infty[$. O espaço $\mathscr{S}^p(X, C)$ é subespaço linear de C^X.*

$\square$ *Demonstração.* Consideramos dois casos. (1) Seja $p \in [1, \infty[$. Para todos $c \in C$ e $f, f' \in \mathscr{S}^p(X, C)$,

$$\begin{aligned}
\mathop{+}_{x \in X} |(cf + f')(x)|^p &= \mathop{+}_{x \in X} |cf(x) + f'(x)|^p \\
&\leq \mathop{+}_{x \in X} (|c|\,|f(x)| + |f'(x)|)^p \\
&\leq 2^{p-1} \left(|c|^p \mathop{+}_{x \in X} |f(x)|^p + \mathop{+}_{x \in X} |f'(x)|^p \right) \\
&< \infty,
\end{aligned}$$

pois $+_{x \in X} |f(x)|^p < \infty$ e $+_{x \in X} |f'(x)|^p < \infty$. Isso mostra que $cf + f' \in \mathscr{S}^p(X, C)$, portanto que $\mathscr{S}^p(X, C)$ é um espaço linear, subespaço de C^X.

(2) Para todos $c \in C$ e $f, f' \in \mathscr{S}^\infty(X, C)$,

$$\sup_{x \in X} |(cf + f')(x)| \leq |c| \sup_{x \in X} |f(x)| + \sup_{x \in X} |f'(x)| < \infty,$$

pois $\sup_{x \in X} |f(x)| < \infty$ e $\sup_{x \in X} |f'(x)| < \infty$. Isso mostra que $cf + f' \in \mathscr{S}^\infty(X, C)$, portanto que $\mathscr{S}^\infty(X, C)$ é um espaço linear, subespaço de C^X. $\blacksquare$

Embora esses espaços possam ser definidos para quaisquer $p \in [0, \infty]$ e sejam espaços lineares em todos os casos, nem todos esses espaços admitem os mesmos tipos de estrutura além da de espaço linear.

$:\vdash$ **Definição 19.18.** Sejam X um conjunto e $(\boldsymbol{C}, |\cdot|)$ um corpo normado.

1. Para todo $p \in [0, 1[$, a *distância p* entre $f, f' \in \mathscr{S}^p(X, C)$ é

$$|f, f'|_p := \mathop{+}_{x \in X} |f'(x) - f(x)|^p .$$

2. Para todo $p \in [1, \infty[$, a *norma p* de $f \in \mathscr{S}^p(X, C)$ é

$$\|f\|_p := \left(\mathop{+}_{x \in X} |f(x)|^p \right)^{p^{-1}} .$$

A *norma* ∞ de $f \in \mathscr{S}^\infty(X, C)$ é

$$\|f\|_\infty := \sup_{x \in X} |f(x)| .$$

A função $\|\cdot\|_p$ é uma norma em $\mathscr{S}^p(X, C)$, a *norma* p, e p pode ser descrito como a *potência* da norma e do espaço. Quando X é finito, esse espaço é simplesmente o espaço vetorial finito C^X definido anteriormente e as p-normas definidas para espaços de dimensão finita coincidem com essas. Quando $X = \mathbb{N}$ ou $X = \mathbb{Z}$, os espaços são espaços de sequências infinitas unilaterais ou bilaterais, respectivamente.

⊢ **Proposição 19.24.** *Sejam X um conjunto e $(\boldsymbol{C}, |\cdot|)$ um corpo normado completo.*

1. *Para todo $p \in [0, 1[$, o espaço $(\mathscr{S}^p(X, C), |\cdot, \cdot|_p)$ é um espaço métrico completo;*
2. *Para todo $p \in [1, \infty]$, o espaço $(\mathscr{S}^p(X, C), \|\cdot\|_p)$ é um espaço normado completo.*

19.3.4 Espaços de funções absolutamente integráveis

Consideraremos funções de um espaço de medida $\boldsymbol{X}$ para um corpo normado $\boldsymbol{C}$. Esse corpo deve ser entendido, em geral, como $\mathbb{R}$ ou $\mathbb{C}$, pois alguns detalhes não serão especificados, por exemplo qual a estrutura de espaço de medida de um corpo normado qualquer, ou ainda um problema maior, o que é a integral de uma função com valores em um corpo qualquer.

Lembremos que o conjunto de funções mensuráveis de $\boldsymbol{X}$ para $\boldsymbol{C}$ é denotadas $\mathscr{M}(\boldsymbol{X}, \boldsymbol{C})$ e o conjunto das quase funções (classe de equivalência de funções que são iguais a menos de um conjunto de medida nula) mensuráveis é denotado $\mathscr{M}_{\mathring{=}}(\boldsymbol{X}, \boldsymbol{C})$.

Antes da definição a seguir, ressaltamos dois comentários. Primeiro, definimos que elevar um número positivo a 0 dará o seguinte resultado:

$$\begin{aligned} (\cdot)^0 \colon [0, \infty[&\longrightarrow [0, \infty[\\ x &\longmapsto \begin{cases} 0, & x = 0 \\ 1, & x \neq 0. \end{cases} \end{aligned}$$

Isso faz com que, para toda função $f \colon X \to C$,

$$|f|^0 = \mathbf{1}_{\operatorname{supp}(f)}.$$

Segundo, lembremos que o supremo essencial de uma função f é definido por

$$\sup \operatorname{ess}(f) := \inf \left\{ t \in]0, \infty[\;\middle|\; \mathring{\forall}_{x \in X} f(x) \leq t \right\},$$

em que $\mathring{\forall}$ é 'para quase todo', ou seja, existe conjunto nulo N tal que, para todo $x \in X \smallsetminus N$, a propriedade vale.

:⊢ **Definição 19.19.** Sejam $\boldsymbol{X}$ um espaço de medida, $(\boldsymbol{C}, |\cdot|)$ um corpo normado e $p \in [0, \infty[$. Uma função *absolutamente p-integrável*[8] de $\boldsymbol{X}$ para $\boldsymbol{C}$ é uma função $f \in \mathscr{M}(\boldsymbol{X}, \boldsymbol{C})$ tal que

$$\int |f|^p \, \mathrm{d}m < \infty.$$

O conjunto das quase funções absolutamente p-integráveis de $\boldsymbol{X}$ para $\boldsymbol{C}$ é denotado $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$.

Uma função *absolutamente ∞-integrável* é uma função $f \in \mathscr{M}(\boldsymbol{X}, \boldsymbol{C})$ tal que

$$\sup \operatorname{ess}(|f|) < \infty.$$

O conjunto das quase-funções absolutamente ∞-integráveis de $\boldsymbol{X}$ para $\boldsymbol{C}$ é denotado $\mathscr{I}^\infty(\boldsymbol{X}, \boldsymbol{C})$.

Note que, os espaços $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ são espaços de quase funções, ou seja, espaços de classes de equivalência de funções, cuja equivalência é ser igual a menos de um conjunto de medida nula. Na prática, trataremos essas quase funções como funções, mas esse detalhe tem que estar sempre claro para o leitor.

Por definição, para todo $p \in [0, \infty]$, vale que

$$\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}) \subseteq \mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C}),$$

pois $\mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$ é o espaço de quase-funções mensuráveis. As inclusões não são somente de conjuntos, no entanto. De fato, como $\mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$ é um espaço linear, os espaços $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ herdam uma estrutura de espaço linear e pode-se mostrar que eles são subespaços lineares de $\mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$. Inclusões relacionando diferentes espaços $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ e $\mathscr{I}^q(\boldsymbol{X}, \boldsymbol{C})$ não são tão óbvias e não serão abordadas por enquanto. O caso em que $p = 0$ nos dá que $\mathscr{I}^0(\boldsymbol{X}, \boldsymbol{C})$ é o conjunto das funções cujo suporte tem medida finita.

Esta proposição auxiliará a demonstração da proposição seguinte.

⊢ **Proposição 19.25.** *Sejam $a, b \in [0, \infty]$.*

1. *Para todo $p \in [0, 1]$,*
$$(a + b)^p \leq a^p + b^p;$$
2. *Para todo $p \in [1, \infty[$,*
$$(a + b)^p \leq 2^{p-1}(a^p + b^p).$$

[8] Essas funções não recebem esse nome usualmente. O espaço $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ é geralmente chamado de espaço $L^p(\boldsymbol{X}, \boldsymbol{C})$, em homenagem a Henri Lebesgue, embora de acordo com conjunto dos Bourbaki o criador dos espaços tenha sido Frigyes Riesz (https://en.wikipedia.org/wiki/Lp_space).

□ *Demonstração.* 1. Para $p \in [0,1]$, como a função

$$\begin{aligned}(\cdot)^p \colon [0,\infty[&\longrightarrow [0,\infty[\\ t &\longmapsto t^p\end{aligned}$$

é côncava[9] e $0^p = 0 \geq 0$, segue que ela é subaditiva[10].

2. Para $p \in [1,\infty[$, como a função

$$\begin{aligned}(\cdot)^p \colon [0,\infty[&\longrightarrow [0,\infty[\\ t &\longmapsto t^p\end{aligned}$$

é convexa[11], segue que

$$(a+b)^p = 2^p(2^{-1}a + 2^{-1}b)^p \leq 2^p\left(2^{-1}a^p + 2^{-1}b^p\right) = 2^{p-1}(a^p + b^p).$$

■

⊢ **Proposição 19.26.** *Sejam $\boldsymbol{X}$ um espaço de medida, $\boldsymbol{C}$ um corpo normado e $p \in [0,\infty]$. O espaço $\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C})$ é um espaço linear.*

□ *Demonstração.* Demonstraremos que os espaços $\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C})$ são espaços lineares mostrando que são subespaços lineares de $\mathscr{M}_{\mathrm{m}}(\boldsymbol{X},\boldsymbol{C})$. Sejam $c \in C$ e $f, f' \in \mathscr{I}^p(\boldsymbol{X},\boldsymbol{C})$. Consideramos três casos.

1. Seja $p \in [0,1[$. Segue de 19.25 e da homogeneidade absoluta e subaditividade de $|\cdot|$ que
$$|cf + f'|^p \leq ||cf| + |f'||^p \leq |c|^p\,|f|^p + |f'|^p\,.$$
Como $\int |f|^p \,\mathrm{d}m < \infty$ e $\int |f'|^p \,\mathrm{d}m < \infty$,
$$\begin{aligned}\int |cf + f'|^p \,\mathrm{d}m &\leq \int \left(|c|^p\,|f|^p + |f'|^p\right) \mathrm{d}m \\ &= |c|^p \int |f|^p \,\mathrm{d}m + \int |f'|^p \,\mathrm{d}m \\ &< \infty,\end{aligned}$$
o que mostra que $cf + f' \in \mathscr{I}^p(\boldsymbol{X},\boldsymbol{C})$.

[9] Inclusive para $p = 0$ e $p = 1$.

[10] A demonstração é simples e pode ser conferida em https://en.wikipedia.org/wiki/Subadditivity

[11] Note que vale também para $p = 1$.

2. Seja $p \in [1, \infty[$. Segue de 19.25 e da homogeneidade absoluta e subaditividade de $|\cdot|$ que

$$|cf + f'|^p \leq ||cf| + |f'||^p \leq 2^{p-1} \left(|c|^p |f|^p + |f'|^p \right).$$

Como $\int |f|^p \, \mathrm{d}m < \infty$ e $\int |f'|^p \, \mathrm{d}m < \infty$, segue dessa desigualdade que

$$\begin{aligned} \int |cf + f'|^p \, \mathrm{d}m &\leq \int 2^{p-1} \left(|c|^p |f|^p + |f'|^p \right) \mathrm{d}m \\ &= 2^{p-1} \left(|c|^p \int |f|^p \, \mathrm{d}m + \int |f'|^p \, \mathrm{d}m \right) \\ &< \infty, \end{aligned}$$

o que mostra que $cf + f' \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$.

3. Para $p = \infty$, é claro que, como $|cf + f'| \leq |c| \, |f| + |f'|$,

$$\sup \operatorname{ess}(|cf + f'|) \leq |c| \sup \operatorname{ess}(|f|) + \sup \operatorname{ess}(|f'|) < \infty,$$

pois $\sup \operatorname{ess}(|f|) < \infty$ e $\sup \operatorname{ess}(|f'|) < \infty$, o que mostra que $cf + f' \in \mathscr{I}^\infty(\boldsymbol{X}, \boldsymbol{C})$. ■

:⊢ **Definição 19.20.** Sejam $\boldsymbol{X}$ um espaço de medida e $(\boldsymbol{C}, |\cdot|)$ um corpo normado.

1. Para todo $p \in [0, 1[$, a *distância* p entre $f, f' \in \mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$ é

$$|f, f'|_p := \int |f' - f|^p \, \mathrm{d}m.$$

2. Para todo $p \in [1, \infty[$, a *norma* p de $f \in \mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$ é

$$\|f\|_p := \left(\int |f|^p \, \mathrm{d}m \right)^{p^{-1}}.$$

3. A *norma* ∞ de $f \in \mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$ é

$$\|f\|_\infty := \sup \operatorname{ess}(|f|).$$

Esses valores nem sempre são menores que ∞ para qualquer quase-função mensurável. As distâncias p são de fato distâncias quando restritas a $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$, $p \in [0, 1[$, e as normas p são de fato normas quando restritas a $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$, $p \in [1, \infty]$, mas ainda não demonstraremos isso. Para essas demonstrações, precisamos primeiro estabelecer algumas desigualdade clássicas que envolvem esses valores. Essas normas serão avaliadas mais à frente em quase-funções mensuráveis quaisquer,

19.3.4.1 Desigualdades das normas p

Nesta seção, trabalharemos com pares de números $p, q \in [1, \infty]$ tais que

$$p^{-1} + q^{-1} = 1.$$

Esses números p e q são às vezes chamados de 'conjugados de Hölder'. Como $p^{-1} + q^{-1} = 1$, segue que $q = \frac{p}{p-1}$. Além disso, esses são os pares de números cuja média harmônica M_{-1} é igual a 2, já que

$$M_{-1}(p, q) = \frac{2}{p^{-1} + q^{-1}} = 2.$$

Por causa disso, chamaremos esses números de *duais harmônicos*[12]. O fato de 2 ser o dual de 2 será relevante na teoria de espaços com produto interno completos. Sob essa perspectiva, 2 está no 'meio da caminho' entre 1 e ∞. Deve-se comentar a respeito de $p = 1$ ou $p = \infty$. Consideraremos que 1 e ∞ são duais harmônicos, como se tivéssemos $\infty^{-1} = 0$, de modo que $1^{-1} + \infty^{-1} = 1$. Em geral, no entanto, tomaremos cuidado para não realizarmos operações mal definidas com 0 e ∞, e na prática bastará a afirmação de que 1 e ∞ são conjugados.

:⊢ **Definição 19.21.** A função *dual harmônico* em $[1, \infty]$ é

$$\begin{aligned} {}^*\colon [1, \infty] &\longrightarrow [1, \infty] \\ p &\longmapsto p^* = \begin{cases} \infty, & p = 1 \\ \dfrac{p}{p-1}, & p \in \,]1, \infty[\\ 1, & p = \infty \end{cases}. \end{aligned}$$

FIGURA 19.1: Gráfico da função ${}^*\colon [1, \infty] \to [1, \infty], p \mapsto \frac{p}{p-1}$.

[12] A escolha do termo *dual* ficará clara mais adiante, quando forem analisadas as propriedades do espaço dual de $\mathscr{L}^p$

⊢ **Proposição 19.27** ([13]). *Sejam $a, b \in [0, \infty[$ e $p \in]1, \infty[$. Então*

$$ab \leq a^p p^{-1} + b^{p^*} p^{*-1}$$

e a igualdade vale se, e somente se, $a^p = b^{p^*}$.

□ *Demonstração.* Para $a = 0$ ou $b = 0$, a afirmação é claramente verdadeira. Considere $a \neq 0$ e $b \neq 0$. Como $p^{-1} + p^{*-1} = 1$, da concavidade da função logarítmica segue que

$$\log(ab) = \log(a) + \log(b) = p^{-1}\log(a^p) + p^{*-1}\log(b^{p^*}) \leq \log(p^{-1}a^p + p^{*-1}b^{p^*}).$$

e a igualdade vale se, e somente se, $a^p = b^{p^*}$. Como a função exponencial é crescente, conclui-se que $ab \leq a^p p^{-1} + b^{p^*} p^{*-1}$. ■

Essa desigualdade será usada na próxima demonstração.

⊢ **Proposição 19.28** ([14]). *Sejam $\boldsymbol{X}$ um espaço de medida, $(\boldsymbol{C}, |\cdot|)$ um corpo normado e $p \in [1, \infty]$. Para todas funções $f, f' \in \mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$,*

$$\|ff'\|_1 \leq \|f\|_p \|f'\|_{p^*}.$$

□ *Demonstração.* Primeiro, tratamos de alguns casos triviais. Se $\|f\|_p = 0$, então $f = 0$, portanto $ff' = 0$ e segue que $\|ff'\|_1 = 0$, logo a desigualdade vale. O mesmo vale para $\|f'\|_{p^*} = 0$, portanto assumimos que $\|f\|_p \neq 0$ e $\|f'\|_{p^*} \neq 0$. Se $\|f\|_p = \infty$ ou $\|f'\|_{p^*} = \infty$, então $\|f\|_p \|f'\|_{p^*} = \infty$, logo a desigualdade vale. Portanto assumimos também que $\|f\|_p \neq \infty$ e $\|f'\|_{p^*} \neq \infty$. Se $p = \infty$ e $p^* = 1$, então $|ff'| \leq \|f\|_\infty |f'|$ quase sempre, e a desigualdade segue da monotonicidade da integral. O mesmo vale para $p = 1$ e $p^* = \infty$, portanto podemos assumir ainda que $p, p^* \in]1, \infty[$.

Pela desigualdade de produtos 19.27, segue que, para todo $x \in X$,

$$\left|\frac{f(x)}{\|f\|_p}\frac{f'(x)}{\|f'\|_{p^*}}\right| \leq \left|\frac{f(x)}{\|f\|_p}\right|^p p^{-1} + \left|\frac{f'(x)}{\|f'\|_{p^*}}\right|^{p^*} p^{*-1}.$$

Integrando ambos os lados, segue que

$$\frac{\|ff'\|_1}{\|f\|_p \|f'\|_{p^*}} \leq \frac{\|f\|_p^p}{\|f\|_p^p} p^{-1} + \frac{\|f'\|_{p^*}^{p^*}}{\|f'\|_{p^*}^{p^*}} p^{*-1} = p^{-1} + p^{*-1} = 1,$$

portanto $\|ff'\|_1 \leq \|f\|_p \|f'\|_{p^*}$, o que demonstra a proposição. ■

[13] Essa desiguldade é conhecida como 'desigualdade de produtos de Young', em homenagem ao matemático inglês *William Henry Young* (20/10/1863 – 07/07/1942). https://en.wikipedia.org/wiki/Young%27s_inequality_for_products.

[14] Essa desigualdade é conhecida como 'desigualdade de Hölder', em homenagem ao matemático alemão *Otto Ludwig Hölder* (22/12/1859 – 29/08/1937). https://en.wikipedia.org/wiki/H%C3%B6lder%27s_inequality.

19.3.4.2 Os espaços de funções absolutamente integráveis são normados (e completos)

Para todo $p \in [0, \infty]$, os espaços $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ são espaços lineares, pois são subespaços de $\mathscr{M}_{\doteq}(\boldsymbol{X}, \boldsymbol{C})$. Além disso, para $p \in [1, \infty]$, a norma p é de fato uma norma em $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ e, se o corpo normado $\boldsymbol{C}$ for completo, a distância induzida pela norma faz de $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ um espaço normado completo[15]. O ponto mais difícil da demonstração é a subaditividade[16] da norma p.

Para $p \in [0, 1[$, não se pode definir uma norma desse mesmo modo, pois a subaditividade falha, mas a distância p é de fato uma distância em $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ e, se o corpo normado $\boldsymbol{C}$ for completo, $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ é completo.

⊢ **Proposição 19.29.** *Sejam $\boldsymbol{X}$ um espaço de medida e $(\boldsymbol{C}, |\cdot|)$ um corpo normado.*

1. *Para todo $p \in [0, 1[$, o espaço $(\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}), |\cdot, \cdot|_p)$ é um espaço métrico invariante por translação. Se $(\boldsymbol{C}, |\cdot|)$ é completo, então $(\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}), |\cdot, \cdot|_p)$ é completo.*
2. *Para todo $p \in [1, \infty]$. O espaço $(\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}), \|\cdot\|_p)$ é um espaço normado. Se $(\boldsymbol{C}, |\cdot|)$ é completo, então $(\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}), \|\cdot\|_p)$ é completo.*

□ *Demonstração.* 1. Exercício.

2. (Separação) Seja $f \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ tal que $\|f\|_p = 0$. Então $|f|^p = 0$, portanto $f = 0$.

   (Homogeneidade absoluta) Sejam $c \in C$ e $f \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$. Então

$$\begin{aligned}
\|cf\|_p &= \left(\int |cf|^p \,\mathrm{d}m\right)^{p^{-1}} \\
&= \left(|c|^p \int |f|^p \,\mathrm{d}m\right)^{p^{-1}} \\
&= |c| \left(\int |f|^p \,\mathrm{d}m\right)^{p^{-1}} \\
&= |c| \, \|f\|_p \,.
\end{aligned}$$

   (Subaditividade) Sejam $f, f' \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$. Se $\|f + f'\|_p = 0$, então claramente $\|f + f'\|_p = 0 \leq \|f\|_p + \|f'\|_p$. Assumamos então que $\|f + f'\|_p \neq 0$.

[15] Esse resultado de completude é conhecido por 'teorema de Riesz–Fischer', em homenagem ao matemático húngaro *Frigyes Riesz* (22/01/1880 – 28/02/1956) e ao matemático austríaco *Ernst Sigismund Fischer* (12/07/1875 – 14/11/1954). https://en.wikipedia.org/wiki/Riesz%E2%80%93Fischer_theorem

[16] A subaditividade para normas p é conhecida como 'desigualdade de Minkowski', em homenagem ao matemático alemão *Hermann Minkowski* (22/06/1864 – 12/01/1909). https://en.wikipedia.org/wiki/Minkowski_inequality.

Pela subaditividade de $|\cdot|$ e pela desigualdade 19.28, segue que

$$\begin{aligned}
\|f+f'\|_p^p &= \int |f+f'|^p \,\mathrm{d}m \\
&= \int |f+f'|\,|f+f'|^{p-1} \,\mathrm{d}m \\
&\leq \int (|f|+|f'|)\,|f+f'|^{p-1} \,\mathrm{d}m \\
&= \int |f|\,|f+f'|^{p-1} \,\mathrm{d}m + \int |f'|\,|f+f'|^{p-1} \,\mathrm{d}m \\
&= \left\| |f|\,|f+f'|^{p-1} \right\|_1 + \left\| |f'|\,|f+f'|^{p-1} \right\|_1 \\
&\leq \||f|\|_p \left\| |f+f'|^{p-1} \right\|_{p^*} + \||f'|\|_p \left\| |f+f'|^{p-1} \right\|_{p^*} \\
&= \left(\|f\|_p + \|f'\|_p \right) \left\| |f+f'|^{p-1} \right\|_{p^*} \\
&= \left(\|f\|_p + \|f'\|_p \right) \left(\int |f+f'|^{(p-1)p(p-1)^{-1}} \,\mathrm{d}m \right)^{p^{-1}(p-1)} \\
&= \left(\|f\|_p + \|f'\|_p \right) \left(\int |f+f'|^{p} \,\mathrm{d}m \right)^{p^{-1}(p-1)} \\
&= \left(\|f\|_p + \|f'\|_p \right) \left(\|f+f'\|_p \right)^{p-1} .
\end{aligned}$$

Multiplicando por $\left(\|f+f'\|_p\right)^{1-p}$ em ambos os lados, conclui-se que

$$\|f+f'\|_p \leq \|f\|_p + \|f'\|_p .$$

Isso mostra que $(\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C}), \|\cdot\|_p)$ é um espaço normado.
Suponhamos agora que $(\boldsymbol{C}, |\cdot|)$ é completo. Para mostrar que o espaço $(\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C}), \|\cdot\|_p)$ é completo, basta mostrar toda sequência absolutamente somável é somável (19.20). Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência absolutamente somável em $(\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C}), \|\cdot\|_p)$. Consideramos dois casos.

2.1. Para $p \in [1,\infty[$. Notemos que, para todo $k \in \mathbb{N}$, f_k é mensurável, portanto $+_{k\in[n]} |f_k|$ e $\left(+_{k\in[n]} |f_k|\right)^p$ também são mensuráveis. Como

$$0 \leq \mathop{+}_{k\in[n]} |f_k| \leq \mathop{+}_{k\in[n+1]} |f_k| ,$$

e

$$0 \leq \left(\mathop{+}_{k\in[n]} |f_k| \right)^p \leq \left(\mathop{+}_{k\in[n+1]} |f_k| \right)^p ,$$

segue que do Teorema da Convergência Monótona que o limite $+_{k\in\mathbb{N}} |f_k|$

é mensurável e que

$$\begin{aligned}
\left\| \mathop{+}_{k\in\mathbb{N}} |f_k| \right\|_p &= \left(\int \left(\lim_{n\to\infty} \mathop{+}_{k\in[n]} |f_k| \right)^p \mathrm{d}m \right)^{p^{-1}} \\
&= \left(\int \lim_{n\to\infty} \left(\mathop{+}_{k\in[n]} |f_k| \right)^p \mathrm{d}m \right)^{p^{-1}} \\
&= \left(\lim_{n\to\infty} \int \left(\mathop{+}_{k\in[n]} |f_k| \right)^p \mathrm{d}m \right)^{p^{-1}} \\
&= \lim_{n\to\infty} \left(\int \left(\mathop{+}_{k\in[n]} |f_k| \right)^p \mathrm{d}m \right)^{p^{-1}} \\
&= \lim_{n\to\infty} \left\| \mathop{+}_{k\in[n]} |f_k| \right\|_p \\
&\leq \lim_{n\to\infty} \mathop{+}_{k\in[n]} \|f_k\|_p \\
&< \infty,
\end{aligned}$$

em que a segunda igualdade segue da continuidade de $(\cdot)^p$, a terceira segue do Teorema da Convergência Monótona, a quarta segue da continuidade de $(\cdot)^{p^{-1}}$ e a sexta da subaditividade de $\|\cdot\|_p$.
Isso mostra que $+_{k\in\mathbb{N}} |f_k| \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ e, portanto, que para quase todo ponto $x \in X$, $+_{k\in\mathbb{N}} |f_k(x)| \in [0, \infty[$. Assim, para quase todo $x \in X$, a sequência $(f_k(x))_{n\in\mathbb{N}}$ é absolutamente somável em $\boldsymbol{C}$ e, da completude de $\boldsymbol{C}$, segue que é somável e está definido o limite pontual

$$s(x) := \mathop{+}_{k\in\mathbb{N}} f_k(x).$$

A função s está definida para quase todo $x \in X$ e é mensurável. Para todo $n \in \mathbb{N}$, segue da continuidade e da subaditividade de $|\cdot|$ que

$$|s| = \lim_{n\to\infty} \left| \mathop{+}_{k\in[n]} f_k \right| \leq \lim_{n\to\infty} \mathop{+}_{k\in[n]} |f_k| \leq \mathop{+}_{k\in\mathbb{N}} |f_k|,$$

portanto

$$\|s\|_p = \left(\int |s|^p \,\mathrm{d}m \right)^{p^{-1}} \leq \left(\int \left| \mathop{+}_{k\in\mathbb{N}} |f_k| \right|^p \mathrm{d}m \right)^{p^{-1}} = \left\| \mathop{+}_{k\in\mathbb{N}} |f_k| \right\|_p,$$

o que mostra que $s \in \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$.
Temos que

$$\begin{aligned}\left|s - \mathop{+}_{k\in[n]} f_k(x)\right|^p &= \left|\mathop{+}_{k\in\mathbb{N}} f_{n+k}(x)\right|^p \\ &= \left(\mathop{+}_{k\in\mathbb{N}} |f_{n+k}(x)|\right)^p \\ &\leq 2^p \left(\mathop{+}_{k\in\mathbb{N}} |f_k|\right)^p .\end{aligned}$$

Como a função do lado direito é integrável e, para quase todo $x \in X$, $\left|s(x) - +_{k\in[n]} f_k(x)\right|^p \to 0$, do Teorema da Convergência Dominada segue que

$$\int \left|s - \mathop{+}_{k\in[n]} f_k\right|^p \mathrm{d}m \to 0,$$

portanto $\left\|s - +_{k\in[n]} f_k\right\|_p^p \to 0$, o que implica que $\left\|s - +_{k\in[n]} f_k\right\|_p \to 0$.

2.2. Para $p = \infty$, a demonstração se reduz a uma questão de convergência fora de conjuntos quase vazios.

■

19.3.4.3 O caso anterior de espaços de funções absolutamente somáveis

Os espaços $\mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})$ da subseção anterior são casos particulares dos espaços absolutamente p-integráveis. Consideramos o espaço de medida $(X, \wp(X), \#)$, em que X é um conjunto qualquer, o conjunto das partes $\wp(X)$ é a sigma-álgebra de X, e $\#$ é a medida de contagem

$$\begin{aligned}\#\colon \wp(X) &\longrightarrow [0, \infty] \\ C &\longmapsto \begin{cases} \#(M), & \#(M) < \#(\mathbb{N}) \\ \infty, & \#(M) \geq \#(\mathbb{N}) . \end{cases}\end{aligned}$$

Notemos que todas as funções $f \in C^X$ são mensuráveis nesse caso, pois $\wp(X)$ é a maior sigma-álgebra em X. Ainda, temos que

$$\int_X |f|^p \,\mathrm{d}\# = \mathop{+}_{x\in X} |f(x)|^p$$

e

$$\sup \operatorname{ess}(|f|) = \sup_{x\in X} |f(x)| .$$

Também segue que, se $\|f\|_p < \infty$, então $\#(\operatorname{supp}(f)) \leq \#(\mathbb{N})$.

Como o único conjunto de medida nula na medida de contagem é o conjunto vazio, segue que as funções são quase-iguais se, e somente se, elas são iguais, portanto funções e quase-funções representam o mesmo objeto. Sendo assim, as definições dos espaços de funções absolutamente somáveis e de funções absolutamente integráveis coincidem nesse caso, e por isso têm a mesma notação.

19.3.5 Dualidade e mergulho de espaços absolutamente integráveis

Novamente, consideramos $(\boldsymbol{C}, |\cdot|)$ como $\mathbb{R}$ ou $\mathbb{C}$.

⊢ **Proposição 19.30.** *Sejam $\boldsymbol{X}$ um espaço de medida e $(\boldsymbol{C}, |\cdot|)$ um corpo normado.*

1. *Para todo $p \in]1, \infty[$,*

$$\mathscr{I}^{p^*}(\boldsymbol{X}, \boldsymbol{C}) \simeq \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})^*.$$

O isomorfismo de espaços normados é

$$\begin{aligned} I_p \colon \mathscr{I}^{p^*}(\boldsymbol{X}, \boldsymbol{C}) &\longrightarrow \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C})^* \\ f &\longmapsto I_p(f) \colon \mathscr{I}^p(\boldsymbol{X}, \boldsymbol{C}) \longrightarrow C \\ & \qquad\qquad\qquad\qquad f' \longmapsto \int f f' \mathrm{d}m. \end{aligned}$$

2. *Se $\boldsymbol{X}$ é σ-finito,*

$$\mathscr{I}^{\infty}(\boldsymbol{X}, \boldsymbol{C}) \simeq \mathscr{I}^1(\boldsymbol{X}, \boldsymbol{C})^*.$$

□ *Demonstração.* Os detalhes da demonstração não serão explicados aqui, mas a ideia geral é a seguinte. A função I_p é linear e, pela desigualdade 19.28 ela é uma isometria local. Pelo teorema de Radon-Nikodym, pode-se mostrar que essa isometria é sobrejetiva e, portanto, um isomorfismo de espaços normados. ■

A escolha do termo *dual* para p^* fica mais clara agora, já que temos a relação

$$\mathscr{I}^{p^*} \simeq (\mathscr{I}^p)^*.$$

O dual de $\mathscr{I}^{\infty}(\boldsymbol{X}, \boldsymbol{C})$ é um caso mais complicado e não será abordado aqui.

Para $p, q \in [1, \infty]$ tais que $p < q$, devemos entender $\mathscr{I}^p$ como um espaço de funções que estão mais concentradas na origem, que são mais localmente singulares, enquanto as funções de $\mathscr{I}^q$ é um espaço de funções que são mais espalhadas.

⊢ **Proposição 19.31.** *Seja $\boldsymbol{X}$ um espaço de medida e $(\boldsymbol{C}, |\cdot|)$ um corpo normado (completo) e $p, q \in]0, \infty]$ tais que $p < q$.*

1. *$\mathscr{I}^q(\boldsymbol{X},\boldsymbol{C}) \subset \mathscr{I}^p(\boldsymbol{X},\boldsymbol{C})$ se, e somente se, X não contém conjuntos de medida finita mas arbitrariamente grande;*
2. *$\mathscr{I}^p(\boldsymbol{X},\boldsymbol{C}) \subset \mathscr{I}^q(\boldsymbol{X},\boldsymbol{C})$ se, e somente se, X não contém conjuntos de medida não nula mas arbitrariamente pequena.*

19.4 Isometrias lineares

:⊢ **Definição 19.22.** Sejam $(\boldsymbol{E}, \|\cdot\|)$ e $(\boldsymbol{E}', \|\cdot\|')$ espaços normados. Uma *isometria linear local* de $(\boldsymbol{E}, \|\cdot\|)$ para $(\boldsymbol{E}', \|\cdot\|')$ é uma função linear $L\colon E \to E'$ tal que, para todo $v \in E$,

$$\|L(v)\|' = \|v\|\,.$$

O conjunto dessas funções é $\mathscr{L}_{\|\|}(\boldsymbol{E}, \boldsymbol{E}')$. Uma *isometria linear* é uma isometria linear local bijetiva.

Uma isometria linear local é uma isometria local com respeito à distância induzida pela norma, pois, para todos $v, v' \in E$,

$$|L(v), L(v')| = \|L(v) - L(v')\| = \|L(v - v')\| = \|v - v'\| = |v, v'|\,.$$

19.4.1 Os grupos lineares geral e especial de transformações e de isometrias

O espaço normado $\overleftrightarrow{\mathscr{L}}(\boldsymbol{E})$ das transformações lineares contínuas invertíveis é um grupo com respeito à operação de composição, chamado *grupo de transformações lineares de* $(\boldsymbol{E}, \|\cdot\|)$. Esse grupo é geralmente chamado de *grupo linear geral* e denotado $\mathrm{GL}(\boldsymbol{E})$ e, se $\boldsymbol{E} = \boldsymbol{C}^d$, $\boldsymbol{C}$ o corpo de escalares, denota-se $\mathrm{GL}_d(\boldsymbol{C})$.

O conjunto das transformações de $\overleftrightarrow{\mathscr{L}}(\boldsymbol{E})$ que têm determinante unitário é um subgrupo, pois se $\det f = \det f' = 1$, então $\det f' \circ f = \det f' \det f = 1$. Esse grupo é geralmente chamado de *grupo linear especial* e denotado $\mathrm{SL}(\boldsymbol{E})$ e, se $\boldsymbol{E} = \boldsymbol{C}^d$, $\boldsymbol{C}$ o corpo de escalares, denota-se $\mathrm{SL}_d(\boldsymbol{C})$.

O espaço normado $\overleftrightarrow{\mathscr{L}}_{\|\|}(\boldsymbol{E})$ das transformações lineares contínuas invertíveis que preservam a norma é um subgrupo de $\overleftrightarrow{\mathscr{L}}(\boldsymbol{E})$, chamado *grupo linear de isometrias de* $(\boldsymbol{E}, \|\cdot\|)$. Esse grupo é geralmente chamado de *grupo ortogonal* e denotado $\mathrm{O}(\boldsymbol{E})$ e, se $\boldsymbol{E} = \boldsymbol{C}^d$, $\boldsymbol{C}$ o corpo de escalares, denota-se $\mathrm{O}_d(\boldsymbol{C})$.

O conjunto das transformações de $\overleftrightarrow{\mathscr{L}}_{\|\|}(\boldsymbol{E})$ que têm determinante unitário é um subgrupo, pois se $\det f = \det f' = 1$, então $\det f' \circ f = \det f' \det f = 1$. Esse grupo é geralmente chamado de *grupo ortogonal especial* e denotado $\mathrm{SO}(\boldsymbol{E})$ e, se $\boldsymbol{E} = \boldsymbol{C}^d$, C o corpo de escalares, denota-se $\mathrm{SO}_d(\boldsymbol{C})$.

Temos que

$$\begin{aligned}
\mathrm{SL}(\boldsymbol{E}) &\subseteq \mathrm{GL}(\boldsymbol{E}) \\
\mathrm{O}(\boldsymbol{E}) &\subseteq \mathrm{GL}(\boldsymbol{E}) \\
\mathrm{SO}(\boldsymbol{E}) &\subseteq \mathrm{O}(\boldsymbol{E}) \\
\mathrm{SO}(\boldsymbol{E}) &\subseteq \mathrm{SL}(\boldsymbol{E})
\end{aligned}$$

19.4.2 Perturbação de uma transformação linear

⊢ **Proposição 19.32.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *um espaço normado completo,* $X \subseteq E$ *um subconjunto e* $f\colon X \longrightarrow E$ *uma contração* ($\langle\!\langle f \rangle\!\rangle < 1$)*. Para todos* $x \in X$ *e* $r \in \,]0, \infty[$ *tais que* $\bar{\mathrm{B}}\,(x; r) \subseteq X$ *e* $\|f(x) - x\| \leq (1 - \langle\!\langle f \rangle\!\rangle) r$*, a função* f *tem ponto fixo na bola fechada* $\bar{\mathrm{B}}\,(x; r)$*.*

□ *Demonstração.* Basta provar que $f(\bar{\mathrm{B}}\,(x; r)) \subseteq \bar{\mathrm{B}}\,(x; r)$, pois como $\bar{\mathrm{B}}\,(x; r)$ é fechado, é um espaço métrico completo não vazio, e segue do teorema do ponto fixo de contração (16.45) que $f\colon \bar{\mathrm{B}}\,(x; r) \longrightarrow \bar{\mathrm{B}}\,(x; r)$ tem ponto fixo.

Seja $x' \in \bar{\mathrm{B}}\,(x; r)$. Então $\|x' - x\| \leq r$, portanto

$$\begin{aligned}
\|f(x') - x\| &= \|f(x') - f(x) + f(x) - x\| \\
&\leq \|f(x') - f(x)\| + \|f(x) - x\| \\
&\leq \langle\!\langle f \rangle\!\rangle \, \|x' - x\| + (1 - \langle\!\langle f \rangle\!\rangle) r \\
&\leq \langle\!\langle f \rangle\!\rangle r + (1 - \langle\!\langle f \rangle\!\rangle) r \\
&= r,
\end{aligned}$$

o que mostra que $f(x') \in \bar{\mathrm{B}}\,(x; r)$, logo $f(\bar{\mathrm{B}}\,(x; r)) \subseteq \bar{\mathrm{B}}\,(x; r)$. ■

⊢ **Proposição 19.33** (Perturbação da identidade)**.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *um espaço normado completo,* $A \subseteq E$ *aberto e* $f\colon A \longrightarrow E$ *contração* ($\langle\!\langle f \rangle\!\rangle < 1$)*. Então*

$$\begin{aligned}
\mathrm{I} + f\colon A &\longrightarrow (\mathrm{I} + f)(A) \\
v &\longmapsto v + f(v)
\end{aligned}$$

é um mergulho métrico, $\langle\!\langle (\mathrm{I} + f)^{-1} \rangle\!\rangle \leq \frac{\langle\!\langle f \rangle\!\rangle}{1 - \langle\!\langle f \rangle\!\rangle}$ *e* $(\mathrm{I} + f)(A)$ *é aberto. Se* $A = E$*, então* $(\mathrm{I} + f)(A) = E$*.*

□ *Demonstração.* Primeiro mostremos que $\mathrm{I} + f$ é mergulho métrico. Por 19.16 vale $\langle\!\langle \mathrm{I} + f \rangle\!\rangle \leq \langle\!\langle \mathrm{I} \rangle\!\rangle + \langle\!\langle f \rangle\!\rangle = 1 + \langle\!\langle f \rangle\!\rangle$, o que mostra que $\mathrm{I} + f$ é controlada. Para todos $x, x' \in A$,

$$\begin{aligned}
(1 - \langle\!\langle f \rangle\!\rangle) \, \|x' - x\| &= \|x' - x\| - \langle\!\langle f \rangle\!\rangle \, \|x' - x\| \\
&\leq \|x' - x\| - \|f(x') - f(x)\| \\
&\leq \|x' - x + f(x') - f(x)\| \\
&= \|(\mathrm{I} + f)(x') - (\mathrm{I} + f)(x)\|,
\end{aligned}$$

o que mostra que $\mathrm{I}+f$ é injetiva e sua inversa à esquerda é controlada, portanto f é mergulho métrico (16.38). Ainda, $\langle\!\langle(\mathrm{I}+f)^{-1}\rangle\!\rangle \leq \frac{\langle\!\langle f\rangle\!\rangle}{1-\langle\!\langle f\rangle\!\rangle}$.

Agora, mostremos que $(\mathrm{I}+f)(A)$ é aberto. Seja $y_0 \in (\mathrm{I}+f)(A)$. Para algum $x_0 \in A$, $y_0 = (\mathrm{I}+f)(x_0) = x_0 + f(x_0)$. Como A é aberto, para algum $r \in \,]0,\infty[$ vale $\bar{\mathrm{B}}\,(x_0;r) \subseteq A$. Seja $y \in \bar{\mathrm{B}}\,(y_0;(1-\langle\!\langle f\rangle\!\rangle)r)$ e definamos

$$\begin{aligned}\phi_y\colon \bar{\mathrm{B}}\,(x_0;r) &\longrightarrow E\\ x &\longmapsto y - f(x).\end{aligned}$$

A função ϕ_y é uma contração com $\langle\!\langle\phi_y\rangle\!\rangle \leq \langle\!\langle f\rangle\!\rangle$ pois, para todos $x, x' \in \bar{\mathrm{B}}\,(x_0;r)$,

$$\|\phi_y(x') - \phi_y(x)\| = \|y - f(x') - y + f(x)\| = \|f(x') - f(x)\| \leq \langle\!\langle f\rangle\!\rangle\,\|x' - x\|\,.$$

Assim, temos $1 - \langle\!\langle f\rangle\!\rangle \leq 1 - \langle\!\langle\phi_y\rangle\!\rangle$. Como $y \in \bar{\mathrm{B}}\,(y_0;(1-\|f\|)r)$, segue que

$$\begin{aligned}\|\phi_y(x_0) - x_0\| &= \|y - f(x_0) - x_0\|\\ &= \|y - y_0\|\\ &\leq (1-\langle\!\langle f\rangle\!\rangle)r\\ &\leq (1-\langle\!\langle\phi_y\rangle\!\rangle)r.\end{aligned}$$

Assim, por 19.32 existe ponto fixo $\bar{x} \in \bar{\mathrm{B}}\,(x_0;r)$ de ϕ_y, ou seja, $\bar{x} = \phi_y(\bar{x}) = y - f(\bar{x})$, o que equivale a $y = \bar{x} + f(\bar{x}) = (\mathrm{I}+f)(\bar{x})$. Concluímos que $y \in (\mathrm{I}+f)(\bar{\mathrm{B}}\,(x_0;r))$ e portanto que

$$\mathrm{B}\,(y_0;(1-\langle\!\langle f\rangle\!\rangle)r) \subseteq \bar{\mathrm{B}}\,(y_0;(1-\langle\!\langle f\rangle\!\rangle)r) \subseteq (\mathrm{I}+f)(\bar{\mathrm{B}}\,(x_0;r)) \subseteq (\mathrm{I}+f)(A),$$

logo $(\mathrm{I}+f)(A)$ é aberto.

Por fim, se $A = E$, então para todos $x \in A$ e $r \in \,]0,\infty[$ vale $\bar{\mathrm{B}}\,(x;r) \subseteq A$, portanto $\bar{\mathrm{B}}\,(f(x);(1-\langle\!\langle f\rangle\!\rangle)r) \subseteq f(A)$. Para todo $n \in \mathbb{N}$, tomando $r = n \diagup (1-\langle\!\langle f\rangle\!\rangle)$ segue que

$$\bar{\mathrm{B}}\,(f(x);n) = \bar{\mathrm{B}}\,(f(x);(1-\langle\!\langle f\rangle\!\rangle)n \diagup (1-\langle\!\langle f\rangle\!\rangle)) \subseteq f(A),$$

portanto $E = \bigcup_{n\in\mathbb{N}} \bar{\mathrm{B}}\,(f(x);n) \subseteq f(A)$, o que mostra que $f(A) = E$. ■

Em particular, isso mostra que $\mathrm{I}+f\colon A \longrightarrow (\mathrm{I}+f)(A)$ é um homeomorfismo. O conjunto $(\mathrm{I}+f)(A)$ ser aberto equivale à função $\mathrm{I}+f$ ser aberta.

⊢ **Proposição 19.34** (Perturbação de isomorfismo linear)**.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$ *espaços normados completos,* $A \subseteq E$ *aberto,* $L\colon E \longrightarrow E'$ *isomorfismo linear contínuo e* $f\colon A \longrightarrow E'$ *função controlada tal que* $\langle\!\langle L^{-1}\rangle\!\rangle\langle\!\langle f\rangle\!\rangle < 1$*. Então*

$$\begin{aligned}L+f\colon A &\longrightarrow (L+f)(A)\\ v &\longmapsto L(v) + f(v)\end{aligned}$$

é um mergulho métrico e $\langle\!\langle L+f\rangle\!\rangle \leq \frac{\langle\!\langle L\rangle\!\rangle\langle\!\langle f\rangle\!\rangle}{\langle\!\langle L^{-1}\rangle\!\rangle^{-1}-\langle\!\langle f\rangle\!\rangle}$.

□ *Demonstração.* A função $L^{-1} \circ f \colon A \longrightarrow E$ é uma contração pois $\langle\!\langle L^{-1} \circ f \rangle\!\rangle \leq \langle\!\langle L^{-1} \rangle\!\rangle \langle\!\langle f \rangle\!\rangle < 1$. Como $L^{-1} \circ (L+f) = \mathrm{I} + (L^{-1} \circ f)$, segue de 19.33 que $L^{-1} \circ (L+f) \colon A \longrightarrow L^{-1} \circ (L+f)(A)$ é mergulho métrico, $\langle\!\langle L^{-1} \circ (L+f) \rangle\!\rangle \leq \frac{\langle\!\langle L^{-1} \circ f \rangle\!\rangle}{1 - \langle\!\langle L^{-1} \circ f \rangle\!\rangle}$ e $L^{-1} \circ (L+f)(A)$ é aberto. Como L é isomorfismo linear contínuo e $L+f = L \circ L^{-1} \circ (L+f)$, segue que $L+f \colon A \longrightarrow (L+f)(A)$ é mergulho métrico e $(L+f)(A)$ é aberto. Como $\langle\!\langle L^{-1} \circ f \rangle\!\rangle \leq \langle\!\langle L^{-1} \rangle\!\rangle \langle\!\langle f \rangle\!\rangle < 1$, então $1 - \langle\!\langle L^{-1} \circ f \rangle\!\rangle \geq 1 - \langle\!\langle L^{-1} \rangle\!\rangle \langle\!\langle f \rangle\!\rangle > 0$; segue que

$$\langle\!\langle L^{-1} \circ (L+f) \rangle\!\rangle \leq \frac{\langle\!\langle L^{-1} \circ f \rangle\!\rangle}{1 - \langle\!\langle L^{-1} \circ f \rangle\!\rangle} \leq \frac{\langle\!\langle L^{-1} \rangle\!\rangle \langle\!\langle f \rangle\!\rangle}{1 - \langle\!\langle L^{-1} \rangle\!\rangle \langle\!\langle f \rangle\!\rangle} = \frac{\langle\!\langle f \rangle\!\rangle}{\langle\!\langle L^{-1} \rangle\!\rangle^{-1} - \langle\!\langle f \rangle\!\rangle},$$

portanto

$$\langle\!\langle L+f \rangle\!\rangle = \langle\!\langle L \circ L^{-1} \circ (L+f) \rangle\!\rangle \leq \langle\!\langle L \rangle\!\rangle \langle\!\langle L^{-1} \circ (L+f) \rangle\!\rangle \leq \frac{\langle\!\langle L \rangle\!\rangle \langle\!\langle f \rangle\!\rangle}{\langle\!\langle L^{-1} \rangle\!\rangle^{-1} - \langle\!\langle f \rangle\!\rangle}. \quad \blacksquare$$

19.5 Álgebras normadas

Em geral, álgebra são definidas como espaços lineares com um produto bilinear associativo, e nós usaremos a propriedade de associatividade na grande maioria dos casos. No entanto, a propriedade de associatividade é às vezes dispensável, ou não vale (como no caso dos octônios), e por isso ressaltamos abaixo que as definições ainda valem sem a hipótese de associatividade.

:⊢ **Definição 19.23.** Uma *álgebra (associativa) normada* é um par $\mathbb{A} = (\boldsymbol{A}, \|\cdot\|)$ em que $\boldsymbol{A} = (A, \cdot)$ é uma álgebra (associativa) sobre um corpo $\boldsymbol{C} \subseteq \mathbb{C}$ e $\|\cdot\| : A \to \mathbb{R}$ é uma norma em $\boldsymbol{A}$ satisfazendo, para todos $a, a' \in A$,

$$\|aa'\| \leq \|a\| \, \|a'\| \, ;$$

Uma *álgebra unitária normada* é uma álgebra normada $\mathbb{A} = (\boldsymbol{A}, \|\cdot\|)$ tal que $\boldsymbol{A} = (A, \cdot, 1)$ é uma álgebra unitária e

$$\|1\| = 1.$$

Uma *álgebra invertível normada* é uma álgebra unitária normada $\mathbb{A} = (\boldsymbol{A}, \|\cdot\|)$ tal que $(A, \cdot, ^{-1}, 1)$ é uma álgebra invertível e, para todos $a \in A \smallsetminus \{0\}$,

$$\left\|a^{-1}\right\| = \|a\|^{-1} \, .$$

Na proposição a seguir, usamos a associatividade do produto, mas de fato bastaria que o produto fosse alternativo, como no caso dos octônios.

▷ **Exercício 19.6.** *Seja* $\mathbb{A} = (\boldsymbol{A}, \|\cdot\|)$ *uma álgebra (associativa) invertível normada. Para todos* $a, a' \in A$*,*

$$\|aa'\| = \|a\| \, \|a'\| \, .$$

□ *Demonstração.* Se $a = 0$, então $aa' = 0$ e $\|a\| = 0$, logo $\|a\| \, \|a'\| = 0$, o que implica

$$\|aa'\| = \|0\| = 0 = \|a\| \, \|a'\| \, .$$

Se $a \in A \smallsetminus \{0\}$,

$$\begin{aligned}
\|a\| \, \|a'\| &= \|a\| \left\| (a^{-1}a)a' \right\| \\
&= \|a\| \left\| a^{-1}(aa') \right\| \\
&\leq \|a\| \left\| a^{-1} \right\| \|aa'\| \\
&= \|a\| \, \|a\|^{-1} \, \|aa'\| \\
&= \|aa'\| \, .
\end{aligned}$$

Como vale $\|aa'\| \leq \|a\| \, \|a'\|$, concluímos que

$$\|aa'\| = \|a\| \, \|a'\| \, .$$

■

19.5.1 Função exponencial

Consideremos uma álgebra normada unitária completa $\mathbb{A}$. Para todo $x \in A$, a sequência

$$e_n(x) := \mathop{+}_{i=0}^{n} \frac{x^i}{i!}$$

é uma sequência aproximante,[17] já que, para todos $n, n' \in \mathbb{N}$ com $n' > n$,

$$\|e_{n'}(x) - e_n(x)\| = \left\| \mathop{+}_{i=n+1}^{n'} \frac{x^i}{i!} \right\| \leq \mathop{+}_{i=n+1}^{n'} \frac{\|x\|^i}{i!}.$$

Como a álgebra é completa, isso significa que $e_n(x)$ converge para um elemento

$$\mathrm{e}^x := \lim_{n \to \infty} e_n(x) = \mathop{+}_{n \in \mathbb{N}} \frac{x^n}{n!}.$$

:⊢ **Definição 19.24.** Seja $\mathbb{A}$ uma álgebra unitária normada completa. A *função exponencial* em $\mathbb{A}$ é a função

$$\begin{aligned}
\exp \colon A &\longrightarrow A \\
x &\longmapsto \mathrm{e}^x := \mathop{+}_{n \in \mathbb{N}} \frac{x^n}{n!}.
\end{aligned}$$

[17] Sequência de Cauchy.

⊢ **Proposição 19.35.** *Seja $\mathbb{A}$ uma álgebra unitária normada completa.*

1. *Para todo $x \in A$, $\|\exp(x)\| \leq \exp(\|x\|)$;*
2. *A função exponencial* $\exp\colon A \to A$ *é contínua;*
3. $\exp(0) = 1$*;*
4. *Para todos $x, y \in A$ tais que $xy = yx$,*
$$\exp(x+y) = \exp(x)\exp(y);$$
5. *Para todo $x \in A$, $\exp(x)^{-1} = \exp(-x)$;*
6. *Para todo $x \in A$, $\exp(x) = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$.*

□ *Demonstração.* 1. Basta notar que, para todos $x \in A$ e $n \in \mathbb{N}$,
$$\|e_n(x)\| = \left\| \mathop{+}_{i=0}^{n} \frac{x^i}{i!} \right\| \leq \mathop{+}_{i=0}^{n} \frac{\|x\|^i}{i!} = e_n(\|x\|),$$
logo $\|\exp(x)\| \leq \exp(\|x\|)$.

2. Primeiro notamos que, para todos $x, y \in A$,
$$\begin{aligned}\|y^n - x^n\| &= \left\| \mathop{+}_{i=1}^{n} y^{n-i}(y-x)x^{i-1} \right\| \\ &\leq n(\|x\| \vee \|y\|)^{n-1} \|y - x\|,\end{aligned}$$
portanto
$$\begin{aligned}\|\exp(y) - \exp(x)\| &= \left\| \mathop{+}_{n=1}^{\infty} \frac{y^n - x^n}{n!} \right\| \\ &\leq \mathop{+}_{n=1}^{\infty} \frac{\|y^n - x^n\|}{n!} \\ &\leq \mathop{+}_{n=1}^{\infty} \frac{n(\|x\| \vee \|y\|)^{n-1}}{n!} \|y - x\| \\ &= \mathrm{e}^{(\|x\| \vee \|y\|)} \|y - x\|,\end{aligned}$$
o que mostra que exp é contínua.

3. Segue de
$$\exp(0) = \mathop{+}_{n\in\mathbb{N}} \frac{0^n}{n!} = 1 + \mathop{+}_{n\in\mathbb{N}^*} \frac{0^n}{n!} = 1.$$

4. Sejam $n \in \mathbb{N}$ e
$$s_n := e_{2n}(x+y) - e_n(x)e_n(y) = \mathop{+}_{k=0}^{2n} \frac{(x+y)^k}{k!} - \left(\mathop{+}_{i=1}^{n} \frac{x^i}{i!}\right)\left(\mathop{+}_{j=1}^{n} \frac{y^j}{j!}\right).$$

Como $xy = yx$,

$$\frac{(x+y)^k}{k!} = \frac{1}{k!}\mathop{+}_{i=0}^{k}\frac{k!}{i!(k-i)!}x^i y^{k-i} = \mathop{+}_{i+j=k}\frac{x^i y^j}{i!j!}.$$

Portanto

$$\begin{aligned} s_n &= \mathop{+}_{0\le i+j\le 2n}\frac{x^i y^j}{i!j!} - \mathop{+}_{0\le i\le n, 0\le j\le n}\frac{x^i y^j}{i!j!} \\ &= \mathop{+}_{k=0}^{n-1}\frac{x^k}{k!}\mathop{+}_{n+1}^{2n-k}\frac{y^j}{j!} + \mathop{+}_{k=0}^{n-1}\frac{y^k}{k!}\mathop{+}_{n+1}^{2n-k}\frac{x^j}{j!}, \end{aligned}$$

o que implica que

$$\begin{aligned} \|s_n\| &\le \mathop{+}_{k=0}^{n-1}\frac{\|x\|^k}{k!}\mathop{+}_{n+1}^{2n-k}\frac{\|y\|^j}{j!} + \mathop{+}_{k=0}^{n-1}\frac{\|y\|^k}{k!}\mathop{+}_{n+1}^{2n-k}\frac{\|x\|^j}{j!} \\ &= \mathop{+}_{k=0}^{2n}\frac{(\|x\|+\|y\|)^k}{k!} - \left(\mathop{+}_{i=1}^{n}\frac{\|x\|^i}{i!}\right)\left(\mathop{+}_{j=1}^{n}\frac{\|y\|^j}{j!}\right) \\ &= e_{2n}(\|x\|+\|y\|) - e_n(\|x\|)e_n(\|y\|). \end{aligned}$$

Assim segue que

$$\|\exp(x+y) - \exp(x)\exp(y)\| \le \mathrm{e}^{\|x\|\|y\|} - \mathrm{e}^{\|x\|}\mathrm{e}^{\|y\|} = 0.$$

5. Basta notar que, como $x(-x) = (-x)x$,

$$\exp(x)\exp(-x) = \exp(x-x) = \exp(0) = 1.$$

6. Exercício. ■

19.5.1.1 Exponencial dos complexos

Seja $x = x_0 + x_1\mathrm{i} \in \mathbb{R}^2$. Como $\mathbb{R}^2$ é comutativo,

$$\mathrm{e}^x = \mathrm{e}^{x_0 + x_1\mathrm{i}} = \mathrm{e}^{x_0}\mathrm{e}^{x_1\mathrm{i}}.$$

Devemos calcular, portanto, somente a exponencial $\mathrm{e}^{x_1\mathrm{i}}$. Como $\mathrm{i}^2 = -1$, segue que, para todo $n \in \mathbb{N}$,

$$\begin{aligned} (x_1\mathrm{i})^{2n} &= (x_1)^{2n}(\mathrm{i})^{2n} = (-1)^n(x_1)^{2n} \\ (x_1\mathrm{i})^{2n+1} &= (x_1)^{2n+1}(\mathrm{i})^{2n+1} = (-1)^n(x_1)^{2n+1}\mathrm{i}, \end{aligned}$$

o que implica que

$$\begin{aligned}
\mathrm{e}^{x_1\mathrm{i}} &= \mathop{+}_{n\in\mathbb{N}} \frac{(x_1\mathrm{i})^n}{n!} \\
&= \mathop{+}_{n\in\mathbb{N}} \frac{(x_1\mathrm{i})^{2n}}{(2n)!} + \mathop{+}_{n\in\mathbb{N}} \frac{(x_1\mathrm{i})^{2n+1}}{(2n+1)!} \\
&= \mathop{+}_{n\in\mathbb{N}} (-1)^n \frac{(x_1)^{2n}}{(2n)!} + \mathop{+}_{n\in\mathbb{N}} (-1)^n \frac{(x_1)^{2n+1}}{(2n+1)!}\mathrm{i} \\
&= \cos(x_1) + \operatorname{sen}(x_1)\mathrm{i}.
\end{aligned}$$

Concluímos que

$$\mathrm{e}^x = \mathrm{e}^{x_0}\mathrm{e}^{x_1\mathrm{i}} = \mathrm{e}^{x_0}(\cos(x_1) + \operatorname{sen}(x_1)\mathrm{i}).$$

Logaritmo Pela decomposição polar, para todo $x \in \mathbb{R}^2 \setminus \{0\}$,

$$x = \|x\| \left(\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}\right).$$

Assim, definimos

$$\log(x) := \log(\|x\|) + \theta\mathrm{i},$$

Concluímos que

$$\begin{aligned}
x &= \|x\| \left(\cos(\theta) + \operatorname{sen}(\theta)\mathrm{i}\right) \\
&= \mathrm{e}^{\log(\|x\|)}\mathrm{e}^{\theta\mathrm{i}} \\
&= \mathrm{e}^{\log(\|x\|)+\theta\mathrm{i}} \\
&= \mathrm{e}^{\log(x)}.
\end{aligned}$$

No entanto, é importante notar que, para todo $n \in \mathbb{N}$,

$$\mathrm{e}^{(\theta+n\tau)\mathrm{i}} = \cos(\theta + n\tau) + \operatorname{sen}(\theta + n\tau)\mathrm{i} = \cos(\theta) + \operatorname{sen}(\theta)\mathrm{i} = \mathrm{e}^{\theta\mathrm{i}}$$

19.5.1.2 Exponencial dos quatérnios

Seja $x = x_0 + \vec{x} \in \mathbb{R}^4$, em que $x_0 = x_0$ e $\vec{x} = x_1\mathrm{i}_1 + x_2\mathrm{i}_2 + x_3\mathrm{i}_3$. Como $x_0\vec{x} = \vec{x}x_0$, segue que

$$\mathrm{e}^x = \mathrm{e}^{x_0+\vec{x}} = \mathrm{e}^{x_0}\mathrm{e}^{\vec{x}}.$$

Devemos calcular, portanto, somente a exponencial $\mathrm{e}^{\vec{x}}$. Lembremos que $\hat{x} = \frac{\vec{x}}{\|\vec{x}\|}$ e que $\hat{x} \in \vec{\mathbb{S}}^2$, portanto $\hat{x}^2 = -1$. Segue que, para todo $n \in \mathbb{N}$,

$$\begin{aligned}
(\vec{x})^{2n} &= (\|\vec{x}\|\,\hat{x})^{2n} = (-1)^n \|\vec{x}\|^{2n} \\
(\vec{x})^{2n+1} &= (\|\vec{x}\|\,\hat{x})^{2n+1} = (-1)^n \|\vec{x}\|^{2n+1}\,\hat{x},
\end{aligned}$$

o que implica que

$$\begin{aligned}
\mathrm{e}^{\vec{x}} &= \mathop{+}_{n\in\mathbb{N}} \frac{(\vec{x})^n}{n!} \\
&= \mathop{+}_{n\in\mathbb{N}} \frac{(\vec{x})^{2n}}{(2n)!} + \mathop{+}_{n\in\mathbb{N}} \frac{(\vec{x})^{2n+1}}{(2n+1)!} \\
&= \mathop{+}_{n\in\mathbb{N}} (-1)^n \frac{\|\vec{x}\|^{2n}}{(2n)!} + \mathop{+}_{n\in\mathbb{N}} (-1)^n \frac{\|\vec{x}\|^{2n+1}}{(2n+1)!}\hat{x} \\
&= \cos(\|\vec{x}\|) + \operatorname{sen}(\|\vec{x}\|)\hat{x}.
\end{aligned}$$

Concluímos que

$$\mathrm{e}^{x} = \mathrm{e}^{x_0}\mathrm{e}^{\vec{x}} = \mathrm{e}^{x_0}\left(\cos(\|\vec{x}\|) + \operatorname{sen}(\|\vec{x}\|)\hat{x}\right).$$

Logaritmo Pela decomposição polar, para todo $x \in \mathbb{R}^4 \smallsetminus \{0\}$,

$$x = \|x\|\left(\cos(\phi_0 + \operatorname{sen}(\phi_0)\hat{x})\right).$$

Assim, definimos

$$\log(x) := \log(\|x\|) + \phi_0\hat{x}.$$

Denotando $y := \log(x)$, vale $y_0 = \log(\|x\|)$ e $\vec{y} = \phi_0\hat{x}$, portanto $\|\vec{y}\| = \|\phi_0\hat{x}\| = \phi_0$ e

$$\hat{y} = \frac{\vec{y}}{\|\vec{y}\|} = \frac{\phi_0\hat{x}}{\phi_0} = \hat{x}.$$

Concluímos que

$$\begin{aligned}
x &= \|x\|\left(\cos(\phi_0 + \operatorname{sen}(\phi_0)\hat{x})\right) \\
&= \mathrm{e}^{\log(\|x\|)}\mathrm{e}^{\phi_0\hat{x}} \\
&= \mathrm{e}^{\log(\|x\|)+\phi_0\hat{x}} \\
&= \mathrm{e}^{\log(x)}.
\end{aligned}$$

19.6 Funções multilineares

Lembremos que, para espaços lineares topológicos $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$. o conjunto das funções multilineares contínuas de $\boldsymbol{E}_0 \times \cdots \times \boldsymbol{E}_{n-1}$ para $\boldsymbol{E}'$ é denotado $\mathscr{L}(\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}; \boldsymbol{E}')$ e, caso $\boldsymbol{E}_0 = \cdots = \boldsymbol{E}_{n-1} = \boldsymbol{E}$, denota-se $\mathscr{L}^n(\boldsymbol{E}, \boldsymbol{E}')$.

:⊢ **Definição 19.25.** Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados. Um função n-linear *controlada* de $\boldsymbol{E}_0 \times \cdots \times \boldsymbol{E}_{n-1}$ para $\boldsymbol{E}'$ é uma função n-linear

$$L\colon \boldsymbol{E}_0 \times \cdots \times \boldsymbol{E}_{n-1} \longrightarrow \boldsymbol{E}'$$

tal que algum $c \in [0, \infty[$ satisfaz que, para todos $v_0 \in E_0$, …, $v_{n-1} \in E_{n-1}$,

$$\|L(v_0, \ldots, v_{n-1})\| \leq c \|v_0\| \cdots \|v_{n-1}\| .$$

A *norma* de uma função controlada L é

$$\|L\| := \inf \left\{ c \in [0, \infty[\;\middle|\; \forall_{v_0 \in E_0, \ldots, v_{n-1} \in E_{n-1}} \|L(v_0, \ldots, v_{n-1})\| \leq c \|v_0\| \cdots \|v_{n-1}\| \right\} .$$

▷ **Exercício 19.7.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$, $\boldsymbol{E}'$ espaços normados e*

$$L\colon \boldsymbol{E}_0 \times \cdots \times \boldsymbol{E}_{n-1} \longrightarrow \boldsymbol{E}'$$

uma função n-linear. São equivalentes

1. *L é controlada;*
2. *L é contínua;*
3. *L é contínua em 0.*

⊢ **Proposição 19.36.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ espaços normados de dimensão finita, $\boldsymbol{E}$ um espaço normado e $L\colon \boldsymbol{E}_0 \times \cdots \times \boldsymbol{E}_{n-1} \longrightarrow \boldsymbol{E}'$ uma função n-linear. Então L é contínua.*

□ *Demonstração.* Mostraremos que L é controlada. Para todo $i \in [n]$, sejam $d_i := \dim E_i$ e $(b_j^{(i)})_{j \in [d_i]}$ uma base ordenada de $\boldsymbol{E_i}$. Todas normas em $\boldsymbol{E_i}$ são equivalentes, portanto usaremos a norma $\|\cdot\|_\infty$. Assim, para todos $v_1 \in E_1$, …, $v_{n-1} \in E_{n-1}$,

$$\begin{aligned}
\|L(v_0, \ldots, v_{n-1})\| &= \left\| \mathop{\Bigg+}_{\substack{0 \leq k_0 < d_0 \\ \cdots \\ 0 \leq k_{n-1} < d_{n-1}}} v_{(0)}^{k_0} \cdots v_{(n-1)}^{k_{n-1}} L\Big(b_{k_0}^{(0)}, \cdots, b_{k_{n-1}}^{(n-1)}\Big) \right\| \\
&\leq \mathop{\Bigg+}_{\substack{0 \leq k_0 < d_0 \\ \cdots \\ 0 \leq k_{n-1} < d_{n-1}}} \left| v_{(0)}^{k_0} \right| \cdots \left| v_{(n-1)}^{k_{n-1}} \right| \left\| L\Big(b_{k_0}^{(0)}, \cdots, b_{k_{n-1}}^{(n-1)}\Big) \right\| \\
&\leq \|v_0\| \cdots \|v_{n-1}\| \mathop{\Bigg+}_{\substack{0 \leq k_0 < d_0 \\ \cdots \\ 0 \leq k_{n-1} < d_{n-1}}} \left\| L\Big(b_{k_0}^{(0)}, \cdots, b_{k_{n-1}}^{(n-1)}\Big) \right\| .
\end{aligned}$$

Definindo

$$c := \mathop{\Bigg+}_{\substack{0 \leq k_0 < d_0 \\ \cdots \\ 0 \leq k_{n-1} < d_{n-1}}} \left\| L\Big(b_{k_0}^{(0)}, \cdots, b_{k_{n-1}}^{(n-1)}\Big) \right\| ,$$

segue que

$$\|L(v_0, \ldots, v_{n-1})\| \leq c \|v_0\| \cdots \|v_{n-1}\| .$$

■

19.6.1 Espaço de funções multilineares contínuas

Lembremos que para espaços lineares $\boldsymbol{V}$ e $\boldsymbol{V}'$, definimos indutivamente para todo $n \in \mathbb{N}$ os espaços

- $\mathcal{L}^0(\boldsymbol{V}, \boldsymbol{V}') := \boldsymbol{V}'$;
- $\mathcal{L}^m(\boldsymbol{V}, \boldsymbol{V}') := \mathcal{L}(\boldsymbol{V}, \mathcal{L}^{m-1}(\boldsymbol{V}, \boldsymbol{V}'))$.

Cada espaço $\mathcal{L}^m(\boldsymbol{V}, \boldsymbol{V}')$ é linearmente isomorfo ao espaço $\mathcal{L}^{(m)}(\boldsymbol{V}, \boldsymbol{V}')$ das funções m-lineares de V^m para V'. O isomorfismo é dado por $I^{(m)}$, que é definida indutivamente por

- Para $m = 1$, $I^{(1)} := \mathrm{I}\colon \mathcal{L}^1(\boldsymbol{V}, \boldsymbol{V}') \longrightarrow \mathcal{L}^{(1)}(\boldsymbol{V}, \boldsymbol{V}')$;
- Para $m > 1$, $I^{(m)}\colon \mathcal{L}^m(\boldsymbol{V}, \boldsymbol{V}') \longrightarrow \mathcal{L}^{(m)}(\boldsymbol{V}, \boldsymbol{V}')$ é definida, para cada $T \in \mathcal{L}^m(\boldsymbol{V}, \boldsymbol{V}')$, por

$$\begin{aligned} I^{(m)}(T)\colon V \times \cdots \times V &\longrightarrow V' \\ (v_0, \ldots, v_{m-1}) &\longmapsto I^{(m-1)}(T(v_0))(v_1, \ldots, v_{m-1}). \end{aligned}$$

Isso nos dá a relação

$$I^{(m)}(T)(v_0, \ldots, v_{m-1}) = I^{(m-1)}(T(v_0))(v_1, \ldots, v_{m-1}) = T(v_0) \cdots (v_{m-1}).$$

Definimos espaços análogos para funções lineares contínuas.

:⊢ **Definição 19.26.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados. Definimos indutivamente para todo $m \in \mathbb{N}$:

- $\mathscr{L}^0(\boldsymbol{E}, \boldsymbol{E}') := \boldsymbol{E}'$;
- $\mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}') := \mathscr{L}(\boldsymbol{E}, \mathscr{L}^{m-1}(\boldsymbol{E}, \boldsymbol{E}'))$.

Os mesmos isomorfismo lineares $I^{(m)}\colon \mathcal{L}^m(\boldsymbol{E}, \boldsymbol{E}') \longrightarrow \mathcal{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}')$, quando restritos a $\mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}')$ e $\mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}')$, são isomorfismos isométricos, como mostraremos a seguir.

⊢ **Proposição 19.37.** *Sejam $\boldsymbol{E}$, $\boldsymbol{E}'$ e $\boldsymbol{E}''$ espaços normados. A função*

$$\begin{aligned} I\colon \mathscr{L}(\boldsymbol{E}, \mathcal{L}(\boldsymbol{E}', \boldsymbol{E}'')) &\longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'') \\ f &\longmapsto I(f)\colon E \times E' \longrightarrow E'' \\ &\qquad\qquad (v, v') \longmapsto f_v(v') \end{aligned}$$

é um isomorfismo isométrico.

□ *Demonstração.* O isomorfismo linear segue de 9.34 pelas funções

$$\begin{aligned} I\colon \mathcal{L}(\boldsymbol{E}, \mathcal{L}(\boldsymbol{E}', \boldsymbol{E}'')) &\longrightarrow \mathcal{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'') \\ f &\longmapsto I(f)\colon E \times E' \longrightarrow E'' \\ &\qquad\qquad (v, v') \longmapsto f_v(v') \end{aligned}$$

e

$$\begin{aligned} I^{-1}\colon \mathcal{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'') &\longrightarrow \mathcal{L}(\boldsymbol{E}, \mathcal{L}(\boldsymbol{E}', \boldsymbol{E}'')) \\ f &\longmapsto J(f)\colon E \longrightarrow \mathcal{L}(\boldsymbol{E}', \boldsymbol{E}'') \\ &\qquad\qquad v \longmapsto J(f)_v\colon E' \longrightarrow E'' \\ &\qquad\qquad\qquad\qquad v' \longmapsto f(v, v'). \end{aligned}$$

Devemos mostrar que essas funções se restringem a $\mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}''))$ e $\mathscr{L}(\boldsymbol{E}', \boldsymbol{E}'')$ e que I é isometria.

Assim como em 9.34, denotaremos uma função $f \in \mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}''))$ por

$$\begin{aligned} f\colon E &\longrightarrow \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}'') \\ v &\longmapsto f_v\colon E' \longrightarrow E'' \\ &\qquad\qquad v' \longmapsto f_v(v'). \end{aligned}$$

- (I é homometria) Seja $f \in \mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}''))$. Por 9.34 vale que $I(f) \in \mathcal{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'')$. Como f é contínua vale que, para todo $v \in E$, $\|f_v\| \leq \|f\| \, \|v\|$; como $f_v \in \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}'')$ é contínua vale que, para todo $v' \in E'$, $\|f_v(v')\| \leq \|f_v\| \, \|v'\|$. Disso segue que, como $I(f)(v, v') = f_v(v')$,

  $$\|I(f)(v, v')\| = \|f_v(v')\| \leq \|f_v\| \, \|v'\| \leq \|f\| \, \|v\| \, \|v'\| \, .$$

  Portanto $\|I(f)\| \leq \|f\|$, o que mostra que $I(f) \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'')$. Concluímos portanto que

  $$\|I\| = \sup_{f \in \mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}'')) \smallsetminus \{0\}} \frac{\|I(f)\|}{\|f\|} \leq 1.$$

- (I^{-1} é homometria) Seja $f \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'')$. Por 9.34 vale que $J(f) \in \mathcal{L}(\boldsymbol{E}, \mathcal{L}(\boldsymbol{E}', \boldsymbol{E}''))$. Como f é contínua vale que, para todos $v \in E$ e $v' \in E'$, $\|f(v, v')\| \leq \|f\| \, \|v\| \, \|v'\|$. Disso segue que, como $I^{-1}(f)_v(v') = f(v, v')$,

  $$\left\| I^{-1}(f)_v(v') \right\| = \|f(v, v')\| \leq \|f\| \, \|v\| \, \|v'\| \, .$$

  Portanto $\|I^{-1}(f)_v\| \leq \|f\| \, \|v\|$, o que mostra que $I^{-1}(f)_v \in \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}'')$, e $\|I^{-1}(f)\| \leq \|f\|$, o que mostra que $I^{-1}(f) \in \mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}', \boldsymbol{E}''))$. Concluímos portanto que

  $$\left\| I^{-1} \right\| = \sup_{f \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}'; \boldsymbol{E}'') \smallsetminus \{0\}} \frac{\|I^{-1}(f)\|}{\|f\|} \leq 1.$$

■

Analogamente a 9.44, segue por indução que, para todos espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$ e todo $m \in \mathbb{N}$, são isometricamente isomorfismos

$$\mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}') \simeq \mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}').$$

⊢ **Proposição 19.38.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados e $m \in \mathbb{N} \smallsetminus \{0\}$. A função*

$$I^{(m)} \colon \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}') \longrightarrow \mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}')$$

é isomorfismo isométrico.

□ *Demonstração.* Segue de 19.37 por indução em m. ■

19.7 Diferenciação em espaços normados

A derivada de uma função real $f \colon I \longrightarrow \mathbb{R}$, em que $I \subseteq \mathbb{R}$ é um intervalo, pode ser generalizada para funções entre espaços normados de duas formas principais:

1. a diferencial $\mathrm{D}f$ de uma função $f \colon A \longrightarrow E'$, em que $A \subseteq E$ é um aberto e $\boldsymbol{E}, \boldsymbol{E}'$ são espaços normados, é uma transformação linear em cada ponto do domínio de f;
2. a derivada $\dot{\partial} f$ de uma função (ou trajetória) $f \colon I \longrightarrow E$, em que $I \subseteq \mathbb{R}$ é intervalo e $\boldsymbol{E}$ é um espaço normado, que é também uma função (ou trajetória) em $\boldsymbol{E}$.

Esses dois conceitos estão intimamente relacionados, mas são distintos, pois a diferencial produz, a partir de uma função entre espaços normados, uma família de transformações lineares entre esses espaços, uma para cada ponto do domínio, enquanto a derivada produz, a partir de uma trajetória em um espaço normado, uma outra trajetória no mesmo espaço.

Nesta seção, definiremos esses dois conceitos e mostraremos como eles se relacionam.

19.7.1 Diferenciabilidade

19.7.1.1 Diferencial e propriedades

Consideremos espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$ e um aberto $A \subseteq E$. A ideia por trás dessa definição de diferenciabilidade é a de que a função $f \colon A \longrightarrow E'$ pode ser aproximada em uma vizinhança de um ponto $x \in A$ por seu valor no ponto mais o valor de uma função linear aplicada num vetor $v \in E$ de variação que mede quanto a função se afastou do ponto x. Ser aproximada, nesse sentido, quer dizer que o erro da aproximação será da ordem da norma do vetor variação v, de modo que a

razão entre os dois vá a zero quando a variação vai a zero. A definição de função contínua pode ser pensada como um caso análogo de ordem menor: a função f numa vizinhança do ponto x pode ser aproximada por seu valor em x, e aproximada aqui quer dizer que a norma da diferença vai a zero quando o vetor variação vai a zero.

:⊢ **Definição 19.27** (Diferenciabilidade)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$. Uma função de A para E' *diferenciável em* X (ou *diferenciável em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para todo $x \in X$, alguma função linear contínua $D\colon E \longrightarrow E'$ satisfaz

$$\lim_{v \to 0} \frac{f(x+v) - f(x) - D(v)}{\|v\|} = 0. \tag{19.1}$$

Uma função *diferenciável* de A para E' é uma função $f\colon A \longrightarrow E'$ diferenciável em A.

Notemos que o limite 19.1 acima também poderia ser calculado como

$$\lim_{x' \to x} \frac{f(x') - f(x) - D(x' - x)}{\|x' - x\|} = 0.$$

Vale notar também que na equação 19.1 poderíamos calcular o limite usando a norma de $\boldsymbol{E}'$ que o resultado seria equivalente, porque a definição de convergência em espaços normados é dada pela métrica induzida pela norma. Na literatura adota-se às vezes o limite com ambas as normas, mas alguns cálculos ficam mais simples se a definição acima é adotada. Usaremos a definição com a norma também quando for conveniente para algum cálculo.

▷ **Exercício 19.8** (Diferenciabilidade — alternativa)**.** *Sejam* $(\boldsymbol{E}, \|\cdot\|)$ *e* $(\boldsymbol{E}', \|\cdot\|')$ *espaços normados,* $A \subseteq E$ *um aberto,* $X \subseteq A$ *e* $f\colon A \longrightarrow E'$*. A função* f *é diferenciável em* X *se, e somente se, para todo* $x \in X$*, alguma função linear contínua* $D\colon E \longrightarrow E'$ *satisfaz*

$$\lim_{v \to 0} \frac{\|f(x+v) - f(x) - D(v)\|'}{\|v\|} = 0.$$

Pode-se mostrar também que a noção de diferenciabilidade coincide para normas topologicamente equivalentes tanto no espaço $\boldsymbol{E}$ como no espaço $\boldsymbol{E}'$. Em particular, em espaços normados de dimensão finita todas normas são equivalentes e portanto a noção de diferenciabilidade é sempre a mesma.

▷ **Exercício 19.9** (Equivalência de normas)**.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$ *espaços normados,* $\|\cdot\|_0$ *e* $\|\cdot\|_1$ *normas equivalentes em* $\boldsymbol{E}$*,* $\|\cdot\|_0'$ *e* $\|\cdot\|_1'$ *normas equivalentes em* $\boldsymbol{E}'$*,* $A \subseteq E$ *um aberto,* $X \subseteq A$ *e* $f\colon A \longrightarrow E'$*. São equivalentes:*

- *Para todo $x \in X$, alguma função linear contínua $D\colon E \longrightarrow E'$ satisfaz*
$$\lim_{v\to 0} \frac{\|f(x+v) - f(x) - D(v)\|_0'}{\|v\|_0} = 0.$$
- *Para todo $x \in X$, alguma função linear contínua $D\colon E \longrightarrow E'$ satisfaz*
$$\lim_{v\to 0} \frac{\|f(x+v) - f(x) - D(v)\|_1'}{\|v\|_1} = 0.$$

A transformação linear na definição de uma função diferenciável é única, como mostraremos a seguir, e é denominada a diferencial da função no ponto. Além disso, vale também notar que permitimos $\dim \boldsymbol{E} = 0$ na nossa definição, mas que nesse caso simplesmente definimos que todas funções são diferenciáveis e suas diferenciais são nulas.

⊢ **Proposição 19.39** (Unicidade da diferencial)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $x \in A$ e $f\colon A \longrightarrow E'$ uma função diferenciável em x . Uma única função linear contínua $D\colon E \longrightarrow E'$ satisfaz 19.1.*

□ *Demonstração.* Sejam $D, D'\colon E \longrightarrow E'$ funções lineares que satisfazem 19.1. Nesse caso,
$$\begin{aligned}
&\lim_{v\to 0} \frac{D(v) - D'(v)}{\|v\|} = \\
&\quad = \lim_{v\to 0} \frac{D(v) - (f(x+v) - f(x)) + (f(x+v) - f(x)) - D'(v)}{\|v\|} \\
&\quad = -\lim_{v\to 0} \frac{f(x+v) - f(x) - D(v)}{\|v\|} + \lim_{v\to 0} \frac{f(x+v) - f(x) - D'(v)}{\|v\|} \\
&\quad = 0.
\end{aligned}$$

Como D e D' são funções lineares, vale que $D(0) = 0 = D'(0)$; resta mostrar que, para todo $v \in E \smallsetminus \{0\}$, $D(v) = D'(v)$. Seja $v \in E \smallsetminus \{0\}$. Então $\|v\| \neq 0$ e, para todo $t \in \mathbb{R}$, como D e D' são lineares vale que $D(tv) = tD(v)$ e $D'(tv) = tD'(v)$. Ainda, como $\lim_{t\to 0} tv = 0$, segue do limite calculado acima que
$$\begin{aligned}
\|D(v) - D'(v)\| &= \|v\| \frac{\|D(v) - D'(v)\|}{\|v\|} \\
&= \|v\| \lim_{t\to 0} \frac{|t|\,\|D(v) - D'(v)\|}{|t|\,\|v\|} \\
&= \|v\| \lim_{t\to 0} \frac{\|D(tv) - D'(tv)\|}{\|tv\|} \\
&= 0.
\end{aligned}$$
o que implica $D(v) = D'(v)$. Portanto $D = D'$. ■

:⊢ **Definição 19.28** (Diferencial)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função diferenciável em X. Para todo $x \in X$, a *diferencial de f em x* é a única função linear contínua

$$\mathrm{D}f|_x\colon E \longrightarrow E'$$

que satisfaz

$$\lim_{v\to 0} \frac{f(x+v) - f(x) - \mathrm{D}f|_x(v)}{\|v\|} = 0.$$

A *diferencial* de f é a função

$$\begin{aligned} \mathrm{D}f\colon X &\longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}') \\ x &\longmapsto \mathrm{D}f|_x. \end{aligned}$$

Como comentado anteriormente, a definição de diferenciabilidade pode ser entendida como uma generalização da continuidade, mas para uma ordem superior de aproximação. A seguir, mostramos que se uma função é diferenciável — ou seja, pode ser aproximada dessa forma superior — então ela é contínua —pode ser aproximada da forma inferior.

⊢ **Proposição 19.40** (Diferenciabilidade implica continuidade)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função diferenciável em X. Então f é contínua em X.*

□ *Demonstração.* Seja $x \in X$. Como $\mathrm{D}f|_x$ é linear e contínua, $\lim_{v\to 0} \mathrm{D}f|_x(v) = 0$, portanto segue que

$$\begin{aligned} \lim_{v\to 0}(f(x+v) - f(x)) &= \lim_{v\to 0}(f(x+v) - f(x) - \mathrm{D}f|_x(v)) \\ &= \lim_{v\to 0} \|v\| \, \frac{f(x+v) - f(x) - \mathrm{D}f|_x(v)}{\|v\|} \\ &= 0. \end{aligned}$$

■

19.7.1.2 Operações com funções diferenciáveis

A partir da estrutura algébrica dos espaços vetoriais (ou álgebras) dos espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$, podem-se induzir operações entre funções desses espaços, como a adição de funções, a multiplicação de função por escalar e, no caso do espaço imagem ser uma álgebra, a multiplicação de funções. Quando realizamos essas operações entre funções diferenciáveis, as funções resultantes ainda são diferenciáveis e a expressão de sua diferencial em termos das diferenciais das funções usadas na operação são muito simples. Essas propriedades são de fato análogas a propriedades de uma derivação em uma álgebra (ou anel): são lineares e satisfazem uma regra

do produto, ou regra de derivação. No entanto, para que a diferencial seja de fato uma derivação, algumas restrições devem ser feitas; as propriedades a seguir podem portanto ser compreendidas como casos mais gerais das propriedades de uma derivação, pois são pontuais.

⊢ **Proposição 19.41** (Propriedades de derivação)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$.*

1. *Para todas funções $f, f' \colon A \longrightarrow E'$ diferenciáveis em X, a função*

$$f + f' \colon A \longrightarrow E'$$

é diferenciável em X e, para todo $x \in X$,

$$\mathrm{D}(f + f')|_x = \mathrm{D}f|_x + \mathrm{D}f'|_x;$$

2. *Para todo $c \in \mathbb{R}$ e toda função $f \colon A \longrightarrow E'$ diferenciável em X, a função*

$$cf \colon A \longrightarrow E'$$

é diferenciável em X e, para todo $x \in X$,

$$\mathrm{D}(cf)|_x = c\mathrm{D}f|_x;$$

3. *Se $\boldsymbol{E}'$ é uma álgebra normada, para todas funções $f, f' \colon A \longrightarrow E'$ diferenciáveis em X, a função*

$$ff' \colon A \longrightarrow E'$$

é diferenciável em X e, para todo $x \in X$ e todo $v \in E$,

$$\mathrm{D}(ff')|_x(v) = \mathrm{D}f|_x(v)f'(x) + f(x)\mathrm{D}f'|_x(v);$$

4. *Toda função constante $f \colon A \longrightarrow E'$ é diferenciável e, para todo $x \in A$,*

$$\mathrm{D}f|_x = 0;$$

5. *Toda função linear $l \colon E \longrightarrow E'$ é diferenciável e, para todo $x \in E$,*

$$\mathrm{D}l|_x = l.$$

□ *Demonstração.* Exercício.

■

Além das operações entre funções diferenciáveis provenientes da estrutura algébrica dos espaços vetoriais (ou álgebras) em questão, uma operação possível entre funções diferenciáveis é a composição. Como é de se esperar, a composição de funções diferenciáveis é diferenciável e a expressão de sua diferencial em termos das diferenciais da funções envolvidas na composição é conhecida como a *regra da cadeia*, pois as diferenciais são 'encadeadas', calculadas uma em seguida da outra e compostas. Essa regra é uma das fórmulas mais úteis na teoria de diferenciação.

$\vdash$ **Proposição 19.42** (Diferenciabilidade da composição ou Regra da cadeia)**.** *Sejam $\boldsymbol{E}$, $\boldsymbol{E}'$ e $\boldsymbol{E}''$ espaços normados, $A \subseteq E$ e $A' \subseteq E'$ abertos, $X \subseteq A$, $f\colon A \longrightarrow E'$ diferenciável em X tal que $f(A) \subseteq A'$ e $f\colon A' \longrightarrow E''$ diferenciável em $f(X)$. A função composta $f' \circ f\colon A \longrightarrow E''$ é diferenciável em X e, para todo $x \in X$,*

$$\mathrm{D}(f' \circ f)|_x = \mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x.$$

$\square$ *Demonstração.* Seja $x \in X$. Definamos, para todo $v \in E$,

$$\begin{aligned} r(v) &:= f(x+v) - f(x) - \mathrm{D}f|_x(v), \\ r'(v) &:= f'(f(x)+v) - f'(f(x)) - \mathrm{D}f'|_{f(x)}(v), \\ r''(v) &:= (f' \circ f)(x+v) - (f' \circ f)(x) - \mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x(v), \end{aligned}$$

de modo que as diferenciabilidades de f em x e de f' em $f(x)$ equivalem a

$$\lim_{v \to 0} \frac{r(v)}{\|v\|} = \lim_{v \to 0} \frac{r'(v)}{\|v\|} = 0$$

e, como $\mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x\colon E \longrightarrow E''$ é linear e contínua, o limite

$$\lim_{v \to 0} \frac{r''(v)}{\|v\|} = 0$$

implica a diferenciabilidade de $f' \circ f$ em x. Calculemos esse limite.

Calculando $(f' \circ f)(x+v)$, obtemos

$$\begin{aligned} (f' \circ f)(x+v) &= f'(f(x+v)) \\ &= f'(f(x) + \mathrm{D}f|_x(v) + r(v)) \\ &= f'(f(x)) + \mathrm{D}f'|_{f(x)}(\mathrm{D}f|_x(v) + r(v)) + r'(\mathrm{D}f|_x(v) + r(v)) \\ &= (f' \circ f)(x) + \mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x(v) + \mathrm{D}f'|_{f(x)}(r(v)) \\ &\quad + r'(\mathrm{D}f|_x(v) + r(v)), \end{aligned}$$

portanto

$$\begin{aligned} r''(v) &= (f' \circ f)(x+v) - (f' \circ f)(x) - \mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x(v) \\ &= \mathrm{D}f'|_{f(x)}(r(v)) + r'(\mathrm{D}f|_x(v) + r(v)). \end{aligned}$$

e

$$\lim_{v\to 0}\frac{r''(v)}{\|v\|}=\lim_{v\to 0}\frac{\mathrm{D}f'|_{f(x)}(r(v))+r'(\mathrm{D}f|_x(v)+r(v))}{\|v\|}.$$

A primeira parcela do limite é 0, pois da linearidade e continuidade de $\mathrm{D}f|_x$ segue que

$$\lim_{v\to 0}\frac{\mathrm{D}f'|_{f(x)}(r(v))}{\|v\|}=\lim_{v\to 0}\mathrm{D}f'|_{f(x)}\left(\frac{r(v)}{\|v\|}\right)=\mathrm{D}f'|_{f(x)}\left(\lim_{v\to 0}\frac{r(v)}{\|v\|}\right)=0.$$

A segunda parcela do limite também é 0, pois como $\lim_{v\to 0}\mathrm{D}f|_x(v)+r(v)=0$ e $\mathrm{D}f|_x\left(\frac{v}{\|v\|}\right)$ é limitado, segue que

$$\begin{aligned}\lim_{v\to 0}\frac{r'(\mathrm{D}f|_x(v)+r(v))}{\|v\|}&=\lim_{v\to 0}\frac{r'(\mathrm{D}f|_x(v)+r(v))}{\|\mathrm{D}f|_x(v)+r(v)\|}\frac{\|\mathrm{D}f|_x(v)+r(v)\|}{\|v\|}\\&=\lim_{v\to 0}\frac{r'(\mathrm{D}f|_x(v)+r(v))}{\|\mathrm{D}f|_x(v)+r(v)\|}\left\|\mathrm{D}f|_x\left(\frac{v}{\|v\|}\right)+\frac{r(v)}{\|v\|}\right\|\\&=0.\end{aligned}$$

Segue então que

$$\lim_{v\to 0}\frac{(f'\circ f)(x+v)-(f'\circ f)(x)-\mathrm{D}f'|_{f(x)}\circ\mathrm{D}f|_x(v)}{\|v\|}=0,$$

e concluímos que $f'\circ f$ é diferenciável em x e $\mathrm{D}f'|_{f(x)}\circ\mathrm{D}f|_x$ é sua diferencial em x. ■

⊢ **Proposição 19.43** (Restrição de contradomínio)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $\boldsymbol{S}'\subseteq\boldsymbol{E}'$ subespaço fechado, $A\subseteq E$ um aberto, $X\subseteq A$ e $f\colon A\longrightarrow E'$ uma função diferenciável em X tal que $f(A)\subseteq S'$. Então $f\colon A\longrightarrow S'$ é diferenciável em X.*

□ *Demonstração.* Exercício. ■

19.7.1.3 Diferenciabilidade contínua

Como diferenciabilidade implica continuidade, se f é diferenciável então f é contínua, ou seja, $f\in\mathscr{C}(\boldsymbol{A},\boldsymbol{E})$. O espaço $\mathscr{L}(\boldsymbol{E},\boldsymbol{E}')$ das funções lineares contínuas de $\boldsymbol{E}$ para $\boldsymbol{E}'$ é um espaço normado e então podemos nos perguntar se a diferencial $\mathrm{D}f\colon A\longrightarrow\mathscr{L}(\boldsymbol{E},\boldsymbol{E}')$ de f é também contínua. Isso nem sempre vale.

A seguir, quando falamos de uma vizinhança de um conjunto X, queremos dizer um conjunto V tal que $X\subseteq V^\circ$, o que equivale a unirmos vizinhanças de cada ponto $x\in X$.

:⊢ **Definição 19.29** (Diferenciabilidade contínua)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$. Uma função de A para E' $\mathscr{C}^1$*-diferenciável* (ou *continuamente diferenciável*) *em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para alguma vizinhança $V \subseteq A$ de X, f é diferenciável em V e sua diferencial $\mathrm{D}f\colon V \longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$ é contínua em X.

Uma função $\mathscr{C}^1$*-diferenciável* (ou *continuamente diferenciável*) de A para E' é uma função $f\colon A \longrightarrow E'$ $\mathscr{C}^1$-diferenciável em A e o conjunto dessas funções é denotado $\mathscr{C}^1(\boldsymbol{A}, \boldsymbol{E}')$.

▷ **Exercício 19.10.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$ *espaços normados,* $A \subseteq E$ *um aberto e* $X \subseteq A$*.*

1. *Para todas funções* $f, f'\colon A \longrightarrow E'$ $\mathscr{C}^1$*-diferenciáveis em* X*, a função* $f + f'\colon A \longrightarrow E'$ *é* $\mathscr{C}^1$*-diferenciável em* X*;*
2. *Para todo* $c \in \mathbb{R}$ *e toda função* $f\colon A \longrightarrow E'$ $\mathscr{C}^1$*-diferenciável em* X*, a função* $cf\colon A \longrightarrow E'$ *é* $\mathscr{C}^1$*-diferenciável em* X*;*
3. *Se* $\boldsymbol{E}'$ *é uma álgebra normada, para todas funções* $f, f'\colon A \longrightarrow E'$ $\mathscr{C}^1$*-diferenciáveis em* X*, a função* $ff'\colon A \longrightarrow E'$ *é* $\mathscr{C}^1$*-diferenciável em* X*;*
4. *Toda função constante* $f\colon A \longrightarrow E'$ *é* $\mathscr{C}^1$*-diferenciável;*
5. *Toda função linear* $l\colon E \longrightarrow E'$ *é* $\mathscr{C}^1$*-diferenciável.*

⊢ **Proposição 19.44** (Diferenciabilidade contínua da composição)**.** *Sejam* $\boldsymbol{E}$*,* $\boldsymbol{E}'$ *e* $\boldsymbol{E}'$ *espaços normados,* $A \subseteq E$ *e* $A' \subseteq E'$ *abertos,* $X \subseteq A$*,* $f\colon A \longrightarrow E'$ $\mathscr{C}^1$*-diferenciável em* X *tal que* $f(A) \subseteq A'$ *e* $f\colon A' \longrightarrow E''$ $\mathscr{C}^1$*-diferenciável em* $f(X)$*. A função composta* $f' \circ f\colon A \longrightarrow E''$ *é* $\mathscr{C}^1$*-diferenciável em* X*.*

□ *Demonstração.* Segue de 19.42 e do fato de que composição de funções contínuas é contínua. ■

19.7.1.4 Diferenciabilidade forte

Quando definimos diferenciabilidade de uma função, definimos primeiro para um ponto ou conjunto de pontos no domínio, sem levar em consideração a diferenciabilidade dos pontos na vizinhança desse ponto. A definição de diferenciabilidade contínua, porém, não pode ser definida somente pontualmente, pois para que a diferencial $\mathrm{D}f$ de uma função f seja contínua em um ponto, ela primeiro precisa estar definida em uma vizinhança desse ponto. Assim, se torna necessário que a função f seja diferenciável numa vizinhança do ponto em consideração. Isso faz com que as proposições envolvendo diferenciabilidade contínua da função não possam ser enunciadas pontualmente, mas sim para funções diferenciáveis em uma vizinhança do ponto.

Existe, no entanto, uma definição mais forte de diferenciabilidade que pode ser definida pontualmente e que, quando válida numa vizinhança do ponto, coincide

com a propriedade de diferenciabilidade contínua. Isso permite que essa propriedade estenda resultados geralmente apresentados para funções $\mathscr{C}^1$-diferenciáveis. Casos notáveis são o Teorema da Função Inversa e o Teorema da Função Implícita, que serão apresentados mais adiante.

:⊢ **Definição 19.30** (Diferenciabilidade forte)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$. Uma função de A para $\boldsymbol{E}'$ *fortemente diferenciável* em X (ou em x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ diferenciável em X tal que, para todo $x \in X$ e $\varepsilon \in \mathbb{R}_{>0}$, algum $\delta \in \mathbb{R}_{>0}$ satisfaz que a função resto restrita a $\mathrm{B}\,(0;\delta)$, dada por

$$\begin{aligned} r_x\colon \mathrm{B}\,(0;\delta) &\longrightarrow E' \\ v &\longmapsto f(x+v) - f(x) - \mathrm{D}f|_x(v), \end{aligned}$$

é controlada e sua distorção satisfaz $\langle\!\langle r_x \rangle\!\rangle \leq \varepsilon$.

A definição adotada por Lima em seu *Curso de Análise* é equivalente a essa, mas como usaremos a propriedade acima mais usualmente, essa equivalência pode ser conferida em [Lim14, p. 279].

A proposição a seguir é a mais importante aqui, pois mostra a equivalência dos conceito de diferenciabilidade contínua e diferenciabilidade forte quando em uma vizinhança do ponto, portanto justifica a afirmação de que o segunda generaliza o primeira.

⊢ **Proposição 19.45.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função diferenciável. A função f é fortemente diferenciável em X se, e somente se, é continuamente diferenciável em X.*

□ *Demonstração.* Seja $x \in X$. Suponhamos que f é continuamente diferenciável em x, o que significa que $\mathrm{D}f$ é contínua em x. A função resto em x é dada por

$$r_x(v) = f(x+v) - f(x) - \mathrm{D}f|_x(v)$$

e, como é soma de funções diferenciáveis, é diferenciável e

$$\mathrm{D}r_x|_v = \mathrm{D}f|_{x+v} - \mathrm{D}f|_x.$$

Isso implica que $\mathrm{D}r_x|_0 = 0$ e, como $\mathrm{D}f$ é contínua em x, $\mathrm{D}r_x$ é contínua em 0. Assim, para todo $\varepsilon \in \mathbb{R}_{>0}$, algum $\delta \in \mathbb{R}_{>0}$ satisfaz que, para todo $v \in \mathrm{B}\,(0;\delta)$,

$$\|\mathrm{D}r_x|_v\| = \|\mathrm{D}r_x|_v - \mathrm{D}r_x|_0\| < \varepsilon.$$

Segue então de 19.65 que r_x é controlada e sua distorção satisfaz $\langle\!\langle r \rangle\!\rangle \leq \varepsilon$, o que mostra que f é fortemente diferenciável em x.

A recíproca é exercício, mas sua demonstração pode ser conferida em [Lim14, p. 279]. ■

19.7.1.5 Diferenciabilidade da norma

A norma é uma função $\|\cdot\| : E \longrightarrow \mathbb{R}$ e portanto pode-se considerar sua diferenciabilidade. Em um ponto $x \in E$, a diferencial de $\|\cdot\|$ em x é denotada $\mathrm{D}\,\|\cdot\|\,|_x$. Em geral, a diferenciabilidade vai depender da norma que se adota no espaço normado $\boldsymbol{E}$. Mesmo normas equivalentes, que de acordo com 19.9 geram a mesma classe de funções diferenciáveis, podem diferir quanto à própria diferenciabilidade em algum ponto. Como veremos mais adiante, para $p \in]1, \infty[$, as normas p em $\mathbb{R}^d$ são diferenciáveis em todos os pontos além da origem. Para $p = 1$ e $p = \infty$, a situação é um pouco mais complexa, embora todas essas normas sejam equivalentes. No entanto, sempre podemos afirmar que a norma não é diferenciável na origem. Os resultados dessa seção serão apresentados sem as demonstrações, mas todas podem ser conferidas em [Col12, §2.6].

$\vdash$ **Proposição 19.46.** *Seja $\boldsymbol{E} = (E, \|\cdot\|)$ um espaço normado. A norma $\|\cdot\| : E \longrightarrow \mathbb{R}$ não é diferenciável em 0.*

Para normas em espaços arbitrários, não podemos estabelecer resultados muito específicos, mas a seguinte proposição dá uma condição necessária para que a norma seja diferenciável em algum ponto: ela deve ser diferenciável na semirreta positiva iniciada na origem e passando pelo ponto em questão. Isso resume o estudo da diferenciabilidade para os pontos na esfera de norma 1, por exemplo. Além disso, a norma de operador dessa da diferencial, uma norma no espaço dual E^*, é igual a 1.

$\vdash$ **Proposição 19.47.** *Sejam $\boldsymbol{E} = (E, \|\cdot\|)$ um espaço normado e $x \in E \smallsetminus \{0\}$. Se a norma $\|\cdot\|$ é diferenciável em x, então, para todo $c \in \mathbb{R}_{>0}$, a norma $\|\cdot\|$ é diferenciável em cx*

$$\|\mathrm{D}\,\|\cdot\|\,|_x\|_{E^*} = 1.$$

19.7.2 Derivabilidade

19.7.2.1 Derivada de uma trajetória

Para a definição de derivada, definiremos antes o conceito de uma trajetória em um espaço normado. A ideia é formalizar as propriedades do movimento de uma partícula ou objeto em um espaço que seja parametrizado pelo tempo. Aqui, definiremos uma trajetória simplesmente como uma função de um intervalo real, que representará o parâmetro temporal, para um espaço normado, que representará o espaço. Em geral, espera-se que a trajetória de uma partícula seja de certa forma suave, ou ao menos contínua, mas por simplicidade de notação não imporemos nenhuma condição sobre a função, para que a partir de definições adicionais sobre a continuidade e derivabilidade de uma tal trajetória possamos obter condições adequadas para tais trajetórias. Na prática, quase todas trajetórias que considerarmos serão contínuas ou mesmo deriváveis.

:⊢ **Definição 19.31.** Sejam $\boldsymbol{E}$ um espaço normado. Uma *trajetória em* E é uma função $\gamma\colon I \longrightarrow E$ de um intervalo aberto $I \subseteq \mathbb{R}$ para E.

Na literatura matemática, o termo *curva* às vezes denomina o que aqui chamamos de trajetória, em geral impondo a condição de continuidade, e outras vezes denomina o conjunto imagem de uma trajetória que seja bem comportada. Aqui, reservaremos o nome *curva* para variedades 1-dimensionais, que serão definidas mais adiante; em particular, o conjunto imagem de uma trajetória suficientemente bem comportada é uma curva. Um outro termo também usado na literatura é *curva parametrizada*.

O conceito de derivada de uma trajetória vem de sua interpretação física. A derivada em um tempo $t \in I$ da trajetória é a 'taxa de variação instantânea' ou a *velocidade* do objeto que se move conforme a trajetória. Esse é o conceito usado na definição de derivada de funções reais e se estende facilmente para o caso de espaços normados.

:⊢ **Definição 19.32** (Derivabilidade)**.** Sejam $\boldsymbol{E}$ um espaço normado, $I \subseteq \mathbb{R}$ um intervalo aberto e $X \subseteq I$. Uma trajetória de I para E *derivável em* X (ou *derivável em* t para $X = \{t\}$) é uma trajetória $\gamma\colon I \longrightarrow E$ tal que, para todo $t \in X$, algum $v \in E$ satisfaz

$$v = \lim_{\delta \to 0} \frac{\gamma(t+\delta) - \gamma(t)}{\delta}.$$

Nesse caso, a *derivada de* γ *em* t é o vetor

$$\dot{\gamma}(t) := \lim_{\delta \to 0} \frac{\gamma(t+\delta) - \gamma(t)}{\delta}.$$

Caso seja necessário explicitar a operação de derivação, denota-se

$$\dot{\partial}\gamma(t) := \dot{\gamma}(t);$$

caso seja necessário explicitar a variável da função γ, denota-se

$$\dot{\partial}_{t'}\gamma(t')|_{t'=t} := \dot{\partial}\gamma(t) = \dot{\gamma}(t).$$

Uma trajetória *derivável* de I para E é uma trajetória $\gamma\colon I \longrightarrow E$ derivável em I. Nesse caso, a *derivada* de f é a trajetória

$$\begin{aligned} \dot{\partial}\gamma = \dot{\gamma}\colon I &\longrightarrow E \\ t &\longmapsto \dot{\gamma}(t). \end{aligned}$$

Mais geralmente, pode-se definir derivabilidade inferior e superior de uma trajetória em um ponto através de limites superior e inferior, já que o domínio de uma trajetória é um subconjunto de $\mathbb{R}$. Não abordaremos essas definições mais gerais aqui, mas elas podem ser conferidas em [Car71].

⊢ **Proposição 19.48** (Derivabilidade implica continuidade)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $I \subseteq \mathbb{R}$ *um intervalo aberto,* $X \subseteq I$ *e* $\gamma\colon I \longrightarrow E$ *uma trajetória derivável em* X*. Então* γ *é contínua em* X*.*

□ *Demonstração.* Basta notar que, para todo $t \in X$,

$$\lim_{\delta\to 0}(\gamma(t+\delta)-\gamma(t)) = \lim_{\delta\to 0}\delta\frac{\gamma(t+\delta)-\gamma(t)}{\delta} = 0\dot{\gamma}(t) = 0,$$

portanto $\lim_{\delta\to 0}\gamma(t+\delta) = \gamma(t)$, o que mostra que γ é contínua em t. ■

19.7.2.2 Operaões entre trajetórias deriváveis

⊢ **Proposição 19.49** (Propriedades de derivação)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $I \subseteq \mathbb{R}$ *um intervalo aberto e* $X \subseteq I$*.*

1. *Para todas trajetórias* $\gamma, \gamma'\colon I \longrightarrow E$ *deriváveis em* X*, a trajetória* $\gamma+\gamma'\colon I \longrightarrow E$ *é derivável em* X *e, para todo* $t \in X$*,*

$$\dot{\partial}(\gamma+\gamma')(t) = \dot{\partial}\gamma(t) + \dot{\partial}\gamma'(t);$$

2. *Para todo* $c \in \mathbb{R}$ *e toda trajetória* $\gamma\colon I \longrightarrow E$ *derivável em* X*, a trajetória* $c\gamma\colon I \longrightarrow E$ *é derivável em* X *e, para todo* $t \in X$*,*

$$\dot{\partial}(c\gamma)(t) = c\dot{\partial}\gamma(t);$$

3. *Se* $\boldsymbol{E}$ *é uma álgebra normada, para todas trajetórias* $\gamma, \gamma'\colon I \longrightarrow E$ *deriváveis em* X*, a trajetória* $\gamma\gamma'\colon I \longrightarrow E$ *é derivável em* X *e, para todo* $t \in X$*,*

$$\dot{\partial}(\gamma\gamma')(t) = \dot{\partial}\gamma(t)\gamma'(t) + \gamma(t)\dot{\partial}\gamma'(t).$$

□ *Demonstração.* 1. Para todo $t \in X$,

$$\begin{aligned}
\dot{\partial}(\gamma+\gamma')(t) &= \lim_{t\to 0}\frac{(\gamma+\gamma')(t+\delta)-(\gamma+\gamma')(t)}{\delta} \\
&= \lim_{\delta\to 0}\frac{\gamma(t+\delta)+\gamma'(t+\delta)-\gamma(t)-\gamma'(t)}{\delta} \\
&= \lim_{\delta\to 0}\frac{\gamma(t+\delta)-\gamma(t)}{t} + \lim_{\delta\to 0}\frac{\gamma'(t+\delta)-\gamma'(t)}{\delta} \\
&= \dot{\partial}\gamma(t) + \dot{\partial}\gamma'(t).
\end{aligned}$$

2. Para todo $t \in X$,

$$\begin{aligned}\dot{\partial}(c\gamma)(t) &= \lim_{\delta\to 0}\frac{(c\gamma)(t+\delta)-(c\gamma)(t)}{\delta}\\ &= \lim_{\delta\to 0}\frac{c\gamma(t+\delta)-c\gamma(t)}{\delta}\\ &= c\lim_{\delta\to 0}\frac{\gamma(t+\delta)-\gamma(t)}{\delta}\\ &= c\dot{\partial}\gamma(t).\end{aligned}$$

3. Para todo $t \in X$, da continuidade do produto da álgebra e da continuidade de γ' em t segue que

$$\begin{aligned}\dot{\partial}(\gamma\gamma')(t) &= \lim_{\delta\to 0}\frac{(\gamma\gamma')(t+\delta)-(\gamma\gamma')(t)}{\delta}\\ &= \lim_{\delta\to 0}\frac{\gamma(t+\delta)\gamma'(t+\delta)-\gamma(t)\gamma'(t)}{\delta}\\ &= \lim_{\delta\to 0}\frac{\gamma(t+\delta)\gamma'(t+\delta)-\gamma(t)\gamma'(t+\delta)+\gamma(t)\gamma'(t+\delta)-\gamma(t)\gamma'(t)}{\delta}\\ &= \lim_{\delta\to 0}\frac{\gamma(t+\delta)\gamma'(t+\delta)-\gamma(t)\gamma'(t+\delta)}{\delta}\\ &+ \lim_{\delta\to 0}\frac{\gamma(t)\gamma'(t+\delta)-\gamma(t)\gamma'(t)}{\delta}\\ &= \left(\lim_{\delta\to 0}\frac{\gamma(t+\delta)-\gamma(t)}{\delta}\right)\left(\lim_{\delta\to 0}\gamma'(t+\delta)\right)\\ &+ \lim_{\delta\to 0}\gamma(t)\left(\lim_{\delta\to 0}\frac{\gamma'(t+\delta)-\gamma'(t)}{\delta}\right)\\ &= \dot{\partial}\gamma(t)\gamma'(t)+\gamma(t)\dot{\partial}\gamma'(t).\end{aligned}$$

■

▷ **Exercício 19.11.** *Se* $p\colon \mathbb{R} \longrightarrow \mathbb{R}$ *é uma função polinomial dada por* $p(x) = +_{i=0}^{n} c_i x^i$, *então* $\dot{p}(x) = +_{i=0}^{n-1} c_i i x^{i-1}$.

⊢ **Proposição 19.50.** *Sejam* $\boldsymbol{A}$ *uma álgebra unitária normada completa e* $x \in A$.

$$\dot{\partial}\mathrm{e}^{tx} = x\mathrm{e}^{tx} = \mathrm{e}^{tx}x.$$

□ *Demonstração.* Como para todos $t, \delta \in \mathbb{R}$ vale

$$\mathrm{e}^{(t+\delta)x} - \mathrm{e}^{tx} = \mathrm{e}^{tx}\mathrm{e}^{\delta x} - \mathrm{e}^{tx} = \mathrm{e}^{tx}(\mathrm{e}^{\delta x} - 1_A) = \mathrm{e}^{tx}\left(\mathop{+}_{n=1}^{\infty}\frac{\delta^n x^n}{n!}\right),$$

segue que

$$\frac{\mathrm{e}^{(t+\delta)x} - \mathrm{e}^{tx}}{\delta} = \mathrm{e}^{tx}\left(x + \mathop{+}_{n=2}^{\infty}\frac{\delta^{n-1}x^n}{n!}\right),$$

o que implica que

$$\dot{\partial}\mathrm{e}^{tx} = \lim_{\delta\to 0} \frac{\mathrm{e}^{(t+\delta)x} - \mathrm{e}^{tx}}{\delta} = \mathrm{e}^{tx}x.$$

Como x comuta com todo elemento da forma $\frac{\delta^n x^n}{n!}$ da série de e^{tx}, segue que $\mathrm{e}^{tx}x = x\mathrm{e}^{tx}$. ■

19.7.2.3 Diferenciabilidade e derivabilidade

A proposição a seguir mostra que os conceitos de diferenciabilidade e derivabilidade coincidem para trajetórias. Mais à frente, definiremos as derivadas direcionais de uma função entre espaços normados, que estendem o conceito de derivação de trajetórias para essas funções. Nesse caso, nem toda função derivável em todas as direções será diferenciável, mas a recíproca será verdadeira.

⊢ **Proposição 19.51.** *Sejam $\boldsymbol{E}$ um espaço normado, $I \subseteq \mathbb{R}$ aberto e $X \subseteq I$. Uma função $\gamma\colon I \longrightarrow E$ é diferenciável em X se, e somente se, é derivável em X. Nesse caso, para todo $t \in X$ vale $\|\mathrm{D}\gamma|_t\| = \|\dot{\gamma}(t)\|$ e, para todo $\delta \in \mathbb{R}$,*

$$\mathrm{D}\gamma|_t(\delta) = \delta\dot{\gamma}(t).$$

□ *Demonstração.*

- (⇒) Suponhamos que γ é diferenciável e seja $t \in X$. Então

$$\begin{aligned}\lim_{\delta\to 0}\left\|\frac{\gamma(t+\delta)-\gamma(t)}{\delta} - \mathrm{D}\gamma|_t(1)\right\| &= \lim_{\delta\to 0}\left\|\frac{\gamma(t+\delta)-\gamma(t)-\delta\mathrm{D}\gamma|_t(1)}{\delta}\right\| \\ &= \lim_{\delta\to 0}\frac{\|\gamma(t+\delta)-\gamma(t)-\mathrm{D}\gamma|_t(\delta)\|}{|\delta|} \\ &= 0,\end{aligned}$$

  o que mostra que $\mathrm{D}\gamma|_t(1)$ é a derivada de γ em t.

- (⇐) Suponhamos que γ é derivável e seja $t \in X$. Defina

$$\begin{aligned} D\colon \mathbb{R} &\longrightarrow E \\ \delta &\longmapsto \delta\dot{\gamma}(t). \end{aligned}$$

  A função D é claramente linear e contínua. Da derivabilidade de γ em t segue que

$$\begin{aligned}\lim_{\delta\to 0}\frac{\|\gamma(t+\delta)-\gamma(t)-D(\delta)\|}{|\delta|} &= \lim_{\delta\to 0}\left\|\frac{\gamma(t+\delta)-\gamma(t)-\delta\dot{\gamma}(t)}{\delta}\right\| \\ &= \lim_{\delta\to 0}\left\|\frac{\gamma(t+\delta)-\gamma(t)}{\delta} - \dot{\gamma}(t)\right\| \\ &= 0,\end{aligned}$$

o que mostra que D é a diferencial de γ em t.

No caso em que γ é diferenciável, para todos $t \in X$ e $\delta \in \mathbb{R}$ vale

$$\mathrm{D}\gamma|_t(\delta) = \delta\dot{\gamma}(t)$$

e portanto

$$\|\mathrm{D}\gamma|_t\| = \sup_{\delta\in\mathbb{R}\smallsetminus\{0\}} \frac{\|\mathrm{D}\gamma|_t(\delta)\|}{|\delta|} = \sup_{\delta\in\mathbb{R}\smallsetminus\{0\}} \frac{\|\delta\dot{\gamma}(t)\|}{\|\delta\|} = \sup_{\delta\in\mathbb{R}\smallsetminus\{0\}} \frac{|\delta|\,\|\dot{\gamma}(t)\|}{\|\delta\|} = \|\dot{\gamma}(t)\|\,. \quad \blacksquare$$

Note que as proposições sobre derivabilidade implicar continuidade e propriedades de derivação demonstradas nesta seção seguem diretamente como corolário dessa caracterização da derivada de uma trajetória pela diferencial e das respectivas propriedades da diferencial, mas elas foram demonstradas independentemente para motivar o conceito de derivada.

Usando essa proposição, segue da regra da cadeia (19.42) as seguintes versões da regra da cadeia para trajetórias.

▷ **Exercício 19.12** (Regra da cadeia para trajetórias reparametrizadas)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $I, I' \subseteq \mathbb{R}$ *abertos,* $X \subseteq I$*,* $\alpha\colon I \longrightarrow \mathbb{R}$ *derivável em* X *tal que* $\alpha(I) \subseteq I'$ *e* $\gamma\colon I' \longrightarrow E$ *derivável em* $\alpha(X)$*. A trajetória composta* $\gamma \circ \alpha\colon I \longrightarrow E$ *é derivável em* X *e, para todo* $t \in X$*,*

$$\dot{\partial}(\gamma \circ \alpha)(t) = \dot{\alpha}(t)\dot{\gamma}(\alpha(t)).$$

▷ **Exercício 19.13** (Regra da cadeia para trajetórias empurradas)**.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$ *espaços normados,* $I \subseteq \mathbb{R}$ *e* $A \subseteq E$ *abertos,* $X \subseteq I$*,* $\gamma\colon I \longrightarrow \mathbb{R}$ *derivável em* X *tal que* $\gamma(I) \subseteq A$ *e* $f\colon A \longrightarrow E'$ *diferenciável em* $\gamma(X)$*. A trajetória composta* $f \circ \gamma\colon I \longrightarrow E$ *é derivável em* X *e, para todo* $t \in X$*,*

$$\dot{\partial}(f \circ \gamma)(t) = \mathrm{D}f|_{\gamma(t)}(\dot{\gamma}(t)).$$

19.7.3 Derivabilidade direcional

Antes de definirmos a derivada direcional, definiremos o fluxo em um espaço normado E por um vetor $v \in E$. Essa ferramenta a princípio parece trivial, mas o benefício de introduzir a ideia de um fluxo é que quando os conceitos de derivação forem generalizados para variedades o uso do fluxo, que não será trivial como neste caso, será essencial.

19.7.3.1 Fluxo constante

:⊢ **Definição 19.33.** Sejam $\boldsymbol{E}$ um espaço normado e $v \in E$. O *fluxo de* v em $\boldsymbol{E}$ é a função

$$\begin{aligned} \Phi^v \colon \mathbb{R} &\longrightarrow \mathscr{C}(\boldsymbol{E}, \boldsymbol{E}) \\ t &\longmapsto \Phi^{(t)v} \colon E \longrightarrow E \\ & \qquad\qquad x \longmapsto \Phi^{(t)v}(x) := x + tv. \end{aligned}$$

Para todo $x \in E$, a *trajetória* de Φ^v baseada em x é a trajetória

$$\begin{aligned} \Phi_x^{(\cdot)v} \colon \mathbb{R} &\longrightarrow E \\ t &\longmapsto \Phi^{(t)v}(x). \end{aligned}$$

O fluxo de v também pode ser visto como uma função

$$\begin{aligned} \Phi^v \colon \mathbb{R} \times E &\longrightarrow E \\ (t, x) &\longmapsto \Phi^{(t)v}(x) = x + tv. \end{aligned}$$

Esse ponto de vista será importante na generalização dos fluxos. Dado $x \in E$, denotaremos a derivada da trajetória $\Phi_x^{(\cdot)v}$ em $t \in \mathbb{R}$ por

$$\dot{\Phi}_x^{(t)v} := \dot{\partial}\Phi_x^{(t)v} = \lim_{\delta \to 0} \frac{\Phi_x^{(t+\delta)v} - \Phi_x^{(t)v}}{\delta}.$$

Ressaltamos na proposição seguinte as principais propriedades do fluxo.

⊢ **Proposição 19.52.** *Seja* $\boldsymbol{E}$ *um espaço normado.*

1. *Para todo* $v \in E$ *e todos* $t, t' \in \mathbb{R}$,
$$\Phi^{(t)t'v} = \Phi^{(tt')v};$$
2. *Para todo* $v \in E$,
$$\Phi^{(0)v} = \mathrm{I};$$
3. *Para todo* $v \in E$ *e todos* $t, t' \in \mathbb{R}$,
$$\Phi^{(t+t')v} = \Phi^{(t)v} \circ \Phi^{(t')v};$$
4. *Para todo* $v \in E$, *todo* $x \in E$ *e todo* $t \in \mathbb{R}$,
$$\dot{\Phi}_x^{(t)v} = v;$$

5. *Para todos $v, v' \in E$ e todo $t \in \mathbb{R}$,*

$$\Phi^{(t)(v+v')} = \Phi^{(t)v} \circ \Phi^{(t)v'}.$$

□ *Demonstração.* 1. Para todo $x \in E$,

$$\Phi^{(t)(t'v)}(x) = x + t(t'v) = x + (tt')v = \Phi^{(tt')v}(x).$$

2. Para todo $x \in E$,

$$\Phi^{(0)v}(x) = x + 0v = x = \mathrm{I}(x).$$

3. Para todo $x \in E$,

$$\Phi^{(t+t')v}(x) = x + (t+t')v = (x + tv) + t'v = \Phi^{(t)v} \circ \Phi^{(t')v}(x).$$

4. Para todo $x \in E$,

$$\begin{aligned}
\dot{\Phi}_x^{(t)v} &= \lim_{\delta \to 0} \frac{\Phi_x^{(t+\delta)v} - \Phi_x^{(t)v}}{\delta} \\
&= \lim_{\delta \to 0} \frac{x + (t+\delta)v - (x + tv)}{\delta} \\
&= \lim_{\delta \to 0} \frac{\delta v}{\delta} \\
&= v.
\end{aligned}$$

5. Para todo $x \in E$ e todo $t \in \mathbb{R}$,

$$\begin{aligned}
\Phi^{t(v+v')}(x) &= \Phi^{t(v'+v)}(x) \\
&= x + t(v' + v) \\
&= x + tv' + tv \\
&= \Phi^{tv}(x + tv') \\
&= \Phi^{tv}(\Phi^{tv'}(x)) \\
&= \Phi^{tv} \circ \Phi^{tv}(x),
\end{aligned}$$

o que mostra que $\Phi^{t(v+v')} = \Phi^{tv} \circ \Phi^{tv}$. ■

A primeira proposição esclarece uma possível ambiguidade de notação e serve como justificativa para a escolha de notação de fluxo. Como para todo $v \in E$ e todos $t, t' \in \mathbb{R}$ vale

$$\Phi^{(t)(t'v)} = \Phi^{(tt')v},$$

então a notação $\Phi^{tt'v}$ não é ambígua e podemos em geral denotar

$$\Phi^{tv} := \Phi^{(t)v}.$$

A segunda e a terceira proposições são de fato a propriedade de um fluxo. Elas significam que Φ^v é um homomorfismo de grupos de $\mathbb{R}$ para $\mathscr{C}(\boldsymbol{E}, \boldsymbol{E})$, o que mostra que na verdade temos

$$\begin{aligned} \Phi^v \colon \mathbb{R} &\longrightarrow \overset{\leftrightarrow}{\mathscr{C}}(\boldsymbol{E}) \\ t &\longmapsto \Phi^{tv} \end{aligned}$$

Consideremos espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$, $A \subseteq E$ um aberto, $v \in E$ e $x \in A$. Como A é aberto, para algum $r \in \left]0, \infty\right[$ vale que $\mathrm{B}\,(x; r) \subseteq A$, o que implica que, para todo $t \in \left]-r, r\right[$, $x + tv \in \mathrm{B}\,(x; r)$ se, e somente $\|(x + tv) - x\| < r$; como

$$|t|\,\|v\| = \|tv\| = \|(x + tv) - x\| < r,$$

isso equivale a $t \in \left]-r\,\|v\|^{-1}, r\,\|v\|^{-1}\right[$. Tomando $\varepsilon := r\,\|v\|^{-1}$, para toda função $f \colon A \longrightarrow E'$ a função

$$\begin{aligned} f \circ \Phi^{(\cdot)v}(x) \colon \left]-\varepsilon, \varepsilon\right[&\longrightarrow E' \\ t &\longmapsto f \circ \Phi^{tv}(x) = f(x + tv) \end{aligned}$$

está bem definida.

19.7.3.2 Derivada direcional

:⊢ **Definição 19.34** (Derivabilidade direcional)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $v \in E$ e $X \subseteq A$. Uma função de A para E' *derivável em X na direção v* (ou *derivável em x na direção v* para $X = \{x\}$) é uma função $f \colon A \longrightarrow E'$ tal que, para todo $x \in X$, a trajetória

$$\begin{aligned} f \circ \Phi_x^{(\cdot)v} \colon \left]-\varepsilon, \varepsilon\right[&\longrightarrow E' \\ t &\longmapsto f \circ \Phi^{tv}(x) = f(x + tv) \end{aligned}$$

é derivável em 0.

Nesse caso, para todo $x \in X$, a *derivada direcional de f em x na direção v* é

$$\partial_v f(x) := \dot{\partial}_t (f \circ \Phi^{(t)v}(x))|_{t=0} = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Listamos a seguir algumas das propriedades da derivada direcional.

⊢ **Proposição 19.53.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $v \in E$ e $x \in A$.*

1. *Para todas funções $f, f' \colon A \longrightarrow E'$ deriváveis em x na direção v, a função $f + f' \colon A \longrightarrow E'$ é derivável em x na direção v e*

$$\partial_v (f + f')(x) = \partial_v f(x) + \partial_v f'(x);$$

2. *Para todo $c \in \mathbb{R}$ e toda função $f\colon A \longrightarrow E'$ derivável em x na direção v, a função $cf\colon A \longrightarrow E'$ é derivável em x na direção v e*
$$\partial_v(cf)(x) = c\partial_v f(x);$$
3. *Para todas funções $f, f'\colon A \longrightarrow \mathbb{R}$ deriváveis em x na direção v, a função $ff'\colon A \longrightarrow \mathbb{R}$ é derivável em x na direção v e*
$$\partial_v(ff')(x) = \partial_v f(x)f'(x) + f(x)\partial_v f'(x);$$
4. *Para todo $c \in \mathbb{R}$ e toda função $f\colon A \longrightarrow E'$ derivável em x na direção v, a função $f\colon A \longrightarrow E'$ é derivável em x na direção cv e*
$$\partial_{cv} f(x) = c\partial_v f(x).$$

□ *Demonstração.* 1. Segue direto de 19.49.
2. Segue direto de 19.49.
3. Segue direto de 19.49.
4. Se $c = 0$, então
$$\partial_{0v} f(x) = \lim_{t\to 0} \frac{f(x + t0v) - f(x)}{t} = \lim_{t\to 0} \frac{f(x) - f(x)}{t} = 0 = 0\partial_v f(x).$$
Se $c \neq 0$, como $\lim_{t\to 0} ct = 0$, tomando $t' := ct$ segue que
$$\begin{aligned}\partial_{cv} f(x) &= \lim_{t\to 0} \frac{f(x + tcv) - f(x)}{t}\\ &= \lim_{t'\to 0} \frac{f(x + t'v) - f(x)}{(t' \diagup c)}\\ &= c \lim_{t'\to 0} \frac{f(x + t'v) - f(x)}{t'}\\ &= c\partial_v f(x).\end{aligned}$$
■

⊢ **Proposição 19.54.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E'}$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função diferenciável em X. Para todo $v \in E$, f é derivável em X na direção v e, para todo $x \in X$,*
$$\partial_v f(x) = \mathrm{D}f|_x(v).$$

□ *Demonstração.* Seja $x \in X$. Da diferenciabilidade de f em x segue que, como $\lim_{t\to 0} tv = 0$,
$$\begin{aligned}\lim_{t\to 0} \frac{f(x + tv) - f(x)}{t} &= \lim_{t\to 0} \frac{f(x + tv) - f(x)}{t} - \mathrm{D}f|_x(v) + \mathrm{D}f|_x(v)\\ &= \lim_{t\to 0} \frac{f(x + tv) - f(x) - \mathrm{D}f|_x(tv)}{t} + \mathrm{D}f|_x(v)\\ &= \mathrm{D}f|_x(v),\end{aligned}$$
o que mostra que f é derivável em x na direção v e $\partial_v f(x) = \mathrm{D}f|_x(v)$. ■

Em particular, no caso de $E = \mathbb{R}^d$, cuja base canônica é $(\mathrm{i}_k)_{k\in[d]}$ e cuja base canônica dual é $(\mathrm{i}_k^*)_{k\in[d]}$, vale

$$\mathrm{D}f|_x(v) = \mathrm{D}f|_x\left(\mathop{+}_{k\in[d]} v_k\mathrm{i}_k\right) = \mathop{+}_{k\in[d]} v_k\mathrm{D}f|_x(\mathrm{i}_k) = \mathop{+}_{k\in[d]} v_k\partial_{\mathrm{i}_k}f(x) = \mathop{+}_{k\in[d]} \partial_{\mathrm{i}_k}f(x)\mathrm{i}_k^*(v),$$

portanto

$$\mathrm{D}f|_x = \mathop{+}_{k\in[d]} \partial_{\mathrm{i}_k}f(x)\mathrm{i}_k^*.$$

19.7.4 Extremos locais e pontos críticos

19.7.4.1 Campos escalares e pontos extremos locais

Definimos aqui uma nomenclatura para um tipo específico de função entre espaços normados, aquelas que tem como contradomínio o corpo dos números reais. Elas são chamadas de campos escalares.

:⊢ **Definição 19.35.** Sejam $\boldsymbol{E}$ um espaço normado e $X \subseteq E$ um conjunto. Um *campo escalar* em X é uma função $f\colon X \longrightarrow \mathbb{R}$.

Os campos escalares são objetos que formalizam a ideia de função densidade e de campo de energia potencial, entre outros conceitos físicos relevantes. Mais à frente, definiremos outros objetos que recebem também o nome de 'campos', os 'campos tensoriais'; os campos escalares aqui definidos são casos especiais desses campos.

No contexto físico, uma propriedade importante dessas funções é se elas têm pontos de máximo e mínimo: pontos em que elas atingem valores reais máximos ou mínimos da sua imagem.

:⊢ **Definição 19.36** (Ponto extremo)**.** Sejam $\boldsymbol{E}$ um espaço normado, $X \subseteq E$ um conjunto e $f\colon X \longrightarrow \mathbb{R}$ um campo escalar.

- Um *ponto de máximo* de f é um ponto $x \in X$ tal que, para todo $x' \in X$,
$$f(x') \leq f(x).$$
Um *ponto de máximo estrito* de f é um ponto $x \in X$ tal que, para todo $x' \in X \smallsetminus \{x\}$,
$$f(x') < f(x).$$
- Um *ponto de mínimo* de f é um ponto $x \in X$ tal que, para todo $x' \in X$,
$$f(x) \leq f(x').$$
Um *ponto de mínimo estrito* de f é um ponto $x \in X$ tal que, para todo $x' \in X \smallsetminus \{x\}$,
$$f(x) < f(x').$$

Um *ponto extremo* de f é um ponto de máximo ou de mínimo de f e um *ponto extremo estrito* de f é um ponto é um ponto de máximo estrito ou de mínimo estrito de f.

Mais geralmente, podem-se considerar pontos que têm localmente essa propriedades, como definidos a seguir.

:⊢ **Definição 19.37** (Ponto extremo local)**.** Sejam $\boldsymbol{E}$ um espaço normado, $X \subseteq E$ um conjunto e $f\colon X \longrightarrow \mathbb{R}$ um campo escalar.

- Um *ponto de máximo local (estrito)* de f é um ponto $x \in X$ tal que, para alguma vizinhança $V \subseteq E$ de x, x é ponto de máximo (estrito) de $f|_{V\cap X}$.
- Um *ponto de mínimo local (estrito)* de f é um ponto $x \in X$ tal que, para alguma vizinhança $V \subseteq E$ de x, x é ponto de mínimo (estrito) de $f|_{V\cap X}$.

Um *ponto extremo local (estrito)* de f é um ponto é um ponto de máximo local (estrito) ou de mínimo local (estrito) de f.

Quando $K \subseteq E$ é compacto, uma função contínua $f\colon K \longrightarrow \mathbb{R}$ tem pontos de máximo e mínimo: como f é contínua, sua imagem é compacta, portanto fechada e limitada, o que implica que o supremo e o ínfimo da imagem pertencem à imagem. Esse resultado, que depende somente da continuidade da função e das propriedades de completude de $\mathbb{R}$, não nos permite achar os pontos de máximo e mínimo, somente afirmar sua existência. Usando diferenciação, podem-se estudar algumas características desses pontos que facilitam encontrá-los localmente.

Para a proposição seguinte, usaremos o conceito de conjunto convexo. Lembremos que, dados $x, x' \in E$, o 1-simplexo (ou intervalo) de vértices x, x' é o conjunto

$$[x, x']_{\Delta^1} = \{(1-t)x + tx' \mid t \in [0,1]\}$$

e o interior desse simplexo é

$$]x, x'[_{\Delta^1} = \{(1-t)x + tx' \mid t \in]0,1[\}\,.$$

Ainda, um conjunto convexo em E é um conjunto $X \subseteq E$ que, para todos $x, x' \in X$, satisfaz que $[x, x']_{\Delta^1} \subseteq X$.

⊢ **Proposição 19.55.** *Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ aberto, $X \subseteq A$ convexo, $x \in X$ e $f\colon A \longrightarrow \mathbb{R}$ um campo escalar diferenciável em x.*

- *Se x é ponto de máximo local de $f|_X\colon X \longrightarrow \mathbb{R}$, então, para todo $x' \in X$,*

$$\mathrm{D}f|_x(x') \leq \mathrm{D}f|_x(x);$$

- *Se x é ponto de mínimo local de $f|_X \colon X \longrightarrow \mathbb{R}$, então, para todo $x' \in X$,*

$$\mathrm{D}f|_x(x) \leq \mathrm{D}f|_x(x').$$

□ *Demonstração.* O caso em que x é ponto de mínimo local de $f|_X$ segue do caso em que x ponto de máximo local de $-f|_X$, portanto demonstraremos só o caso em que x é máximo local.

Sejam x um ponto de máximo local, $V \subseteq E$ uma vizinhança convexa de x tal que x é ponto de máximo em $V \cap X$, $t \in \,]0,1[$ e $x' \in V \cap X$. O teorema é evidente para $x' = x$, portanto consideramos $x' \neq x$. Como $V \cap X$ é convexo, vale que $x + t(x' - x) = (1-t)x + tx' \in V \cap X$; como f é diferenciável em x, definimos o resto

$$r(v) := f(x+v) - f(x) - \mathrm{D}f|_x(v)$$

e vale que $\lim_{v \to 0} \frac{r(v)}{\|v\|} = 0$ e

$$\mathrm{D}f|_x(v) + r(v) = f(x+v) - f(x).$$

Para $v = t(x' - x)$,

$$\begin{aligned} r(t(x'-x)) &= f(x + t(x'-x)) - f(x) - \mathrm{D}f|_x(t(x'-x)) \\ &= f(x + t(x'-x)) - f(x) - t\mathrm{D}f|_x(x'-x), \end{aligned}$$

portanto, como x é ponto de máximo local

$$\mathrm{D}f|_x(x'-x) + \frac{r(t(x'-x))}{t} = \frac{f(x+t(x'-x)) - f(x)}{t} \leq 0.$$

Assim segue que

$$\begin{aligned} \mathrm{D}f|_x(x'-x) &= \mathrm{D}f|_x(x'-x) + \|x'-x\| \lim_{t \to 0} \frac{|t|}{t} \frac{r(t(x'-x))}{\|t(x'-x)\|} \\ &= \mathrm{D}f|_x(x'-x) + \lim_{t \to 0} \|x'-x\| \frac{r(t(x'-x))}{t\,\|x'-x\|} \\ &= \lim_{t \to 0} \left(\mathrm{D}f|_x(x'-x) + \frac{r(t(x'-x))}{t} \right) \\ &\leq 0 \end{aligned}$$

e da linearidade de $\mathrm{D}f|_x$ resulta $\mathrm{D}f|_x(x') \leq \mathrm{D}f|_x(x)$. ■

Vale ressaltar que a desigualdade $\mathrm{D}f|_x(x') \leq \mathrm{D}f|_x(x)$ da proposição é equivalente à desigualdade

$$\mathrm{D}f|_x(x'-x) \leq 0,$$

que foi a obtida durante a demonstração. Assim, vale também o seguinte resultado.

⊢ **Proposição 19.56.** *Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ aberto, $X \subseteq A$ convexo, $x \in X$ e $f\colon A \longrightarrow \mathbb{R}$ um campo escalar diferenciável em x.*

- *Se x é ponto de máximo local de $f|_X\colon X \longrightarrow \mathbb{R}$, então, para todo $v \in E$ tal que $x + v \in X$,*
$$\mathrm{D}f|_x(v) \leq 0;$$
- *Se x é ponto de mínimo local de $f|_X\colon X \longrightarrow \mathbb{R}$, então, para todo $v \in E$ tal que $x + v \in X$,*
$$\mathrm{D}f|_x(v) \geq 0.$$

Para a proposição a seguir, lembremos que um subespaço afim de $S \subseteq \boldsymbol{E}$ pode ser escrito como $S = s + V$ para algum $s \in S$ e $V \subseteq E$ é um subespaço linear, e escrevemos $V = \vec{S}$.

⊢ **Proposição 19.57.** *Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ aberto, $S \subseteq E$ um subespaço afim, $x \in S \cap A$ e $f\colon A \longrightarrow \mathbb{R}$ um campo escalar diferenciável em x. Se x é um ponto de extremo local de $f|_S$, então, para todo $v \in \vec{S}$,*

$$\mathrm{D}f|_x(v) = 0.$$

□ *Demonstração.* O caso em que x é ponto de mínimo local de $f|_X$ segue do caso em que x ponto de máximo local de $-f|_X$, portanto demonstraremos só o caso em que x é máximo local.

Seja $v \in \vec{S}$ e suponhamos que $\mathrm{D}f|_x(v) \neq 0$. Então $\mathrm{D}f|_x(v) > 0$ ou $\mathrm{D}f|_x(v) < 0$. No primeiro caso, defina $v' := -v$ e, no segundo caso, $v' := v$, de modo que em ambos os casos $\mathrm{D}f|_x(v') < 0$. Seja $r \in \,]0, \infty[$ tal que $\mathrm{B}\,(x; r) \subseteq A$ e $X := \mathrm{B}\,(x; r) \cap S$. Como $\mathrm{B}\,(x; r)$ e S são convexos, então X é convexo. Disso segue que, para algum $t \in \,]0, \infty[$, vale $x - tv' \in X$ e então segue de 19.55 que

$$-t\mathrm{D}f|_x(v') = \mathrm{D}f|_x((x - tv') - x) \leq 0,$$

portanto $\mathrm{D}f|_x(v') \geq 0$, o que contradiz $\mathrm{D}f|_x(v') < 0$. Assim concluímos que $\mathrm{D}f|_x(v) = 0$. ■

19.7.4.2 Pontos críticos

Definimos aqui o conceito de ponto crítico de campos escalares. Esse conceito será posteriormente generalizado para quaisquer funções entre espaços normados, mas no momento usaremos esse caso específico.

:⊢ **Definição 19.38.** Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ aberto e $f\colon A \longrightarrow \mathbb{R}$ um campo escalar. Um *ponto crítico* de f é um ponto $x \in A$ tal que f é diferenciável em x e $\mathrm{D}f|_x = 0$.

⊢ **Proposição 19.58.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *aberto,* $x \in A$ *e* $f\colon A \longrightarrow \mathbb{R}$ *um campo escalar diferenciável em* x*. Se* x *é um ponto de extremo local de* f*, então* x *é um ponto crítico de* f*.*

□ *Demonstração.* Segue direto de 19.57 para $S = E$. ■

No entanto, nem todo ponto crítico é um ponto de extremo local. Consideremos por exemplo a função

$$\begin{aligned} f\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^3 \end{aligned}$$

que é diferenciável e cuja derivada é

$$\begin{aligned} \dot{\partial} f\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto 3x^2. \end{aligned}$$

O ponto $0 \in \mathbb{R}$ é um ponto crítico, pois $\dot{\partial} f(0) = 3 \times 0^2 = 0$, mas não é um ponto de extremo local pois, para todo $x > 0$, vale $f(x) > 0$ e, para todo $x < 0$, vale $f(x) < 0$. Esse ponto é o que se chama de ponto de sela, que definiremos mais adiante.

⊢ **Proposição 19.59.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $K \subseteq E$ *compacto tal que* $K^\circ \neq \emptyset$ *e* $f\colon K \longrightarrow \mathbb{R}$ *um campo escalar contínuo que é diferenciável em* K°*. Se* f *é constante em* ∂K*, então* f *tem ponto crítico em* K°*.*

□ *Demonstração.* Seja $c \in \mathbb{R}$ o valor de f em ∂K. Se f é constante, então para todo $x \in K^\circ$ vale $\mathrm{D}f|_x = 0$, logo todo ponto de K° é crítico. Se f não é constante, seja $x \in K$ tal que $f(x) \neq c$. Como $K \smallsetminus \partial K = K^\circ$, isso implica que $x \in K^\circ$. Como K é compacto e f é contínua, f tem ponto de mínimo e de máximo, que em particular são pontos de mínimo e máximo locais. Se $f(x) > c$, seja $m \in K$ o ponto de máximo de f. Nesse caso $m \in K^\circ$, pois $f(m) \geq f(x) > c$, e como f é diferenciável em K° segue de 19.58 que m é ponto crítico de f em K°. Se $f(x) < c$, seja $m \in K$ o ponto de mínimo de f. Nesse caso $m \in K^\circ$, pois $f(m) \leq f(x) < c$, e como f é diferenciável em K° segue de 19.58 que m é ponto crítico de f em K°. ■

19.7.5 Valor médio

19.7.5.1 Valor médio para campos escalares

A proposição a seguir é uma generalização direta do teorema do valor médio para funções de um intervalo compacto real para o corpo dos números reais.

⊢ **Proposição 19.60** (Valor médio em campos escalares)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $C \subseteq E$ *um conjunto,* $f\colon C \longrightarrow \mathbb{R}$ *um campo escalar e* $x, x' \in C$ *tais que* $x \neq x'$, $[x, x']_{\triangle^1} \subseteq C$ *e* f *é contínuo em* $[x, x']_{\triangle^1}$ *e diferenciável em* $]x, x'[_{\triangle^1}$. *Então, para algum* $m \in]x, x'[_{\triangle^1}$,

$$f(x') - f(x) = \mathrm{D}f|_m(x' - x).$$

□ *Demonstração.* A trajetória

$$\begin{aligned} \gamma\colon [0,1] &\longrightarrow [x, x']_{\triangle^1} \\ t &\longmapsto x + t(x' - x). \end{aligned}$$

é contínua em $[0, 1]$, diferenciável em $]0, 1[$, para todo $t \in]0, 1[$, $\dot{\gamma}(t) = x' - x$.

Definamos o campo escalar

$$\begin{aligned} f'\colon [0,1] &\longrightarrow \mathbb{R} \\ t &\longmapsto f \circ \gamma(t) - (f(x') - f(x))t. \end{aligned}$$

Esse campo f' é contínuo porque é a soma de $f \circ \gamma$, que é composição de funções contínuas, com um múltiplo da identidade, que também é contínuo. Analogamente, o campo escalar f' é diferenciável em $]0, 1[$, pois γ é diferenciável em $]0, 1[$ e, para todo $t \in]0, 1[$, $\gamma(t) \in]x, x'[_{\triangle^1}$, logo f é diferenciável em $\gamma(t)$, o que implica que a composta $f \circ \gamma$ é diferenciável em $]0, 1[$. Segue da regra da cadeia que

$$\begin{aligned} \dot{\partial} f'(t) &= \mathrm{D}f'|_t(1) \\ &= \mathrm{D}(f \circ \gamma)|_t(1) - (f(x') - f(x)) \\ &= \mathrm{D}f|_{\gamma(t)}(\dot{\partial}\gamma(t)) - (f(x') - f(x)) \\ &= \mathrm{D}f|_{\gamma(t)}(x' - x) - (f(x') - f(x)). \end{aligned}$$

Além disso, como $\gamma(0) = x$ e $\gamma(1) = x'$,

$$f'(1) = f(x') - (f(x') - f(x))1 = f(x) - (f(x') - f(x))0 = f'(0),$$

o que mostra que f' é constante em $\partial[0, 1] = \{0, 1\}$. Como $[0, 1]$ é compacto, segue de 19.59 que f' tem um ponto crítico $t_m \in]0, 1[$; ou seja, esse ponto satisfaz

$$\dot{\partial} f'(t_m) = \mathrm{D}f'|_{t_m}(1) = 0.$$

Segue que

$$0 = \dot{\partial} f'(t_m) = \mathrm{D}f|_{\gamma(t_m)}(x' - x) - (f(x') - f(x))$$

e, definindo $m := \gamma(t_m) \in]x, x'[_{\triangle^1}$, conclui-se que

$$f(x') - f(x) = \mathrm{D}f|_m(x' - x).$$ ■

19.7.5.2 Majoração de trajetória

A proposição 19.60 não pode ser generalizada para qualquer função entre espaços normados. Em especial, nem mesmo para trajetórias ela é válida. Consideremos por exemplo a trajetória

$$\begin{aligned} f\colon t &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (\cos(t), \operatorname{sen}(t)) \end{aligned}$$

em $\mathbb{R}^2$, que é diferenciável. Sua derivada é

$$\begin{aligned} \dot{f}\colon t &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (-\operatorname{sen}(t), \cos(t)). \end{aligned}$$

Vale[18] $f(0) = (1, 0) = f(\tau)$, logo $f(\tau) - f(0) = 0$, mas não existe $t \in \,]0, \tau[$ tal que $\dot{f}(t) = 0$. No entanto, podemos obter um tipo de desigualdade do valor médio, que descreveremos a seguir.

A proposição a seguir será a ferramenta para demonstrar várias proposições relacionadas ao teorema do valor médio.

$\vdash$ **Proposição 19.61** (Majoração de trajetória)**.** *Sejam $I = [a, a'] \subseteq \mathbb{R}$ um intervalo compacto, $\boldsymbol{E}$ um espaço normado e $\gamma\colon I \longrightarrow E$ e $f\colon I \longrightarrow \mathbb{R}$ trajetórias contínuas que são diferenciáveis em I°. Se, para todo $t \in I^\circ$, $\|\dot{\gamma}(t)\| \leq \dot{f}(t)$, então*

$$\|\gamma(a') - \gamma(a)\| \leq f(a') - f(a).$$

$\square$ *Demonstração.* Seja $\varepsilon \in \,]0, \infty[$ e I_ε o conjunto dos $x \in I = [a, a']$ tais que, para todos $x' \in [a, x]$, vale

$$\|\gamma(x') - \gamma(a)\| \leq f(x') - f(a) + \varepsilon(x' - a).$$

Primeiro notemos que $I_\varepsilon \neq \emptyset$ porque $a \in I_\varepsilon$. Agora, para todo $x \in I_\varepsilon$, se $x' \in [a, x]$ então por definição $x' \in I_\varepsilon$, portanto I_ε é um intervalo. Seja $s_\varepsilon := \sup I_\varepsilon$ o supremo desse intervalo. Mostremos que $s_\varepsilon \in I_\varepsilon$. Se $s_\varepsilon = a$, está provado, Caso contrário, seja $(x_n)_{n\in\infty}$ uma sequência em $]a, s_\varepsilon[$ tal que $\lim_{n\to\infty} x_n = s_\varepsilon$. Como γ e f são contínuas, segue que

$$\begin{aligned} \|\gamma(s_\varepsilon) - \gamma(a)\| &= \left\|\gamma(\lim_{x\to\infty} x_n) - \gamma(a)\right\| \\ &= \lim_{x\to\infty} \|\gamma(x_n) - \gamma(a)\| \\ &\leq \lim_{x\to\infty} f(x_n) - f(a) + \varepsilon(x' - a) \\ &\leq f(\lim_{x\to\infty} x_n) - f(a) + \varepsilon(x' - a) \\ &\leq f(s_\varepsilon) - f(a) + \varepsilon(x' - a), \end{aligned}$$

[18] A constante $\tau \in \mathbb{R}$ aqui é o período de cos e sen.

o que mostra que $s_\varepsilon \in I_\varepsilon$. Mostraremos agora que $s_\varepsilon = a'$. Suponhamos por absurdo que valesse $s_\varepsilon < a'$. Pela hipótese das derivadas, seguiria que, para algum $\delta \in]0, a' - s_\varepsilon[$, valeria para todo $x \in]s_\varepsilon, s_\varepsilon + \delta[$ que

$$\left\| \frac{\gamma(x) - \gamma(s_\varepsilon)}{x - s_\varepsilon} \right\| \leq \frac{f(x) - f(s_\varepsilon)}{x - s_\varepsilon} + \varepsilon,$$

portanto, como $\|x - c\| = x - c$, seguiria que

$$\|\gamma(x) - \gamma(s_\varepsilon)\| \leq f(x) - f(s_\varepsilon) + \varepsilon(x - s_\varepsilon).$$

Assim, a desigualdade triangular implicaria que

$$\begin{aligned} \|\gamma(x) - \gamma(a)\| &\leq \|\gamma(x) - \gamma(s_\varepsilon)\| + \|\gamma(s_\varepsilon) - \gamma(a)\| \\ &\leq f(x) - f(s_\varepsilon) + \varepsilon(x - s_\varepsilon) + f(s_\varepsilon) - f(a) + \varepsilon(s_\varepsilon - a) \\ &= f(x) - f(a) + \varepsilon(x - a), \end{aligned}$$

logo que $x \in I_\varepsilon$, o que contradiria s_ε ser supremo. Conclui-se então que $s_\varepsilon = a'$, logo que $I_\varepsilon = I$. Como isso vale para todo ε, segue que

$$\|\gamma(a') - \gamma(a)\| \leq f(a') - f(a). \qquad \blacksquare$$

Essa proposição tem algumas formulações um pouco mais gerais que não abordaremos aqui. Em vez da condição ser entre as derivadas de γ e f, poderíamos considerar as derivadas à esquerda ou à direita definidas com limites inferior ou superior, respectivamente. Além disso, poderíamos considerar que as trajetórias são deriváveis em um conjunto co-contável, ou seja, derivável em todo ponto a menos de um conjunto contável de pontos. Essas generalizações podem ser conferidas em [Car71].

A proposição 19.61 tem várias consequências importantes. A primeira é o caso (degenerado) em que $E = \{0\}$, que implica que f é crescente. A recíproca também é válida.

$\vdash$ **Proposição 19.62** (Derivada positiva equivale a crescente)**.** *Sejam $I \subseteq \mathbb{R}$ um intervalo e $f\colon I \longrightarrow \mathbb{R}$ uma função contínua que é diferenciável em I°. Então $\dot{f} \geq 0$ se, e somente se, f é crescente.*

$\square$ *Demonstração.*

- ($\Rightarrow$) Suponhamos que $\dot{f} \geq 0$. Para todos $x, x' \in I$, $[x, x']$ é um intervalo compacto. Definimos $E = \{0\}$ e $\gamma\colon [x, x'] \to E; t \mapsto 0$. Como é constante, a trajetória γ é contínua em $[x, x']$, diferenciável em $]x, x'[$ e, para todo $t \in I^\circ$, vale $\dot{\gamma}(t) = 0$. Portanto $\|\dot{\gamma}(t)\| = 0 \leq \dot{f}(t)$ e segue de 19.61 que

$$0 = \|\gamma(x') - \gamma(x)\| \leq f(x') - f(x),$$

o que mostra que $f(x) \leq f(x')$, logo que f é crescente.

- ($\Rightarrow$) Suponhamos que f é crescente e seja $t \in I^\circ$. Para todo $\delta \in [0, \infty[$, vale $f(t+\delta) \geq f(t)$, logo
$$0 \leq \frac{f(t+\delta) - f(t)}{\delta}.$$
Como f é diferenciável, segue que $\dot{f}(t) \geq 0$.

■

Uma outra consequência é que se a derivada de uma trajetória é limitada (por uma constante), então a trajetória é controlada (pela mesma constante).

⊢ **Proposição 19.63** (Controle diferencial de trajetória)**.** *Sejam $I \subseteq \mathbb{R}$ um intervalo, $\boldsymbol{E}$ um espaço normado e $\gamma\colon I \longrightarrow E$ uma trajetória contínua que é diferenciável em I°. Se $\dot{\gamma}$ é limitada por alguma constante $c \in \mathbb{R}_{\geq 0}$ (para todo $t \in I^\circ$, $\|\dot{\gamma}(t)\| \leq c$), então γ é controlada e sua distorção satisfaz $\langle\!\langle \gamma \rangle\!\rangle \leq c$; isto é, para todos $x, x' \in I$,*
$$\|\gamma(x') - \gamma(x)\| \leq c\,\|x' - x\|\,.$$

□ *Demonstração.* Sejam $x, x' \in I$. Se $x = x'$, a desigualdade segue trivialmente. Se $x \neq x'$, consideremos o caso $x < x'$ e o intervalo compacto $[x, x']$, a restrição $\gamma|_{[x,x']}$ e a função $f\colon [x, x'] \to \mathbb{R}; t \mapsto ct$. Como é um múltiplo da identidade, a função f é contínua em $[x, x']$, diferenciável em $]x, x'[$ e, para todo $t \in]x, x'[$, vale $\dot{f}(t) = c$. Portanto $\|\dot{\gamma}(t)\| \leq c = \dot{f}(t)$ e segue de 19.61 que
$$\|\gamma(x') - \gamma(x)\| \leq f(x') - f(x) = c(x' - x) = c\,\|x' - x\|\,.$$
Por fim, caso $x' < x$, o resultado segue trivialmente do caso anterior. ■

19.7.5.3 Desigualdade do valor médio e consequências

Consideramos agora o caso de uma função entre espaços normados quaisquer. A seguinte proposição é uma alternativa à proposição 19.60 e é conhecida como *desigualdade do valor médio*; algumas versões levemente diferentes também existem e recebem o mesmo nome.

⊢ **Proposição 19.64** (Desigualdade do valor médio)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $C \subseteq E$ um conjunto, $f\colon C \longrightarrow E'$ uma função e $x, x' \in C$ tais que $[x, x']_{\mathbb{1}} \subseteq C$, f é contínua em $[x, x']_{\mathbb{1}}$ e diferenciável em $]x, x'[_{\mathbb{1}}$. Então*
$$\|f(x') - f(x)\| \leq \sup_{m \in]x,x'[_{\mathbb{1}}} \|\mathrm{D}f|_m\|\,\|x' - x\|\,.$$

□ *Demonstração.* Se $\sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| = \infty$, a desigualdade de enunciado vale trivialmente. Suponhamos então que $\sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| < \infty$. A trajetória

$$\begin{aligned} \gamma\colon [0,\infty] &\longrightarrow [x,x']_{\mathbb{A}^1} \\ t &\longmapsto x + t(x'-x) \end{aligned}$$

é contínua em $[0,1]$, diferenciável em $]0,1[$ e, para todo $t \in]0,1[$, vale $\dot\gamma(t) = x'-x$. Disso segue que $f\circ\gamma$ é contínua em $[0,1]$, diferenciável em $]0,1[$ e, para todo $t\in]0,1[$, segue da regra da cadeia que

$$\dot\partial(f\circ\gamma)(t) = \mathrm{D}f_{\gamma(t)}(\dot\gamma(t)) = \mathrm{D}f|_{\gamma(t)}(x'-x).$$

Isso implica que

$$\left\|\dot\partial(f\circ\gamma)(t)\right\| \leq \sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| \, \|x'-x\|$$

e, como $\gamma(0)=x$ e $\gamma(1)=x'$, segue de 19.63 que

$$\begin{aligned} \|f(x')-f(x)\| &= \|f(\gamma(1))-f(\gamma(0))\| \\ &\leq \sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| \, \|x'-x\| \, (1-0) \\ &= \sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| \, \|x'-x\| \, . \end{aligned}$$ ■

Um corolário imediato é uma versão multidimensional de 19.63.

⊢ **Proposição 19.65** (Controle diferencial de função)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A\subseteq E$ um aberto convexo e $f\colon A\longrightarrow E'$ uma função diferenciável. Se, para algum $c\in[0,\infty[$, todo $x\in A$ satisfaz $\|\mathrm{D}f|_x\|\leq c$, então f é controlada e sua distorção satisfaz $\langle\!\langle f\rangle\!\rangle \leq c$; isto é, para todos $x,x'\in A$,*

$$\|f(x')-f(x)\| \leq c\,\|x'-x\|\,.$$

□ *Demonstração.* Sejam $x,x'\in A$. Como A é convexo, $[x,x']_{\mathbb{A}^1}\subseteq A$; como f é diferenciável, é contínua em $[x,x']_{\mathbb{A}^1}$ e diferenciável em $]x,x'[_{\mathbb{A}^1}$. Pela hipótese da norma da diferencial, segue de 19.64 que

$$\|f(x')-f(x)\| \leq \sup_{m\in]x,x'[_{\mathbb{A}^1}} \|\mathrm{D}f|_m\| \, \|x'-x\| \leq c\,\|x'-x\|\,.$$ ■

Um outro corolário importante da desigualdade do valor médio é o seguinte, que será usado adiante na demonstração de uma condição suficiente para diferenciabilidade de uma função em um ponto em termos de suas diferenciais parciais.

⊢ **Proposição 19.66.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto convexo, $f\colon A \longrightarrow E'$ uma função diferenciável e $T\colon E \longrightarrow E'$ um função linear contínua. Se, para algum $\varepsilon \in \left]0, \infty\right[$, todo $x \in A$ satisfaz $\|\mathrm{D}f|_x - T\| \leq \varepsilon$, então, para todos $x \in A$ e $v \in E$ tais que $x + v \in A$,*

$$\|f(x+v) - f(x) - T(v)\| \leq \varepsilon \|v\|.$$

□ *Demonstração.* Definamos $f' := f - T$. Então f' é diferenciável em A, pois é soma de funções diferenciáveis, e, para todo $x \in A$,

$$\mathrm{D}f'|_x = \mathrm{D}f|_x - T.$$

Portanto $\|\mathrm{D}f'|_x\| \leq \varepsilon$ e então segue de 19.65 que, para todos $x \in A$ e $v \in E$ tal que $x + v \in A$,

$$\begin{aligned}\|f(x+v) - f(x) - T(v)\| &= \|f(x+v) - f(x) - T(x+v) + T(x)\| \\ &= \|f'(x+v) - f'(x)\| \\ &= \varepsilon \|v\|.\end{aligned}$$

■

19.7.5.4 Diferencial nula

Por fim, uma consequência importante da desigualdade do valor médio é que uma função com diferencial zero é constante em cada componente conexa. Enunciaremos o caso em que f é definida em um conjunto conexo, pois a generalização para cada componente conexa é evidente. Essa proposição é um recíproca da propriedade de que se f é constante, sua diferencial é 0.

⊢ **Proposição 19.67.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto conexo não vazio e $f\colon A \longrightarrow E'$ uma função diferenciável. Se $\mathrm{D}f = 0$, então f é constante.*

□ *Demonstração.* Sejam $x \in A$ e $r \in \left]0, \infty\right[$ tal que $\mathrm{B}(x; r) \subseteq A$. Como $\mathrm{B}(x; r)$ é convexa e, para todo $x' \in \mathrm{B}(x; r)$, $\|\mathrm{D}f|_{x'}\| = 0$, segue de 19.65 que, para todos $x', x'' \in \mathrm{B}(x; r)$,

$$\|f(x'') - f(x')\| \leq 0 \|x'' - x'\| = 0,$$

portanto $f(x') = f(x'')$, o que mostra que f é constante em $\mathrm{B}(x; r)$, logo que f é localmente constante. Como A é um espaço topológico não vazio e conexo e E' é um espaço topológico separado, segue de 14.62 que f é constante. ■

19.7.6 Diferenciabilidade parcial

Nessa seção estudaremos a diferenciabilidade de funções entre produtos de espaços normados. Lembremos que dados $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ espaços normados, o produto $\boldsymbol{E}_0 \times \ldots \times \boldsymbol{E}_{n-1}$ é um espaço normado com qualquer uma das normas p, e essas normas são todas equivalentes. No que segue, usaremos qualquer uma dessas normas.

19.7.6.1 Contradomínio produto

Primeiro, estudamos o caso mais simples em que o contradomínio de uma função é um produto de espaços normados. Relembremos alguma notações relacionadas á projeção canônica e ressaltemos alguns resultados.

Sejam $\boldsymbol{E}'_0, \dots, \boldsymbol{E}'_{n-1}$ espaços normados e $k \in [n]$. A *projeção* de $E'_0 \times \cdots \times E'_{n-1}$ sobre E'_k é a função

$$\begin{aligned} \mathrm{p}_k \colon E'_0 \times \cdots \times E'_{n-1} &\longrightarrow E'_k \\ (x_0, \dots, v_{n-1}) &\longmapsto v_k. \end{aligned}$$

A projeção p_k é linear e contínua. Como p_k é linear, sua diferencial, para todo $x \in E'_0 \times \cdots \times E'_{n-1}$, é

$$\mathrm{D}\mathrm{p}_k|_x = \mathrm{p}_k.$$

Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto e $f \colon A \longrightarrow \boldsymbol{E}'_0 \times \ldots \times \boldsymbol{E}'_{n-1}$ uma função. Definimos a função coordenada $f_k := \mathrm{p}_k \circ f \colon A \longrightarrow E'_k$. A função f pode ser escrita como

$$\begin{aligned} f \colon A &\longrightarrow E'_0 \times \cdots \times E'_{n-1} \\ x &\longmapsto (f_0(x), \dots, f_{n-1}(x)). \end{aligned}$$

⊢ **Proposição 19.68.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'_0, \dots, \boldsymbol{E}'_{n-1}$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f \colon A \longrightarrow E'_0 \times \cdots \times E'_{n-1}$ uma função. Então f é diferenciável em X se, e somente se, para todo $k \in [n]$, $f_k \colon A \longrightarrow E'_k$ é diferenciável em X. Nesse caso, para todo $x \in X$,*

$$\mathrm{D}f|_x = (\mathrm{D}f_0|_x, \dots, \mathrm{D}f_{n-1}|_x) = \bigplus_{k \in [n]} \iota_k \circ \mathrm{D}f_k|_x.$$

□ *Demonstração.*

- (⇒) Suponhamos que f é diferenciável em X e sejam $k \in [n]$ e $x \in X$. Como f é diferenciável em x e p_k é diferenciável em $f(x)$, então $f_k := \mathrm{p}_k \circ f$ é diferenciável em x e

$$\mathrm{D}f_k|_x = \mathrm{D}(\mathrm{p}_k \circ f)|_x = \mathrm{D}\mathrm{p}_k|_{f(x)} \circ \mathrm{D}f|_x = \mathrm{p}_k \circ \mathrm{D}f|_x.$$

- (⇐) Suponhamos que, para todo $k \in [n]$, $f_k \colon A \longrightarrow E_k$ é diferenciável em X e seja $x \in X$. Defina

$$\begin{aligned} D \colon E &\longrightarrow E'_0 \times \cdots \times E'_{n-1} \\ v &\longmapsto (\mathrm{D}f_0|_x(v), \dots, \mathrm{D}f_{n-1}|_x(v)). \end{aligned}$$

Claramente D é linear e contínua. Da diferenciabilidade das f_k em x segue que

$$\begin{aligned}
&\lim_{v\to 0}\frac{f(x+v)-f(x)-D(v)}{\|v\|}\\
&\quad=\lim_{v\to 0}\frac{(f_k(x+v))_{k\in[n]}-(f_k(x))_{k\in[n]}-(\mathrm{D}f_k|_x(v))_{k\in[n]}}{\|v\|}\\
&\quad=\left(\lim_{v\to 0}\frac{f_k(x+v)-f_k(x)-\mathrm{D}f_k|_x(v)}{\|v\|}\right)_{k\in[n]}\\
&\quad=0
\end{aligned}$$

o que mostra que f é diferenciável em x e sua diferencial é

$$\mathrm{D}f|_x=(\mathrm{D}f_0|_x,\dots,\mathrm{D}f_{n-1}|_x). \qquad \blacksquare$$

19.7.6.2 Domínio produto

Para tratar da diferenciabilidade de uma função definida em um subconjunto de um produto de espaços normados, precisamos relembrar noções e notações de espaços e subespaços afins, e definir mais alguns conceitos: a noção de diferenciabilidade parcial e da diferencial parcial de uma função.

Sejam $\boldsymbol{E}_0,\dots,\boldsymbol{E}_{n-1}$ espaços normados e $k\in[n]$. A *inclusão* de E_k em $E_0\times\cdots\times E_{n-1}$ é a função

$$\begin{aligned}
\iota_k\colon E_k &\longrightarrow E_0\times\cdots\times E_k\times\cdots\times E_{n-1}\\
y &\longmapsto (0,\dots,0,\underbrace{y}_{k},0,\dots,0).
\end{aligned}$$

A inclusão ι_k é linear e contínua.

Sejam $x\in E_0\times\cdots\times E_{n-1}$ e $x_{\hat{k}}:=(x_0,\dots,x_{k-1},0,x_{k+1},\dots,x_{n-1})\in E_0\times\cdots\times E_{n-1}$. A *inclusão afim transladada por* $x_{\hat{k}}$ de E_k em $E_0\times\cdots\times E_{n-1}$ é a função $\iota_{k,x}:=x_{\hat{k}}+\iota_k$, dada explicitamente por

$$\begin{aligned}
\iota_{k,x}\colon E_k &\longrightarrow E_0\times\cdots\times E_{n-1}\\
y &\longmapsto x_{\hat{k}}+\iota_k(y)=(x_0,\dots,x_{k-1},y,x_{k+1},\dots,x_{n-1}).
\end{aligned}$$

A função $\iota_{k,x}$ é contínua. A imagem de $\iota_{k,x}$ é o subespaço afim de $E_0\times\cdots\times E_{n-1}$

$$\iota_{k,x}(E_k)=x_{\hat{k}}+\iota_k(E_k)=\{(x_0,\dots,x_{k-1},y,x_{k+1},\dots,x_{n-1})\mid y\in E_k\}$$

e $\iota_{k,x}(x_k)=x$. Como $\iota_{k,x}$ é afim, então é diferenciável e sua diferencial é a diferencial da função linear ι_k, ou seja

$$\mathrm{D}\iota_{k,x}|_y=\iota_k.$$

Seja $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto e $x \in A$. O conjunto

$$A_{k,x} := {\iota_{k,x}}^{-1}(A) = \{y \in E_k \mid (x_0, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_{n-1}) \in A\}$$

é aberto em E_k, pois A é aberto e $\iota_{k,x}$ é contínua. Além disso, $x_k \in A_{k,x}$.

Sejam $\boldsymbol{E}'$ um espaço normado e $f\colon A \longrightarrow E'$ uma função. Definimos $f_{k,x} := f \circ \iota_{k,x}$, dada explicitamente por

$$\begin{aligned} f_{k,x}\colon A_k &\longrightarrow E' \\ y &\longmapsto f(x_0, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_{n-1}). \end{aligned}$$

:⊢ **Definição 19.39** (Diferenciabilidade parcial (local))**.** Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto, $X \subseteq A$ e $k \in [n]$. Uma função de A para E' *k-parcialmente diferenciável em X* é uma função $f\colon A \longrightarrow E'$ tal que, para todo $x \in X$, $f_{k,x}\colon A_{k,x} \longrightarrow E'$ é diferenciável em x_k. A *diferencial k-parcial* de f em x é a diferencial

$$\mathrm{D}_k f|_x := \mathrm{D} f_{k,x}|_{x_k}\colon E_k \longrightarrow E'.$$

⊢ **Proposição 19.69.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função diferenciável em X. Para todo $k \in [n]$, f é k-parcialmente diferenciável em X e, para todo $x \in X$,*

$$\mathrm{D}_k f|_x = \mathrm{D} f|_x \circ \iota_k.$$

Para todo $v \in E_0 \times \cdots \times E_{n-1}$,

$$\mathrm{D} f|_x(v) = \mathop{\bigplus}_{k \in [n]} \mathrm{D}_k f|_x(v_k).$$

□ *Demonstração.* Sejam $k \in [n]$ $x \in X$. Como f é diferenciável em x e $\iota_{k,x}$ é diferenciável em x_k, segue da regra da cadeia que $f_{k,x} = f \circ \iota_{k,x}$ é diferenciável em x e que

$$\mathrm{D}_k f|_x = \mathrm{D} f_{k,x}|_{x_k} = \mathrm{D}(f \circ \iota_{k,x})|_x = \mathrm{D} f|_{\iota_{k,x}(x_k)} \circ \mathrm{D}\iota_{k,x}|_{x_k} = \mathrm{D} f|_x \circ \iota_k.$$

Para todo $v \in E_0 \times \cdots \times E_{n-1}$, segue da linearidade de $\mathrm{D} f|_x$ que

$$\begin{aligned} \mathrm{D} f|_x(v) &= \mathrm{D} f|_x(v_0, \ldots, v_{n-1}) \\ &= \mathop{\bigplus}_{k \in [n]} \mathrm{D} f|_x(0, \ldots, 0, v_k, 0, \ldots, 0) \\ &= \mathop{\bigplus}_{k \in [n]} \mathrm{D} f|_x \circ \iota_k(v_k) \\ &= \mathop{\bigplus}_{k \in [n]} \mathrm{D}_k f|_x(v_k). \end{aligned}$$

■

A recíproca, no entanto, não é verdadeira. Algumas funções são parcialmente diferenciáveis sobre todos E_k em um ponto x sem serem diferenciáveis em x. Para achar uma recíproca, devemos definir uma propriedade um pouco mais forte de diferenciabilidade parcial, que é chamada de diferenciabilidade contínua parcial.

:⊢ **Definição 19.40** ($\mathscr{C}^1$-Diferenciabilidade parcial (local))**.** Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto, $X \subseteq A$ e $k \in [n]$. Uma função de A para E' *k-parcialmente $\mathscr{C}^1$-diferenciável em X* (ou *k-parcialmente continuamente diferenciável em X*) é uma função $f\colon A \longrightarrow E'$ que, para todo $x \in X$, satisfaz seguinte: para alguma vizinhança $V \subseteq A$ de x, f é k-parcialmente diferenciável em V e

$$\begin{aligned} \mathrm{D}_k f\colon V &\longrightarrow \mathscr{L}(\boldsymbol{E}_k, \boldsymbol{E}') \\ y &\longmapsto \mathrm{D}f|_y \end{aligned}$$

é contínua em x.

Primeiro mostramos que vale uma versão de 19.69 para $\mathscr{C}^1$-diferenciabilidade.

⊢ **Proposição 19.70.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função $\mathscr{C}^1$-diferenciável em X. Para todo $k \in [n]$, f é k-parcialmente $\mathscr{C}^1$-diferenciável em X.*

□ *Demonstração.* Sejam $\boldsymbol{E} := \boldsymbol{E}_0 \times \ldots \times \boldsymbol{E}_{n-1}$, $k \in [n]$, $x \in X$ e $V \subseteq A$ uma vizinhança de x em que f é diferenciável e tal que $\mathrm{D}f$ é contínua em x. Por 19.69, para todo $x' \in V$ vale que f é k-parcialmente diferenciável em x' e $\mathrm{D}_k f|_{x'} = \mathrm{D}f|_{x'} \circ \iota_k$.

Seja $\varepsilon \in \,]0, \infty[$. Como $\mathrm{D}f$ é contínua em x, para algum $\delta \in \,]0, \infty[$ vale que, para todo $x' \in \mathrm{B}(x; \delta) \cap V$,

$$\|\mathrm{D}f|_{x'} - \mathrm{D}f|_x\|_{\mathscr{L}(E,E')} \leq \varepsilon.$$

Disso segue que, para todo $k \in [n]$, como $\|\iota_k\|_{\mathscr{L}(E_k,E)} = 1$ (em qualquer norma p no produto), então

$$\begin{aligned} \|\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x\|_{\mathscr{L}(E_k,E')} &= \|(\mathrm{D}f|_{x'} - \mathrm{D}f|_x) \circ \iota_k\|_{\mathscr{L}(E_k,E')} \\ &\leq \|\mathrm{D}f|_{x'} - \mathrm{D}f|_x\|_{\mathscr{L}(E,E')} \|\iota_k\|_{\mathscr{L}(E_k,E)} \\ &= \|\mathrm{D}f|_{x'} - \mathrm{D}f|_x\|_{\mathscr{L}(E,E')} \\ &\leq \varepsilon, \end{aligned}$$

o que mostra que $\mathrm{D}_k f$ é contínua em x. ■

$\vdash$ **Proposição 19.71.** *Sejam $\boldsymbol{E}_0, \dots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função tal que, para todo $k \in [n]$, é k-parcialmente $\mathscr{C}^1$-diferenciável em X. Então a função f é diferenciável em X e, para todo $x \in X$,*

$$\mathrm{D}f|_x = \mathop{\mathbin{+}}_{k\in[n]} \mathrm{D}_k f|_x \circ \mathrm{p}_k.$$

$\square$ *Demonstração.* Sejam $\boldsymbol{E} := \boldsymbol{E}_0, \dots, \boldsymbol{E}_{n-1}$, $x \in X$ e, para cada $k \in [n]$, $V_k \subseteq A$ uma vizinhança aberta de x na qual f é k-parcialmente diferenciável e $\mathrm{D}_k f\colon V_k \longrightarrow \mathscr{L}(\boldsymbol{E}_k, \boldsymbol{E}')$ é contínua em x. Definamos $V := \bigcap_{k\in[n]} V_k$, de modo que V é uma vizinhança aberta de x na qual, para todo $k \in [n]$, f é k-parcialmente diferenciável e $\mathrm{D}_k f\colon V \longrightarrow \mathscr{L}(\boldsymbol{E}_k, \boldsymbol{E}')$ é contínua em x.

Seja $\delta' \in \,]0, \infty[$ tal que $\mathrm{B}\,(x; \delta') \subseteq V$. Para todo $x' \in \mathrm{B}\,(x; \delta')$, definamos

$$g(x') := f(x') - f(x) - \mathop{\mathbin{+}}_{k\in[n]} \mathrm{D}_k f|_x (x'_k - x_k)$$

Mostraremos que $\lim_{x'\to x} \frac{\|g(x')\|}{\|x'-x\|} = 0$, o que implica que f é diferenciável em x e $\mathrm{D}f|_x = +_{k\in[n]} \mathrm{D}_k f|_x \circ \mathrm{p}_k$.

Para todo $k \in [n]$, a função $g\colon \mathrm{B}\,(x; \delta') \subseteq E \longrightarrow E'$ é k-parcialmente diferenciável em $\mathrm{B}\,(x; \delta')$. Isso vale porque, para todo $k' \in [n]$,

$$\mathrm{p}_{k'} \circ \iota_{k,x'}(y) = \mathrm{p}_{k'}({x'}_0, \dots, {x'}_{k-1}, y, {x'}_{k+1}, \dots, {x'}_{n-1}) = \begin{cases} y, & k = k' \\ {x'}_{k'}, & k \neq k', \end{cases}$$

logo

$$\begin{aligned} g_{k,x'}(y) &= g(\iota_{k,x'}(y)) \\ &= f(\iota_{k,x'}(y)) - f(x) - \mathop{\mathbin{+}}_{k'\in[n]} \mathrm{D}_{k'} f|_x (\mathrm{p}_{k'} \circ \iota_{k,x'}(y) - x_{k'}) \\ &= f_{k,x'}(y) - f(x) - \mathrm{D}_k f|_x(y) + \mathrm{D}_k f|_x(x_k) - \mathop{\mathbin{+}}_{k'\in[n]\smallsetminus\{k\}} \mathrm{D}_{k'} f|_x ({x'}_{k'} - x_{k'}), \end{aligned}$$

e então, como $-f(x) + \mathrm{D}_k f|_x(x_k) - +_{k'\in[n]\smallsetminus\{k\}} \mathrm{D}_{k'} f|_x ({x'}_{k'} - x_{k'})$ é constante (com relação a y) e $\mathrm{D}_k f|_x$ é linear, segue que

$$\mathrm{D}_k g|_{x'} = \mathrm{D}g_{k,x'}|_{{x'}_k} = \mathrm{D}f_{k,x'}|_{{x'}_k} - \mathrm{D}_k f|_x = \mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x.$$

Seja $\varepsilon \in \,]0, \infty[$. Para todo $k \in [n]$, a função $\mathrm{D}_k f$ é contínua em x, logo para algum $\delta \in \,]0, \delta'[$ vale que, para todo $x' \in \mathrm{B}\,(x; \delta)$,

$$\|\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x\| < \frac{\varepsilon}{n}.$$

Seja $v \in E$ tal que $\|v\| \leq \delta'$ e, para todo $k \in [n]$, $v_{(0)} := (0, \ldots, 0)$ e $v_{(k+1)} := (v_0, \ldots, v_k, 0, \ldots, 0)$. Vale $v_{(n)} = v$, $v_{(k+1)} - v_{(k)} = v_k$ e $\left\|v_{(k)}\right\| \leq \|v\|$ (para qualquer norma p).

Da desigualdade do valor médio (19.64) segue que, para todo $k \in [n]$,

$$\left\|g_{k,x'_{(k)}}(x_k + v_k) - g_{k,x'_{(k)}}(x_k)\right\| \leq \sup_{y \in]x_k, x_k + v_k[_{\triangle 1}} \|\mathrm{D}g_{k,x'}|_y\| \, \|x_k + v_k - x_k\| \leq \frac{\varepsilon}{n} \|v_k\|$$

Como $g(x) = 0$, vale que

$$g(x + v) = \mathop{+}_{k \in [n]} g(x + v_{(k+1)}) - g(x + v_{(k)}),$$

portanto

$$\begin{aligned}
\|g(x + v)\| &\leq \mathop{+}_{k \in [n]} \left\|g(x + v_{(k+1)}) - g(x + v_{(k)})\right\| \\
&= \mathop{+}_{k \in [n]} \left\|g_{k,x'_{(k)}}(x_k + v_k) - g_{k,x'_{(k)}}(x_k)\right\| \\
&\leq \mathop{+}_{k \in [n]} \frac{\varepsilon}{n} \|v_k\| \\
&\leq \mathop{+}_{k \in [n]} \frac{\varepsilon}{n} \|v\| \\
&= \varepsilon \|v\| .
\end{aligned}$$

Disso segue que f é diferenciável em x e, para todo $v \in E$,

$$\mathrm{D}f|_x(v) = \mathop{+}_{k \in [n]} \mathrm{D}_k f|_x(v_k). \qquad \blacksquare$$

⊢ **Proposição 19.72.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E_0 \times \cdots \times E_{n-1}$ um aberto e $f \colon A \longrightarrow E'$ uma função. A função f é $\mathscr{C}^1$-diferenciável se, e somente se, para todo $k \in [n]$, f é k-parcialmente $\mathscr{C}^1$-diferenciável. Nesse caso, para todo $x \in A$,*

$$\mathrm{D}f|_x = \mathop{+}_{k \in [n]} \mathrm{D}_k f|_x \circ \mathrm{p}_k.$$

□ *Demonstração.* A ida segue direto de 19.70. Para a volta, segue de de 19.71 que f é diferenciável e que, para todo $x \in A$,

$$\mathrm{D}f|_x = \mathop{+}_{k \in [n]} \mathrm{D}_k f|_x \circ \mathrm{p}_k.$$

Seja $\varepsilon \in]0, \infty[$. Como para todo $k \in [n]$ as funções $\mathrm{D}_k f$ são contínuas em x, algum δ_k satisfaz que, para todo $x' \in \mathrm{B}(x; \delta_k) \cap A$,

$$\|\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x\| \leq \frac{\varepsilon}{n}.$$

Logo, para $\delta := \inf_{k \in [n]} \delta_k$, segue que, para todo $x' \in \mathrm{B}(x; \delta_k) \cap A$,

$$\begin{aligned}
\|\mathrm{D}f|_{x'} - \mathrm{D}f|_x\| &= \left\| \mathop{+}_{k \in [n]} (\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x) \circ \mathrm{p}_k \right\| \\
&\leq \mathop{+}_{k \in [n]} \|(\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x) \circ \mathrm{p}_k\| \\
&\leq \mathop{+}_{k \in [n]} \|\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x\| \, \|\mathrm{p}_k\| \\
&\leq \mathop{+}_{k \in [n]} \|\mathrm{D}_k f|_{x'} - \mathrm{D}_k f|_x\| \\
&= \varepsilon,
\end{aligned}$$

o que mostra que $\mathrm{D}f$ é contínua em x. ■

Combinando os dois casos, em que

$$f \colon A \subseteq E_0 \times \cdots \times E_{n-1} \longrightarrow E'_0 \times \cdots \times E'_{n'-1}$$

é uma função $\mathscr{C}^1$-diferenciável, para todo $x \in A$ vale que

$$\mathrm{D}f|_x = \mathop{+}_{(k,k') \in [n] \times [n']} \iota_{k'} \circ \mathrm{D}_k f_{k'}|_x \circ \mathrm{p}_k.$$

No caso em que $f \colon A \subseteq \mathbb{R}^d \longrightarrow E'$, as diferenciais parciais estão relacionadas com as derivadas direcionais ao longo de cada vetor da base canônica. A relação é

$$\mathrm{D}_k f|_x(c) = c\partial_{\mathrm{i}_k} f(x).$$

Uma consequência direta desses resultados é a diferenciabilidade contínua de uma função multilinear.

⊢ **Proposição 19.73** (Diferencial de função multilinear)**.** *Sejam $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{n-1}$ e $\boldsymbol{E}'$ espaços normados e $l \colon E_0 \times \cdots \times E_{n-1} \longrightarrow E'$ uma função n-linear. A função l é $\mathscr{C}^1$-diferenciável e, para todos $x, v \in E_0 \times \cdots \times E_{n-1}$,*

$$\mathrm{D}l|_x(v) = \mathop{+}_{k \in [n]} l(x_0, \ldots, x_{k-1}, v_k, x_{k+1}, \ldots, x_{n-1}).$$

□ *Demonstração.* Sejam $x \in E_0 \times \cdots \times E_{n-1}$ e $k \in [n]$. Da definição multilinearidade, $l \circ \iota_{k,x} \colon E_k \longrightarrow E'$ é linear, portanto $\mathscr{C}^1$-diferenciável e, para todo $w \in E_k$,

$$l \circ \iota_{k,x}(w) = l(x_0, \ldots, x_{k-1}, w, x_{k+1}, \ldots, x_{n-1}).$$

Segue de 19.72 que l é $\mathscr{C}^1$-diferenciável e

$$\mathrm{D}l|_x(v) = \mathop{+}_{k\in[n]} l(x_0, \ldots, x_{k-1}, v_k, x_{k+1}, \ldots, x_{n-1}).$$ ■

19.7.6.3 Diferenciabilidade das normas p

Nesta seção, analisamos o caso das normas p em $\mathbb{R}^d$. Consideraremos $\mathbb{R}^d$ como produto de d espaços $\mathbb{R}$ para poder calcular as diferenciais parciais D_k. As demonstrações não serão apresentadas, mas podem ser conferidas em [Col12, §2.6].

⊢ **Proposição 19.74.** *Sejam $d \in \mathbb{N} \smallsetminus \{0\}$ e $p \in [1, \infty]$.*

1. *Se $p = 1$, a norma $\|\cdot\|_1$ é $\mathscr{C}^1$-diferenciável no aberto*

$$A := \left\{ x \in \mathbb{R}^d \;\middle|\; \forall_{k\in[d]} x_k \neq 0 \right\}$$

e, para todos $x \in A$ e $k \in [d]$,

$$\mathrm{D}_k(\|\cdot\|_1)|_x = \begin{cases} 1, & x_k > 0 \\ -1, & x_k < 0. \end{cases}$$

A norma $\|\cdot\|_1$ não é diferenciável em $E \smallsetminus A$.

2. *Se $p \in \,]1, \infty[$, a norma $\|\cdot\|_p$ é $\mathscr{C}^1$-diferenciável em $E \smallsetminus \{0\}$ e, para todos $x \in E \smallsetminus \{0\}$ e $k \in [d]$,*

$$\mathrm{D}_k(\|\cdot\|_p)|_x = \begin{cases} \|x\|_p^{1-p}\, x_k^{p-1}, & x_k > 0 \\ 0, & x_k = 0 \\ \|x\|_p^{1-p}\, (-x_k)^{p-1}, & x_k < 0. \end{cases}$$

3. *Se $p = \infty$, a norma $\|\cdot\|_\infty$ é $\mathscr{C}^1$-diferenciável no aberto*

$$A := \left\{ x \in \mathbb{R}^d \;\middle|\; \exists_{k\in[d]} x_k = \|x\|_\infty \wedge \forall_{k'\in[d]\smallsetminus\{k\}} x_{k'} \neq \|x\|_\infty \right\}$$

e, para todos $x \in A$ e $k \in [d]$,

$$\mathrm{D}_k(\|\cdot\|_\infty)|_x = \begin{cases} 0, & x_k \neq \|x\|_\infty\, x_k > 0 \\ 1, & x_k = \|x\|_\infty\,, x_k > 0 \\ -1, & x_k = \|x\|_\infty\,, x_k < 0. \end{cases}$$

A norma $\|\cdot\|_\infty$ não é diferenciável em $E \smallsetminus A$.

19.7.7 Multidiferenciabilidade

Nesta seção, estudaremos a generalização dos tópicos estudados até agora para diferenciais de ordem superior. Brevemente, definiremos uma função multidiferenciável como uma função que pode ser diferenciada múltiplas vezes no sentido de que sua diferencial é diferenciável (para ordem 2), a diferencial de sua diferencial é diferenciável (para ordem 3), e assim por diante. As propriedades de diferenciabilidade relacionadas às operações algébricas, à composição e às diferenciais parciais serão também generalizadas, dessa vez com uma apresentação mais sucinta.

19.7.7.1 Multidiferencial e propriedades

Consideremos espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$, um aberto $A \subseteq E$, $X \subseteq A$ e uma função $f\colon A \longrightarrow E'$ diferenciável em X. Sua diferencial em X é a função

$$\begin{aligned} \mathrm{D}f\colon X &\longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}') \\ x &\longmapsto \mathrm{D}f|_x. \end{aligned}$$

Lembremos que adotamos a seguinte notação

$$\mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}') := \mathscr{L}(\boldsymbol{E}, \mathscr{L}(\boldsymbol{E}, \cdots \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}') \cdots)).$$

e que denotamos o conjunto das transformações m-lineares contínuas de $\boldsymbol{E}$ para $\boldsymbol{E}'$ por

$$\mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}') := \mathscr{L}(\boldsymbol{E}, \cdots, \boldsymbol{E}; \boldsymbol{E}').$$

Esses dois espaços são isometricamente isomorfos (19.38) e o isomorfismo é dado por

$$I^{(m)}\colon \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}') \longrightarrow \mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}').$$

Além disso, o subespaço das funções m-lineares simétricas $\mathscr{L}_S^{(m)}(\boldsymbol{E}, \boldsymbol{E}')$ é um subespaço fechado de $\mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}')$.

Se a função f é diferenciável em uma vizinhança aberta V de X, podemos considerar a diferenciabilidade de sua diferencial

$$\begin{aligned} \mathrm{D}f\colon V &\longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}') \\ x &\longmapsto \mathrm{D}f|_x. \end{aligned}$$

Se $\mathrm{D}f$ for diferenciável em X, ela admite uma diferencial

$$\mathrm{D}^2 f := \mathrm{D}(\mathrm{D}f)\colon X \longrightarrow \mathscr{L}^2(\boldsymbol{E}, \boldsymbol{E}').$$

Se por sua vez a função $\mathrm{D}^2 f$ estiver definida na vizinhança V, então pode-se considerar sua diferencial $\mathrm{D}^3 f := \mathrm{D}(\mathrm{D}^2 f)\colon X \longrightarrow \mathscr{L}^3(\boldsymbol{E}, \boldsymbol{E}')$ e assim por diante até algum $m \in \mathbb{N}$

$$\mathrm{D}^m f := \mathrm{D}(\mathrm{D}^{m-1} f)\colon X \longrightarrow \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}').$$

Definimos a seguir o conceito de multidiferenciabilidade baseado nessa discussão preliminar. No entanto, queremos entender a função $\mathrm{D}^m f|_x \in \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}')$ como uma função m-linear, pois elas têm propriedades melhores. Para isso, usaremos o isomorfismo isométrico $I^{(m)}$.

:⊢ **Definição 19.41** (Multidiferenciabilidade)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$. Para todo $m \in \overset{\infty}{\mathbb{N}}$, definimos indutivamente:

1. Uma função de A para E' 0-*diferenciável em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$. Sua 0-*diferencial composta* é $\mathrm{D}^0 f := f\colon A \longrightarrow E'$. Uma função 0-*diferenciável* de A para E' é uma função $f\colon A \longrightarrow E'$;
2. Uma função de A para E' m-*diferenciável em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para alguma vizinhança $V \subseteq A$ de X, f é $(m-1)$-diferenciável em V e sua $(m-1)$-diferencial composta $\mathrm{D}^{m-1} f\colon V \longrightarrow \mathscr{L}^{m-1}(\boldsymbol{E}, \boldsymbol{E}')$ é diferenciável em X. Sua m-*diferencial composta* é
$$\mathrm{D}^m f := \mathrm{D}(\mathrm{D}^{m-1} f)\colon X \longrightarrow \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}')$$
e sua m-*diferencial multilinear* é
$$\mathrm{D}^{(m)} f := I^{(m)} \circ (D^m f)\colon X \longrightarrow \mathscr{L}^{(m)}(\boldsymbol{E}, \boldsymbol{E}').$$
Uma função m-*diferenciável* de A para E' é uma função $f\colon A \longrightarrow E'$ m-diferenciável em A;
3. Uma função de A para E' ∞-*diferenciável* (ou *suave*) *em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para todo $m \in \mathbb{N}$, f é m-diferenciável em X.
Uma função ∞-*diferenciável* (ou *suave*) de A para E' é uma função $f\colon A \longrightarrow E'$ ∞-diferenciável em A.

Denominamos a função $\mathrm{D}^m f$ de m-diferencial composta de f e a função $\mathrm{D}^{(m)} f$ de m-diferencial multilinear, mas essa nomenclatura não é padrão. Em geral, $\mathrm{D}^{(m)} f$ é considerada a m-diferencial e a função não recebe nome, mas preferimos enfatizar que elas estão relacionadas, embora sejam distintas, e ter como nos referirmos a cada uma delas. A proposição a seguir mostra que a diferenciabilidade de $\mathrm{D}^m f$ é equivalente à diferenciabilidade de $\mathrm{D}^{(m)} f$.

⊢ **Proposição 19.75** (Equivalência das multidiferenciais)**.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$ *espaços normados,* $A \subseteq E$ *um aberto,* $X \subseteq A$ *e* $f\colon A \longrightarrow E'$ *uma função e* $m \in \mathbb{N} \smallsetminus \{0\}$*. A função* f *é* m*-diferenciável em* X *se, e somente se, para alguma vizinhança* V *de* X*,* f *é* $(m-1)$*-diferenciável em* V *e sua* $(m-1)$*-diferencial multilinear* $\mathrm{D}^{(m-1)} f$ *é diferenciável em* X*. Nesse caso, para todo* $x \in X$ *e todos* $v_0, \ldots, v_{m-1} \in E$*,*
$$\mathrm{D}(\mathrm{D}^{(m-1)} f)|_x(v_0)(v_1, \ldots, v_{m-1}) = \mathrm{D}^{(m)} f|_x(v_0, \ldots, v_{m-1}).$$

□ *Demonstração.* Para $m = 1$ é direto. Consideremos $m \geq 2$.

- ($\Rightarrow$) Suponhamos que f é m-diferenciável em X e seja V uma vizinhança de X em que f é $(m-1)$-diferenciável e tal que $\mathrm{D}^{m-1}f\colon V \longrightarrow \mathscr{L}^{m-1}(\boldsymbol{E}, \boldsymbol{E}')$ é diferenciável em X. Como para todo $x' \in V$ vale que $\mathrm{D}^{(m-1)}f|_{x'} = I^{(m-1)} \circ \mathrm{D}^{m-1}f|_{x'}$, então $\mathrm{D}^{(m-1)}f = I^{(m-1)} \circ \mathrm{D}^{m-1}f$ e segue da linearidade e continuidade de $I^{(m)}$ que $\mathrm{D}^{(m-1)}f$ é diferenciável em X e, pela regra da cadeia, para todo $x \in X$,

$$\mathrm{D}(\mathrm{D}^{(m-1)}f)|_x = \mathrm{D}(I^{(m-1)} \circ \mathrm{D}^{m-1}f)|_x = I^{(m-1)} \circ \mathrm{D}^m f|_x.$$

- ($\Leftarrow$) Suponhamos que, para alguma vizinhança V de X, f é $(m-1)$-diferenciável em V e $\mathrm{D}^{(m-1)}f$ é diferenciável em X. Para todo $x' \in V$, vale que $\mathrm{D}^{m-1}f|_{x'} = (I^{(m)})^{-1} \circ \mathrm{D}^{(m-1)}f|_{x'}$, portanto $\mathrm{D}^{m-1}f = (I^{(m)})^{-1} \circ \mathrm{D}^{(m-1)}f$ e segue da linearidade e continuidade de $(I^{(m)})^{-1}$ que $\mathrm{D}^{m-1}f$ é diferenciável em X e, pela regra da cadeia, para todo $x \in X$,

$$\mathrm{D}^m f|_x = \mathrm{D}(\mathrm{D}^{m-1}f)|_x = \mathrm{D}((I^{(m)})^{-1} \circ \mathrm{D}^{(m-1)}f)|_x = (I^{(m)})^{-1} \circ \mathrm{D}(\mathrm{D}^{(m-1)}f)|_x.$$

Assim, concluímos que, caso f seja m-diferenciável em X, vale que, para todo $x \in X$ e todos $v_0, \ldots, v_{m-1} \in E$,

$$\begin{aligned}\mathrm{D}(\mathrm{D}^{(m-1)}f)|_x(v_0)(v_1, \ldots, v_{m-1}) &= I^{(m-1)} \circ \mathrm{D}^m f|_x(v_0)(v_1, \ldots, v_{m-1}) \\ &= I^{(m)} \circ \mathrm{D}^m f|_x(v_0, v_1, \ldots, v_{m-1}) \\ &= \mathrm{D}^{(m)}f|_x(v_0, v_1, \ldots, v_{m-1}).\end{aligned}$$ ■

▷ **Exercício 19.14.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função, $m \in \mathbb{N} \smallsetminus \{0\}$ e $k \in [m]$. A função f é m-diferenciável em X se, e somente se, para alguma vizinhança V de X, f é k-diferenciável em V e $\mathrm{D}^{(k)}f\colon V \longrightarrow \mathscr{L}^{(k)}(\boldsymbol{E}, \boldsymbol{E}')$ é $m-k$-diferenciável em X. Nesse caso, para todo $x \in X$,*

$$\mathrm{D}^{(m-k)}\mathrm{D}^{(k)}f|_x = \mathrm{D}^{(m)}f|_x.$$

Além de ser multilinear, a multidiferencial $\mathrm{D}^{(m)}f|_x$ é simétrica, mais um dos motivos pelos quais ela é mais vantajosa em aplicações. Demonstraremos isso a seguir, primeiro para o caso $m = 2$, então para o caso geral.

⊢ **Proposição 19.76** (Simetria da multidiferencial)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $f\colon A \longrightarrow E'$ uma função m-diferenciável em X. Para todo $x \in X$, a função m-linear $\mathrm{D}^{(m)}f|_x\colon E \times \cdots \times E \longrightarrow E'$ é simétrica.*

□ *Demonstração.* Seja $x \in X$. Demostraremos a proposição por indução em m. Os casos $m = 0$ e $m = 1$ são triviais. Consideramos o caso $m = 2$. Seja $V \subseteq A$

uma vizinhança de x tal que f é diferenciável em V e $\mathrm{D}f\colon V \longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$ é diferenciável em x.

Sejam $v_0, v_1 \in E$. Se $v_0 = 0$ ou $v_1 = 0$, então

$$\mathrm{D}^{(2)} f|_x(v_0, v_1) = 0 = \mathrm{D}^{(2)} f|_x(v_1, v_0).$$

Consideremos $v_0 \neq 0$ e $v_1 \neq 0$.

Seja $\varepsilon \in]0, \infty[$. Como f é 2-diferenciável em x, para algum $\delta \in]0, \infty[$ vale que $\mathrm{B}(x; \delta) \subseteq V$ e, para todo $x' \in \mathrm{B}(x; \delta)$,

$$\left\| \mathrm{D}f|_{x'} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(x' - x) \right\| \leq \frac{\varepsilon}{2^4 \|v_0\| \|v_1\|} \|x' - x\|.$$

Sejam $v, v' \in E$ tais que $x + v, x + v', x + v + v' \in \mathrm{B}(x; \delta)$. Definimos

$$r(v, v') = f(x + v + v') - f(x + v) - f(x + v') + f(x) - \mathrm{D}^{(2)} f(v, v')$$

e então segue que

$$\begin{aligned}
\mathrm{D}^{(2)} f|_x(v, v') - \mathrm{D}^{(2)} f|_x(v', v) &= \mathrm{D}^{(2)} f|_x(v, v') \\
&\quad - f(x + v + v') + f(x + v) + f(x + v') - f(x) \\
&\quad + f(x + v' + v) - f(x + v') - f(x + v) + f(x) \\
&\quad - \mathrm{D}^{(2)} f|_x(v', v) \\
&= -r(v, v') + r(v', v),
\end{aligned}$$

portanto

$$\left\| \mathrm{D}^{(2)} f|_x(v, v') - \mathrm{D}^{(2)} f|_x(v', v) \right\| \leq \|r(v, v')\| + \|r(v', v)\|.$$

Agora, como f é diferenciável em V, definimos

$$r'(v') := f(x + v + v') - f(x + v') - \mathrm{D}f|_{x+v}(v') + \mathrm{D}f|_x(v').$$

Como f é 2-diferenciável em x, segue de 19.75 que $\mathrm{D}^{(2)} f|_x(v, v') = \mathrm{D}^2 f|_x(v)(v')$, logo

$$\begin{aligned}
r(v, v') &= f(x + v + v') - f(x + v') - \mathrm{D}f|_{x+v}(v') + \mathrm{D}f|_x(v') \\
&\quad - (f(x + v) - f(x)) \\
&\quad + (\mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v))(v') \\
&= r'(v') - r'(0) + (\mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v))(v'),
\end{aligned}$$

portanto

$$\|r(v, v')\| \leq \|r'(v') - r'(0)\| + \left\| \mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v) \right\| \|v'\|.$$

Como $\mathrm{B}(0;\delta)$ é convexa, segue da desigualdade do valor médio (19.64) que

$$\|r'(v') - r'(0)\| \leq \sup_{t\in]0,1[} \|\mathrm{D}r'|_{tv'}\| \, \|v'\| \, .$$

Mas r' é diferenciável, pois é soma de funções diferenciáveis, e vale que

$$\begin{aligned}
\mathrm{D}r'|_{v'} &= \mathrm{D}f_{x+v+v'} - \mathrm{D}f|_{x+v} - \mathrm{D}f|_{x+v'} + \mathrm{D}f|_x \\
&= \mathrm{D}f_{x+v+v'} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v+v') \\
&\quad - (\mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v)) \\
&\quad - (\mathrm{D}f|_{x+v'} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v'));
\end{aligned}$$

como v e v' são tais que $x+v, x+v', x+v+v' \in \mathrm{B}(x;\delta)$, segue da 2-diferenciabilidade de f que

$$\begin{aligned}
\|\mathrm{D}r'|_{v'}\| &\leq \left\|\mathrm{D}f_{x+v+v'} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v+v')\right\| \\
&\quad + \left\|\mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v)\right\| \\
&\quad + \left\|\mathrm{D}f|_{x+v'} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v')\right\| \\
&\leq \frac{\varepsilon}{2^4 \|v_0\| \, \|v_1\|}(\|v+v'\| + \|v\| + \|v'\|) \\
&\leq \frac{\varepsilon}{2^3 \|v_0\| \, \|v_1\|}(\|v\| + \|v'\|).
\end{aligned}$$

Para $t \in]0,1[$, vale $\|tv'\| \leq \|v'\|$, o que implica que

$$\begin{aligned}
\|r'(v') - r'(0)\| &\leq \sup_{t\in]0,1[} \|\mathrm{D}r'|_{tv'}\| \, \|v'\| \\
&\leq \frac{\varepsilon}{2^3 \|v_0\| \, \|v_1\|}(\|v\| + \|tv'\|) \, \|v'\| \\
&\leq \frac{\varepsilon}{2^3 \|v_0\| \, \|v_1\|}(\|v\| + \|v'\|) \, \|v'\| \, .
\end{aligned}$$

Disso segue que

$$\begin{aligned}
\|r(v,v')\| &\leq \|r'(v') - r'(0)\| + \left\|\mathrm{D}f|_{x+v} - \mathrm{D}f|_x - \mathrm{D}^2 f|_x(v)\right\| \|v'\| \\
&\leq \frac{\varepsilon}{2^3 \|v_0\| \, \|v_1\|}(\|v\| + \|v'\|) \, \|v'\| + \frac{\varepsilon}{2^4 \|v_0\| \, \|v_1\|} \|v\| \, \|v'\| \\
&= \frac{\varepsilon}{2^4 \|v_0\| \, \|v_1\|} \left(3 \|v\| \, \|v'\| + 2 \|v'\|^2\right) \\
&\leq \frac{\varepsilon}{2^4 \|v_0\| \, \|v_1\|} \left(2 \|v\|^2 + 4 \|v\| \, \|v'\| + 2 \|v'\|^2\right) \\
&= \frac{\varepsilon}{2^3 \|v_0\| \, \|v_1\|}(\|v\| + \|v'\|)^2.
\end{aligned}$$

Assim, obtém-se

$$\begin{aligned}\left\|\mathrm{D}^{(2)}f|_x(v,v') - \mathrm{D}^{(2)}f|_x(v',v)\right\| &\leq \frac{\varepsilon}{2^3\,\|v_0\|\,\|v_1\|}(\|v\|+\|v'\|)^2 \\ &\quad + \frac{\varepsilon}{2^3\,\|v_0\|\,\|v_1\|}(\|v'\|+\|v\|)^2 \\ &= \frac{\varepsilon}{2^2\,\|v_0\|\,\|v_1\|}(\|v\|+\|v'\|)^2.\end{aligned}$$

Finalmente, sejam $c \in \,]0,\delta[$ e $v := \frac{c}{2}\frac{v_0}{\|v_0\|}, v' := \frac{c}{2}\frac{v_0}{\|v_0\|}$, de modo que

$$\|v\| = \|v'\| = \frac{c}{2},$$

logo $x+v, x+v', x+v+v' \in \mathrm{B}\,(x;\delta)$. Da bilinearidade de $\mathrm{D}^{(2)}f|_x$, segue que

$$\begin{aligned}\mathrm{D}^{(2)}f|_x(v_0,v_1) &= \mathrm{D}^{(2)}f|_x\left(\frac{2\,\|v_0\|}{c}v, \frac{2\,\|v_1\|}{c}v'\right) = \frac{2^2\,\|v_0\|\,\|v_1\|}{c^2}\mathrm{D}^{(2)}f|_x(v,v'), \\ \mathrm{D}^{(2)}f|_x(v_1,v_0) &= \mathrm{D}^{(2)}f|_x\left(\frac{2\,\|v_1\|}{c}v', \frac{2\,\|v_0\|}{c}v\right) = \frac{2^2\,\|v_1\|\,\|v_0\|}{c^2}\mathrm{D}^{(2)}f|_x(v',v),\end{aligned}$$

portanto

$$\begin{aligned}\left\|\mathrm{D}^{(2)}f|_x(v_0,v_1) - \mathrm{D}^{(2)}f|_x(v_1,v_0)\right\| &= \frac{2^2\,\|v_0\|\,\|v_1\|}{c^2}\left\|\mathrm{D}^{(2)}f|_x(v,v') - \mathrm{D}^{(2)}f|_x(v',v)\right\| \\ &\leq \frac{2^2\,\|v_0\|\,\|v_1\|}{c^2}\frac{\varepsilon}{2^2\,\|v_0\|\,\|v_1\|}(\|v\|+\|v'\|)^2 \\ &\leq \frac{\varepsilon}{c^2}\left(\frac{c}{2}+\frac{c}{2}\right)^2 \\ &\leq \varepsilon,\end{aligned}$$

o que implica que $\mathrm{D}^{(2)}f|_x(v_0,v_1) = \mathrm{D}^{(2)}f|_x(v_1,v_0)$.

Agora, consideramos o caso em que $m > 2$. Suponhamos que f é m-diferenciável em x e que toda função $(m-1)$-diferenciável tem $(m-1)$-diferencial simétrica. Seja $V \subseteq A$ uma vizinhança convexa de x tal que f é $(m-1)$-diferenciável em V e $\mathrm{D}^{(m-1)}f\colon V \longrightarrow \mathscr{L}(\boldsymbol{E},\boldsymbol{E}')$ é diferenciável em x.

Por hipótese de indução, $\mathrm{D}^{(m-1)}f|_x$ é uma função $(m-1)$-linear simétrica. Como $\mathscr{L}_S^{(m-1)}(\boldsymbol{E},\boldsymbol{E}')$ é um subespaço fechado de $\mathscr{L}^{(m-1)}(\boldsymbol{E},\boldsymbol{E}')$ e $(I^{(m-1)})^{-1}$ é isomorfismo isométrico, então

$$S' := (I^{(m-1)})^{-1}(\mathscr{L}_S^{(m-1)}(\boldsymbol{E},\boldsymbol{E}'))$$

é subespaço fechado de $\mathscr{L}^{m-1}(\boldsymbol{E},\boldsymbol{E}')$, portanto segue de 19.43 que $\mathrm{D}^{m-1}f\colon V \longrightarrow S'$ é diferenciável e $\mathrm{D}^m f|_x = \mathrm{D}(\mathrm{D}^{m-1}f)|_x$ tem imagem em S', ou seja, para todo

$v \in E$ a função $\mathrm{D}(\mathrm{D}^{m-1}f)|_x(v) \in S'$, o que implica que $I^{(m-1)}(\mathrm{D}(\mathrm{D}^{m-1}f)|_x(v))$ é $(m-1)$-linear simétrica. Assim, segue de 19.75 que a função m-linear $\mathrm{D}^{(m)}f|_x$ é simétrica nas entradas $1, \ldots, m-1$; ou seja, para todos $v, v'_0, \ldots, v'_{m-2} \in E$ toda permutação $\sigma' \colon [m-1] \longrightarrow [m-1]$,

$$\mathrm{D}^{(m)}f|_x(v, v'_{\sigma(0)} \ldots, v'_{\sigma(m-2)}) = \mathrm{D}^{(m)}f|_x(v, v'_0 \ldots, v'_{m-2}).$$

Para obter simetria em todas as entradas, consideremos a função $\mathrm{D}^{m-2}f$. Vale que $\mathrm{D}^2(\mathrm{D}^{m-2}f)|_x = \mathrm{D}^m f|_x$ e que $\mathrm{D}^2(\mathrm{D}^{m-2}f)|_x \in \mathscr{L}^2(\boldsymbol{E}, \mathscr{L}^{m-2}(\boldsymbol{E}, \boldsymbol{E}')) = \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}')$. Pelo caso base para $m = 2$, vale que, para todos $v, v' \in E$,

$$\mathrm{D}^{(2)}(\mathrm{D}^{m-2}f)|_x(v, v') = \mathrm{D}^{(2)}(\mathrm{D}^{m-2}f)|_x(v, v'),$$

portanto, para todos $v, v', v''_0, \ldots, v''_{m-3} \in E$,

$$\mathrm{D}^{(m)}f|_x(v, v', v''_0, \ldots, v''_{m-3}) = \mathrm{D}^{(m)}f|_x(v', v, v''_0, \ldots, v''_{m-3}).$$

Finalmente, sejam $v_0, \ldots, v_{m-1} \in E$ e $\sigma \colon [m] \longrightarrow [m]$ uma permutação. Se $\sigma(0) = 0$, então $\sigma|_{\{1,\ldots,m-1\}}$ é uma permutação de $\{1, \ldots, m-1\}$, portanto

$$\begin{aligned}\mathrm{D}^{(m)}f|_x(v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)}) &= \mathrm{D}^{(m)}f|_x(v_0, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)}) \\ &= \mathrm{D}^{(m)}f|_x(v_0, v_1 \ldots, v_{m-1}).\end{aligned}$$

Se $\sigma(0) \neq 0$, seja $k_0 \in \{1, \ldots, m-1\}$ tal que $\sigma(k_0) = 0$. Podemos permutar as entradas $v_{\sigma(1)}$ e $v_{\sigma(k_0)} = v_0$, então as entradas $v_{\sigma(0)}$ com $v_{\sigma(k_0)} = v_0$, e por fim as restantes entradas $v_{\sigma(k)}$ para $k \in [m] \smallsetminus \{k_0\}$, e obtemos

$$\begin{aligned}\mathrm{D}^{(m)}f|_x(v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)}) &= \mathrm{D}^{(m)}f|_x(v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_{\sigma(k_0)}, \ldots, v_{\sigma(m-1)}) \\ &= \mathrm{D}^{(m)}f|_x(v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_0, \ldots, v_{\sigma(m-1)}) \\ &= \mathrm{D}^{(m)}f|_x(v_{\sigma(0)}, v_0, \ldots, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)}) \\ &= \mathrm{D}^{(m)}f|_x(v_0, v_{\sigma(0)}, \ldots, v_{\sigma(1)}, \ldots, v_{\sigma(m-1)}) \\ &= \mathrm{D}^{(m)}f|_x(v_0, v_1, \ldots, v_{m-1}).\end{aligned}$$

■

Assim como em 19.41, também podemos demonstrar que adição, ação por escalar e (no caso de uma álgebra normada como contradomínio) multiplicação de funções preservam a multidiferenciabilidade das funções.

▷ **Exercício 19.15.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $m \in \overset{\infty}{\mathbb{N}}$.*

1. *Para todas funções $f, f' \colon A \longrightarrow E'$ m-diferenciáveis em X, a função*

$$f + f' \colon A \longrightarrow E'$$

é m-diferenciável em X;

2. *Para todo $c \in \mathbb{R}$ e toda função $f\colon A \longrightarrow E'$ m-diferenciável em X, a função*

$$cf\colon A \longrightarrow E'$$

é m-diferenciável em X;

3. *Se $\boldsymbol{E}'$ é uma álgebra normada, para todas funções $f, f'\colon A \longrightarrow E'$ m-diferenciáveis em X, a função*

$$ff'\colon A \longrightarrow E'$$

é m-diferenciável em X;

4. *Toda função constante $f\colon A \longrightarrow E'$ é m-diferenciável;*
5. *Toda função linear $l\colon E \longrightarrow E'$ é m-diferenciável;*
6. *Toda função n-multilinear $l\colon E_0 \times \cdots \times E_{n-1} \longrightarrow E'$ é m-diferenciável.*

Assim como no caso da diferencial (19.42), a composição de funções multidiferenciáveis também é multidiferenciável. A seguir, provamos esse resultado, mas não exibiremos uma expressão análoga à da regra da cadeia para multidiferenciais; tal expressão fica a cargo do leitor.

⊢ **Proposição 19.77** (Multidiferenciabilidade da composição)**.** *Sejam $\boldsymbol{E}$, $\boldsymbol{E}'$ e $\boldsymbol{E}'$ espaços normados, $m \in \mathbb{N} \smallsetminus \{0\}$, $A \subseteq E$ e $A' \subseteq E'$ abertos, $X \subseteq A$, $f\colon A \longrightarrow E'$ m-diferenciável em X tal que $f(A) \subseteq A'$ e $f\colon A' \longrightarrow E''$ m-diferenciável em $f(X)$. A função composta $f' \circ f\colon A \longrightarrow E''$ é m-diferenciável em X.*

□ *Demonstração.* Mostraremos por indução em m. Seja $x \in X$. Para $m = 1$, segue direto de 19.42. Suponhamos que a proposição vale para $m \geq 2$ e mostremos que vale também para $m + 1$. Como f e f' são diferenciáveis, da regra da cadeia segue que

$$\mathrm{D}(f' \circ f)|_x = \mathrm{D}f'|_{f(x)} \circ \mathrm{D}f|_x.$$

Para mostrar que $f' \circ f$ é $(m+1)$-diferenciável, mostraremos que sua diferencial $\mathrm{D}(f' \circ f)$ é m-diferenciável. Note que $\mathrm{D}(f' \circ f)$ é composição de duas funções. A primeira é a composição de funções

$$\begin{aligned} \circ\colon \mathscr{L}(E', E'') \times \mathscr{L}(E, E') &\longrightarrow \mathscr{L}(E, E'') \\ (l', l) &\longmapsto l' \circ l, \end{aligned}$$

que é função bilinear, logo ∞-diferenciável e portanto m-diferenciável; a segunda é a função

$$\begin{aligned} g\colon X &\longrightarrow \mathscr{L}(E', E'') \times \mathscr{L}(E, E') \\ x &\longmapsto (\mathrm{D}f'|_{f(x)}, \mathrm{D}f|_x). \end{aligned}$$

A função f é $(m+1)$-diferenciável por hipótese, logo m-diferenciável; a função $\mathrm{D}f'$ é m-diferenciável pois por hipótese f' é $(m+1)$-diferenciável. Assim, por hipótese de indução, a primeira componente de g, que é a composição $\mathrm{D}f' \circ f$, é m-diferenciável. A segunda componente de g é a função $\mathrm{D}f$, que é m-diferenciável pois por hipótese f é $(m+1)$-diferenciável. Assim segue que a função g é m-diferenciável, pois cada uma de suas componentes é m-diferenciável.

Agora, novamente pela hipótese de indução, a diferencial $\mathrm{D}(f' \circ f)$ é m-diferenciável, pois é composição de duas funções m-diferenciáveis, e portanto $f' \circ f$ é $(m+1)$-diferenciável. ■

19.7.7.2 Multidiferenciabilidade contínua

Assim como no caso da continuidade da diferencial, a multidiferencial de uma função f nem sempre é contínua. Nesta seção, generalizaremos o conceito de diferenciabilidade contínua para multidiferenciabilidade.

:⊢ **Definição 19.42** (Multidiferenciabilidade contínua)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto e $X \subseteq A$. Para todo $m \in \overset{\infty}{\mathbb{N}}$, definimos indutivamente:

1. Uma função de A para E' $\mathscr{C}^0$*-diferenciável* (ou *continuamente* 0*-diferenciável*) *em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ contínua em X.
   Uma função $\mathscr{C}^0$*-diferenciável* (ou *continuamente* 0*-diferenciável*) de A para E' é uma função contínua $f\colon A \longrightarrow E'$ e o conjunto dessas funções é denotado $\mathscr{C}^0(\boldsymbol{A}, \boldsymbol{E}') := \mathscr{C}(\boldsymbol{A}, \boldsymbol{E}')$;
2. Uma função de A para E' $\mathscr{C}^m$*-diferenciável* (ou *continuamente* m*-diferenciável*) *em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para alguma vizinhança $V \subseteq A$ de X, f é m-diferenciável em V e $\mathrm{D}^m f\colon V \longrightarrow \mathscr{L}^m(\boldsymbol{E}, \boldsymbol{E}')$ é contínua em X.
   Uma função $\mathscr{C}^m$*-diferenciável* (ou *continuamente* m*-diferenciável*) de A para E' é uma função $f\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciável em A e o conjunto dessas funções é denotado $\mathscr{C}^m(\boldsymbol{A}, \boldsymbol{E}')$;
3. Uma função de A para E' $\mathscr{C}^\infty$*-diferenciável* (ou *continuamente* ∞*-diferenciável* ou *suave*) *em* X (ou *em* x para $X = \{x\}$) é uma função $f\colon A \longrightarrow E'$ tal que, para todo $m \in \mathbb{N}$, f é $\mathscr{C}^m$-diferenciável em X.
   Uma função $\mathscr{C}^\infty$*-diferenciável* (ou *continuamente* ∞*-diferenciável* ou *suave*) de A para E' é uma função $f\colon A \longrightarrow E'$ $\mathscr{C}^\infty$-diferenciável em A e o conjunto dessas funções é denotado $\mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}')$.

Para todo $m \in \mathbb{N}$, a diferencial é uma função

$$\begin{aligned} \mathrm{D}\colon \mathscr{C}^m(\boldsymbol{A}, \boldsymbol{E}') &\longrightarrow \mathscr{C}^{m-1}(\boldsymbol{A}, \boldsymbol{E}) \\ f &\longmapsto \mathrm{D}f \end{aligned}$$

e, quando nos restringimos a $\mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}')$,

$$\begin{aligned} \mathrm{D}\colon \mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}') &\longrightarrow \mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}) \\ f &\longmapsto \mathrm{D}f. \end{aligned}$$

O espaço $\mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}')$ é um espaço linear sobre $\mathbb{R}$ e 19.41 mostra que D é uma função linear em $\mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}')$. No entanto, ela nem sempre é contínua. No caso em que $\boldsymbol{E}'$ é uma álgebra, também por 19.41 concluímos que D é uma derivação em $\mathscr{C}^\infty(\boldsymbol{A}, \boldsymbol{E}')$. Em especial, para $E' = \mathbb{R}$, D é uma derivação em $\mathscr{C}^\infty(\boldsymbol{A}, \mathbb{R})$. Essa perspectiva será importante no estudo da derivada exterior de formas, abordado mais à frente.

▷ **Exercício 19.16.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$, $f\colon A \longrightarrow E'$ uma função e $m \in \mathbb{N}$. Então f é $\mathscr{C}^m$-diferenciável em X se, e somente se, para todo $k \in [m]$, f é $\mathscr{C}^k$-diferenciável em X e $\mathrm{D}^k f$ é $\mathscr{C}^{m-k}$-diferenciável em X.*

▷ **Exercício 19.17.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ um aberto, $X \subseteq A$ e $m \in \overset{\infty}{\mathbb{N}}$.*

1. *Para todas funções $f, f'\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciáveis em X, a função $f + f'\colon A \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável em X;*
2. *Para todo $c \in \mathbb{R}$ e toda função $f\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciável em X, a função $cf\colon A \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável em X;*
3. *Se $\boldsymbol{E}'$ é uma álgebra normada, para todas funções $f, f'\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciáveis em X, a função $ff'\colon A \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável em X;*
4. *Toda função constante $f\colon A \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável;*
5. *Toda função linear $l\colon E \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável;*
6. *Toda função n-multilinear $l\colon E_0 \times \cdots \times E_{n-1} \longrightarrow E'$ é $\mathscr{C}^m$-diferenciável.*

⊢ **Proposição 19.78** (Multidiferenciabilidade contínua da composição)**.** *Sejam $\boldsymbol{E}$, $\boldsymbol{E}'$ e $\boldsymbol{E}'$ espaços normados, $m \in \mathbb{N} \smallsetminus \{0\}$, $A \subseteq E$ e $A' \subseteq E'$ abertos, $X \subseteq A$, $f\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciável em X tal que $f(A) \subseteq A'$ e $f\colon A' \longrightarrow E''$ $\mathscr{C}^m$-diferenciável em $f(X)$. A função composta $f' \circ f\colon A \longrightarrow E''$ é $\mathscr{C}^m$-diferenciável em X.*

□ *Demonstração.* Segue de 19.77 e do fato de que composição de funções contínuas é contínua. ■

19.7.7.3 Multiderivabilidade

Da mesma forma que podemos repetidamente diferenciar uma função, podemos repetidamente derivar uma trajetória. A seguir definiremos o conceito de derivadas múltiplas e demonstraremos que ele é equivalente à diferenciação múltipla para trajetórias.

:⊢ **Definição 19.43** (Multiderivabilidade)**.** Sejam $\boldsymbol{E}$ um espaço normado, $I \subseteq E$ um intervalo aberto e $X \subseteq I$. Para todo $m \in \overset{\infty}{\mathbb{N}}$, definimos indutivamente:

1. Uma trajetória de I para E 0-*derivável em* X (ou *em* t para $X = \{t\}$) é uma trajetória $\gamma\colon I \longrightarrow E$. Sua 0-*derivada* é $\dot{\partial}^0 f := f\colon A \longrightarrow E'$.
   Uma trajetória 0-*derivável* de I para E é uma trajetória $f\colon I \longrightarrow E$;
2. Uma trajetória de I para E m-*derivável em* X (ou *em* t para $X = \{t\}$) é uma trajetória $\gamma\colon I \longrightarrow E$ tal que, para alguma vizinhança $V \subseteq I$ de X, γ é $(m-1)$-diferenciável em V e sua $(m-1)$-derivada $\dot{\partial}^{m-1} f\colon V \longrightarrow E$ é derivável em X. Sua m-*derivada* é
   $$\dot{\partial}^m f := \dot{\partial}(\dot{\partial}^{m-1} f)\colon X \longrightarrow E$$
   e, para todo $t \in X$, a m-*derivada* de f em x é $\dot{\partial}^m f(t)$.
   Uma trajetória m-*derivável* de I para E é uma trajetória $\gamma\colon I \longrightarrow E$ m-derivável em I;
3. Uma trajetória de I para E ∞-*derivável* (ou *suave*) *em* X (ou *em* t para $X = \{t\}$) é uma trajetória $\gamma\colon I \longrightarrow E$ tal que, para todo $m \in \mathbb{N}$, γ é m-derivável em X.
   Uma trajetória ∞-*derivável* (ou *suave*) de I para E é uma trajetória $\gamma\colon I \longrightarrow E$ ∞-derivável em I.

⊢ **Proposição 19.79.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $I \subseteq \mathbb{R}$ *aberto,* $X \subseteq I$ *e* $m \in \mathbb{N} \smallsetminus \{0\}$*. Uma trajetória* $\gamma\colon I \longrightarrow E$ *é* m*-diferenciável em* X *se, e somente se, é* m*-derivável em* X*. Nesse caso, para todo* $t \in X$*, vale* $\left\|\mathrm{D}^{(m)} f|_t\right\| = \left\|\dot{\partial}^m f(t)\right\|$ *e, para todos* $c_0, \ldots, c_{m-1} \in \mathbb{R}$*,*

$$\mathrm{D}^{(m)} f|_t(c_0, \ldots, c_{m-1}) = c_0 \cdots c_{m-1} \dot{\partial}^m f(t).$$

□ *Demonstração.* Exercício. ■

19.7.7.4 Multiderivabilidade contínua

Como no caso da diferencial, podemos generalizar derivabilidade contínua para multiderivabilidade contínua. Como derivabilidade implica continuidade, se γ é derivável então γ é contínua, ou seja, $\gamma \in \mathscr{C}(\boldsymbol{I}, \boldsymbol{E})$. Podemos nos perguntar se a derivada $\dot{\gamma}$ de γ é também contínua. Isso nem sempre vale, embora seja possível mostrar que uma descontinuidade de uma função derivável sempre é uma descontinuidade essencial e nunca é uma descontinuidade de salto.

:⊢ **Definição 19.44** (Multiderivabilidade contínua)**.** Sejam $\boldsymbol{E}$ um espaço normado, $I \subseteq \mathbb{R}$ um intervalo aberto e $X \subseteq I$. Para todo $m \in \overset{\infty}{\mathbb{N}}$, definimos indutivamente:

1. Uma trajetória de I para E $\mathscr{C}^0$-*derivável* (ou *continuamente* 0-*derivável*) em X (ou *em* t para $X = \{t\}$) é uma trajetória $\gamma\colon I \longrightarrow E$ contínua em X e sua 0-ésima derivada é $\dot{\partial}^0\gamma := \gamma$.
   Uma trajetória $\mathscr{C}^0$-*derivável* de I para E é uma trajetória contínua $\gamma\colon I \longrightarrow E$ e sua 0-ésima derivada é $\dot{\partial}^0\gamma := \gamma$. O conjunto dessas trajetórias é denotado $\mathscr{C}^0(\boldsymbol{I}, \boldsymbol{E}) := \mathscr{C}(\boldsymbol{I}, \boldsymbol{E})$.
2. Uma trajetória $\mathscr{C}^m$-*derivável* de I para E é uma trajetória derivável $\gamma\colon I \longrightarrow E$ tal que $\dot{\partial}\gamma\colon I \longrightarrow E$ é $\mathscr{C}^{m-1}$-derivável. Sua m-ésima derivada é
$$\dot{\partial}^m\gamma := \dot{\partial}^{m-1}(\dot{\partial}\gamma).$$
   O conjunto dessas trajetórias é denotado $\mathscr{C}^m(\boldsymbol{I}, \boldsymbol{E})$;
3. Uma trajetória $\mathscr{C}^\infty$-*derivável* (ou *suave*) de I para E é uma trajetória $\gamma\colon I \longrightarrow E$ tal que, para todo $m \in \mathbb{N}$, γ é $\mathscr{C}^m$-derivável. O conjunto dessas trajetórias é denotado $\mathscr{C}^\infty(\boldsymbol{I}, \boldsymbol{E})$.

Por 19.79, a notação $\mathscr{C}^m(\boldsymbol{A}, \boldsymbol{E}')$ não é ambígua no caso $A = I \subseteq \mathbb{R}$.

19.7.8 Extremos locais e pontos críticos degenerados

Na seção 19.7.4, analisamos pontos extremos locais de campos escalares usando informações sobre a primeira diferencial do campo no ponto. Definimos o conceito de ponto crítico e o relacionamos aos pontos extremos locais. Nesta seção, o objetivo é analisar pontos extremos locais com informações sobre a segunda diferencial do campo escalar; além disso, definimos o conceito de ponto degenerado. As demonstrações serão omitidas, mas podem ser encontradas em [Col12, p.115].

19.7.8.1 Extremos locais e a segunda diferencial

Analogamente à proposição 19.56, a seguinte proposição oferece uma condição necessária para que pontos sejam extremos, mas usando a segunda diferencial.

⊢ **Proposição 19.80.** *Seja* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $x \in A$ *e* $f\colon A \longrightarrow \mathbb{R}$ *um campo escalar* 2-*diferenciável em* x.

- *Se* x *é ponto de máximo local de* f*, então, para todo* $v \in E$*,*
$$\mathrm{D}^{(2)}f|_x(v, v) \leq 0.$$
- *Se* x *é ponto de mínimo local de* f*, então, para todo* $v \in E$*,*
$$\mathrm{D}^{(2)}f|_x(v, v) \geq 0.$$

Apresentamos então duas condições suficientes, mas não necessárias, para que um ponto seja extremo. A primeira apresenta uma condição para que o ponto seja um extremo e a segunda, para que seja um ponto extremo estrito.

⊢ **Proposição 19.81.** *Seja* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $f\colon A \longrightarrow \mathbb{R}$ *um campo escalar* 2*-diferenciável e* $x \in A$ *um ponto crítico de* f*.*

- *Se, para alguma bola aberta* $\mathrm{B}(x;\varepsilon)$*, vale que, para todo* $x' \in \mathrm{B}(x;\varepsilon)$ *e todo* $v \in E$*,*
$$\mathrm{D}^{(2)} f|_{x'}(v,v) \leq 0,$$
então x *é ponto de máximo local de* f*.*
- *Se, para alguma bola aberta* $\mathrm{B}(x;\varepsilon)$*, vale que, para todo* $x' \in \mathrm{B}(x;\varepsilon)$ *e todo* $v \in E$*,*
$$\mathrm{D}^{(2)} f|_{x'}(v,v) \geq 0,$$
então x *é ponto de mínimo local de* f*.*

⊢ **Proposição 19.82.** *Seja* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $f\colon A \longrightarrow \mathbb{R}$ *um campo escalar e* $x \in A$ *um ponto crítico de* f *tal que* f *é* 2*-diferenciável em* x*.*

- *Se, para algum* $\varepsilon \in \mathbb{R}_{>0}$*, todo* $v \in E$ *satisfaz*
$$\mathrm{D}^{(2)} f|_{x}(v,v) \leq \varepsilon \left\|v\right\|^2,$$
então x *é ponto de máximo local estrito de* f*.*
- *Se, para algum* $\varepsilon \in \mathbb{R}_{>0}$*, todo* $v \in E$ *satisfaz*
$$\mathrm{D}^{(2)} f|_{x'}(v,v) \geq \varepsilon \left\|v\right\|^2,$$
então x *é ponto de mínimo local estrito de* f*.*

19.7.8.2 Pontos críticos degenerados

:⊢ **Definição 19.45.** Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ aberto e $f\colon A \longrightarrow \mathbb{R}$ um campo escalar. Um ponto crítico *degenerado* de f é um ponto crítico $x \in A$ de f tal que f é 2-diferenciável em x e
$$\mathrm{D}^2 f|_x \colon E \longrightarrow E^*$$
não é um isomorfismo de espaços lineares topológicos. Um ponto crítico *não-degenerado* de f é um ponto crítico $x \in A$ de f que não é degenerado.

⊢ **Proposição 19.83.** *Seja* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $f\colon A \longrightarrow \mathbb{R}$ *um campo escalar e* $x \in A$ *um ponto crítico não-degenerado de* f*.*

- *Se, para todo $v \in E$,*
$$\mathrm{D}^{(2)} f|_x(v,v) \leq 0,$$
então, para algum $\varepsilon \in \mathbb{R}_{>0}$, todo $v \in E$ satisfaz
$$\mathrm{D}^{(2)} f|_x(v,v) \leq \varepsilon \|v\|^2 .$$
- *Se, para todo $v \in E$,*
$$\mathrm{D}^{(2)} f|_x(v,v) \geq 0,$$
então, para algum $\varepsilon \in \mathbb{R}_{>0}$, todo $v \in E$ satisfaz
$$\mathrm{D}^{(2)} f|_{x'}(v,v) \geq \varepsilon \|v\|^2 .$$

19.7.9 Invertibilidade

Nesta seção, enunciaremos alguns resultados importantes sem prová-los. Demonstrações podem ser encontradas em [Car71] ou [Col12] e são consequências diretas dos resultados das seções anteriores.

⊢ **Proposição 19.84.** *Seja $\boldsymbol{E}$ uma álgebra normada completa. A inversão multiplicativa*
$$\begin{aligned} \diagup \colon E^{\diagup} &\longrightarrow E^{\diagup} \\ x &\longmapsto x^{-1} \end{aligned}$$
é uma função $\mathscr{C}^\infty$-diferenciável e, para todo $x \in E^{\diagup}$ e todo $v \in E$,
$$\mathrm{D}(\diagup)|_x(v) = -x^{-1} v x^{-1}.$$

□ *Demonstração.* A demonstração de que $\diagup$ é $\mathscr{C}^\infty$-diferenciável será indutiva. Mostraremos que, para todo $m \in \mathbb{N}$, $\diagup$ é $\mathscr{C}^m$-diferenciável.

O caso base será mostrar para $m = 1$. Primeiro, mostremos que $\diagup$ é diferenciável. Seja $x \in E^{\diagup}$ e definamos a função
$$\begin{aligned} l_x \colon E &\longrightarrow E \\ v &\longmapsto -x^{-1} v x^{-1}. \end{aligned}$$
A função l_x é linear e, como vale
$$\left\| -x^{-1} v x^{-1} \right\| \leq \left\| -x^{-1} \right\|^2 \|v\| , \tag{19.2}$$

também é contínua. Notemos que, para todo $v \in E$, vale que

$$\begin{aligned}
\|\diagup(x+v) - \diagup(x) - l_x(v)\| &= \left\|(x+v)^{-1} - x^{-1} + x^{-1}vx^{-1}\right\| \\
&= \left\|(x+v)^{-1}(x - (x+v) + (x+v)x^{-1}v)x^{-1}\right\| \\
&= \left\|(x+v)^{-1}(vx^{-1})^2\right\| \\
&\leq \left\|(x+v)^{-1}\right\| \left\|x^{-1}\right\|^2 \|v\|^2 .
\end{aligned}$$

Isso mostra que $\diagup$ é diferenciável e $\mathrm{D}(\diagup)|_x = l_x$.

Mostremos agora que $\diagup$ é $\mathscr{C}^1$-diferenciável. Definamos a função

$$\begin{aligned}
\alpha \colon E \times E &\longrightarrow \mathscr{L}(\boldsymbol{E}) \\
(a,b)) &\longmapsto \alpha(w,w'') \colon E \longrightarrow E \\
&\qquad\qquad\qquad v \longmapsto -wvw'
\end{aligned}$$

A função α é claramente bilinear; para mostrar sua continuidade, notemos que

$$\|\alpha(w,w')(v)\| \leq \|w\| \|w'\| \|v\| ,$$

logo $\|\alpha(w,w')\|_{\mathscr{L}(\boldsymbol{E})} \leq \|w\| \|w'\|$. Assim, como $\mathrm{D}(\diagup)|_x = \alpha(x^{-1}, x^{-1})$, segue que $\mathrm{D}(\diagup)$ é composição de funções contínuas, logo é contínua, o que mostra que $\diagup$ é $\mathscr{C}^1$-diferenciável.

Agora mostraremos o passo indutivo. Seja $m \in \mathbb{N}$ e suponhamos que $\diagup$ seja $\mathscr{C}^m$-diferenciável. A função α é bilinear, logo $\mathscr{C}^m$-diferenciável, o que implica que a diferencial $\mathrm{D}(\diagup)$, dada pela composição $\mathrm{D}(\diagup)_x = \alpha(x^{-1}, x^{-1})$ é $\mathscr{C}^m$-diferenciável. Isso significa que $\diagup$ é $\mathscr{C}^{m+1}$-diferenciável. ■

⊢ **Proposição 19.85.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados completos e $m \in \mathbb{N}$. A inversão de função*

$$\begin{aligned}
(\cdot)^{-1} \colon \overset{\leftrightarrow}{\mathscr{L}}(E,E') &\longrightarrow \overset{\leftrightarrow}{\mathscr{L}}(E',E) \\
f &\longmapsto f^{-1}
\end{aligned}$$

é uma função $\mathscr{C}^\infty$-diferenciável.

19.7.9.1 Difeomorfismos

:⊢ **Definição 19.46** (Difeomorfismo)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados e $A \subseteq E$ e $A' \subseteq E'$ abertos. Um *difeomorfismo* de A para A' é uma bijeção $f \colon A \longrightarrow A'$ tal que f e f^{-1} são diferenciáveis.

:⊢ **Definição 19.47** (Multidifeomorfismo)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados e $A \subseteq E$ e $A' \subseteq E'$ abertos e $m \in \mathbb{N}$. Um *m-difeomorfismo* de A para A' é uma bijeção $f \colon A \longrightarrow A'$ tal que f e f^{-1} são m-diferenciáveis.

:⊢ **Definição 19.48** (Multidifeomorfismo contínuo)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados, $A \subseteq E$ e $A' \subseteq E'$ abertos e $m \in \mathbb{N}$. Um *$\mathscr{C}^m$-difeomorfismo* de A para A' é um difeomorfismo $f\colon A \longrightarrow A'$ tal que f e f^{-1} são $\mathscr{C}^m$-diferenciáveis.

⊢ **Proposição 19.86** (Invertibilidade da diferencial de difeomorfismo)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados completos, $A \subseteq E$ e $A' \subseteq E'$ abertos e $f\colon A \longrightarrow A'$ um difeomorfismo. Para todo $x \in A$, $\mathrm{D}f|_x$ é invertível e*

$$(\mathrm{D}f|_x)^{-1} = \mathrm{D}(f^{-1})|_{f(x)}.$$

□ *Demonstração.* Da invertibilidade de f vale que $f^{-1} \circ f = \mathrm{I}_A$ e $f \circ f^{-1} = \mathrm{I}_{A'}$. Como f é um difeomorfismo, segue da regra da cadeia que, para todo $x \in A$ e todo $x' \in A'$,

$$\begin{aligned} \mathrm{D}(f^{-1})|_{f(x)} \circ \mathrm{D}f|_x &= \mathrm{I}_E \\ \mathrm{D}f|_{f^{-1}(x')} \circ \mathrm{D}(f^{-1})|_{x'} &= \mathrm{I}_{E'}. \end{aligned}$$

Tomando $x' = f(x)$, segue que $f^{-1}(x') = x$, logo que

$$\mathrm{D}f|_x \circ \mathrm{D}(f^{-1})|_{f(x)} = \mathrm{D}f|_{f^{-1}(x')} \circ \mathrm{D}(f^{-1})|_{x'} = \mathrm{I}_{E'}.$$

Disso segue que $\mathrm{D}f|_x$ é invertível e

$$(\mathrm{D}f|_x)^{-1} = \mathrm{D}(f^{-1})|_{f(x)}.$$ ■

Em particular, essa proposição implica que se as dimensões dos espaços $\boldsymbol{E}$ e $\boldsymbol{E}'$ são diferentes, não pode existir um difeomorfismo entre abertos desses espaços, pois a diferencial desses difeomorfismo em cada ponto seria um isomorfismo linear entre espaços de dimensões distintas.

⊢ **Proposição 19.87** (Difeomorfismo multidiferenciável)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados completos, $A \subseteq E$ e $A' \subseteq E'$ abertos, $m \in \mathbb{N}$ e $f\colon A \longrightarrow A'$ um difeomorfismo $\mathscr{C}^m$-diferenciável. A inversa $f^{-1}\colon A' \longrightarrow A$ é $\mathscr{C}^m$-diferenciável, ou seja, f é um $\mathscr{C}^m$-difeomorfismo.*

□ *Demonstração.* A demonstração se dará por indução em m. Primeiro consideramos o caso base $m = 1$. Como f é difeomorfismo, segue de 19.86 que $\mathrm{D}f|_x$ é invertível e $(\mathrm{D}f|_x)^{-1} = \mathrm{D}(f^{-1})|_{f(x)}$. Aplicando isso em $y \in A'$, obtém-se

$$\mathrm{D}(f^{-1})|_y = (\mathrm{D}f|_{f^{-1}(y)})^{-1},$$

portanto vale que

$$\mathrm{D}(f^{-1}) = (\cdot)^{-1} \circ \mathrm{D}f \circ f^{-1},$$

em que $(\cdot)^{-1}\colon \overleftrightarrow{\mathscr{L}}(E,E') \longrightarrow \overleftrightarrow{\mathscr{L}}(E',E)$ é a inversão de função. Como f é $\mathscr{C}^1$-diferenciável, então $\mathrm{D}f$ é contínua, logo $\mathrm{D}(f^{-1})$ é composição de funções contínuas. Isso implica que f^{-1} é $\mathscr{C}^1$-diferenciável, o que significa que f é um $\mathscr{C}^1$-difeomorfismo.

Para demonstrarmos o passo indutivo, seja $m \in \mathbb{N}$ e suponhamos que a proposição vale para m. Suponhamos que f seja um difeomorfismo $\mathscr{C}^{m+1}$-diferenciável. Devemos mostrar que f^{-1} é também $\mathscr{C}^{m+1}$-diferenciável. Como f é uma função $\mathscr{C}^{m+1}$-diferenciável, é também $\mathscr{C}^m$-diferenciável, portanto segue da hipótese de indução que f^{-1} é $\mathscr{C}^m$-diferenciável. Além disso, $\mathrm{D}f$ é $\mathscr{C}^m$-diferenciável e a inversão de função $^{-1}\colon \overleftrightarrow{\mathscr{L}}(E,E') \longrightarrow \overleftrightarrow{\mathscr{L}}(E',E)$ é $\mathscr{C}^m$-diferenciável (19.85). Como vale $\mathrm{D}(f^{-1}) = (\cdot)^{-1} \circ \mathrm{D}f \circ f^{-1}$ e todas a funções da composição são $\mathscr{C}^m$-diferenciáveis, segue que $\mathrm{D}(f^{-1})$ é $\mathscr{C}^m$-diferenciável, logo f^{-1} é $\mathscr{C}^{m+1}$-diferenciável. ■

19.7.9.2 Teorema da Função Inversa

Na seção anterior, demonstramos que a diferencial de um difeomorfismo é invertível em cada ponto do domínio. Agora, procuramos uma recíproca. Se a diferencial de uma função é invertível em um ponto, é possível afirmar que a função em si é invertível? Isso nem sempre vale, pois diferenciabilidade é um conceito local, mas sob certas condições é possível garantir a invertibilidade local da função. Para isso, antes definimos um conceito local de difeomorfismo.

:⊢ **Definição 19.49** (Difeomorfismo local)**.** Sejam $\boldsymbol{E}$ e $\boldsymbol{E'}$ espaços normados, $A \subseteq E$ um aberto e $x \in A$. Um *difeomorfismo local em* x de A para E' é uma função diferenciável $f\colon A \longrightarrow E'$ tal que, para alguma vizinhança $V \subseteq A$ de x, $f(V)$ é aberto em E' e $f|_V\colon V \longrightarrow f(V)$ é um difeomorfismo. Um *difeomorfismo local* de A para E' é uma função diferenciável $f\colon A \longrightarrow E'$ que é um difeomorfismo local em todos pontos de seu domínio.

Seja $m \in \mathbb{N}$. Um $\mathscr{C}^m$*-difeomorfismo local* de A para E' é um difeomorfismo local $f\colon A \longrightarrow E'$ $\mathscr{C}^m$-diferenciável.

As proposições apresentadas nesta seção vão em geral considerar funções $\mathscr{C}^1$-diferenciáveis ou com uma regularidade maior. A menos de alguns resultados, não será suficiente supor a diferenciabilidade somente em um ponto do domínio. No entanto, resultados análogos podem ser obtidos supondo diferenciabilidade forte em somente um ponto do domínio, que de certa forma representa uma diferenciabilidade contínua somente naquele único ponto. Essa abordagem é adotada por Lima em seu *Curso de Análise* e tais resultados podem ser conferidos lá [Lim14, pp. 286, 287].

⊢ **Proposição 19.88** (Diferenciabilidade do homeomorfismo inverso)**.** *Sejam* $\boldsymbol{E}$ *e* $\boldsymbol{E'}$ *espaços normados completos,* $A \subseteq E$ *e* $A' \subseteq E'$ *abertos,* $x \in A$ *e* $f\colon A \longrightarrow A'$

um homeomorfismo diferenciável em x. Se $\mathrm{D}f|_x$ é invertível, então a inversa f^{-1} é diferenciável em $f(x)$ e

$$\mathrm{D}f^{-1}|_{f(x)} = (\mathrm{D}f|_x)^{-1}.$$

Para todo $m \in \mathbb{N} \smallsetminus \{0\}$, se f é $\mathscr{C}^m$-diferenciável, então f é $\mathscr{C}^m$-difeomorfismo local em x.

□ *Demonstração.* Mostremos que a inversa diferencial de f^{-1} em $f(x)$ é $(\mathrm{D}f|_x)^{-1}$. Como $\mathrm{D}f|_x$ é linear e invertível, sua inversa $(\mathrm{D}f|_x)^{-1}$ é linear e, como os espaços $\boldsymbol{E}$ e $\boldsymbol{E}'$ são completos, contínua. Como A' é aberto, algum $\delta \in \mathbb{R}_{>0}$ satisfaz que, para todo $v' \in \mathrm{B}\,(0;\delta)$, $f(x) + v' \in A'$. Pela bijetividade de f, para cada $v' \in \mathrm{B}\,(0;\delta)$ um único $v \in E$ satisfaz $f(x+v) = f(x) + v'$, o que significa que v é função de v' e implica que

$$f^{-1}(f(x) + v') - f^{-1}(f(x)) = f^{-1}(f(x+v)) - x = x + v - x = v.$$

Da continuidade de f^{-1}, segue $\lim_{v' \to 0} v = 0$. Da diferenciabilidade de f em x, a função resto $r(v)$ definida por

$$r(v) := f(x+v) - f(x) - \mathrm{D}f|_x(v)$$

satisfaz $\lim_{v \to 0} r(v) \diagup \|v\| = 0$. De

$$v' = f(x+v) - f(x) = \mathrm{D}f|_x(v) + r(v)$$

e da expressão anterior para v em termos de f^{-1} segue que

$$\begin{aligned}(\mathrm{D}f|_x)^{-1}(v') &= v + (\mathrm{D}f|_x)^{-1}(r(v)) \\ &= f^{-1}(f(x) + v') - f^{-1}(f(x)) + (\mathrm{D}f|_x)^{-1}(r(v)).\end{aligned}$$

Definamos

$$r'(v') := f^{-1}(f(x) + v') - f^{-1}(f(x)) - (\mathrm{D}f|_x)^{-1}(v'),$$

de modo que (lembrando que v é função de v')

$$r'(v') = -(\mathrm{D}f|_x)^{-1}(r(v)),$$

e então, para mostrar que f^{-1} é diferenciável em $f(x)$, devemos mostrar que

$$\lim_{v' \to 0} \frac{r'(v')}{\|v'\|} = \lim_{v' \to 0} \frac{f^{-1}(f(x) + v') - f^{-1}(f(x)) - (\mathrm{D}f|_x)^{-1}(v')}{\|v'\|} = 0.$$

Notemos que

$$\frac{r'(v')}{\|v'\|} = -(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right)\frac{\|v\|}{\|v'\|}$$

Como $\lim_{v\to 0} r(v) \diagup \|v\| = 0$ e $(\mathrm{D}f|_x)^{-1}$ é contínua, vale $\lim_{v\to 0}(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right) = 0$. Resta mostrar que $\frac{\|v\|}{\|v'\|}$ é limitado nas proximidades de $v = 0$.

Da expressão $(\mathrm{D}f|_x)^{-1}(v') = v + (\mathrm{D}f|_x)^{-1}(r(v))$ segue que

$$\begin{aligned}\|v\| &\leq \left\|(\mathrm{D}f|_x)^{-1}(r(v))\right\| + \left\|(\mathrm{D}f|_x)^{-1}(v')\right\| \\ &\leq \|v\| \left\|(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right)\right\| + \left\|(\mathrm{D}f|_x)^{-1}\right\| \|v'\|,\end{aligned}$$

portanto

$$\frac{\|v\|}{\|v'\|} \leq \frac{\|(\mathrm{D}f|_x)^{-1}\|}{1 - \left\|(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right)\right\|}.$$

Como $\lim_{v\to 0} r(v) \diagup \|v\| = 0$ e $(\mathrm{D}f|_x)^{-1}$ é contínua, segue que

$$\lim_{v\to 0}(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right) = 0,$$

logo

$$\lim_{v\to 0}\frac{\|v\|}{\|v'\|} \leq \left\|(\mathrm{D}f|_x)^{-1}\right\|.$$

Assim segue que

$$\lim_{v\to 0}\frac{r'(v')}{\|v'\|} = \lim_{v\to 0} -(\mathrm{D}f|_x)^{-1}\left(\frac{r(v)}{\|v\|}\right)\frac{\|v\|}{\|v'\|} = 0$$

e, como $\lim_{v'\to 0} v = 0$, segue que

$$\lim_{v'\to 0}\frac{r'(v')}{\|v'\|} = 0,$$

o que mostra que f^{-1} é diferenciável em $f(x)$.

Agora, suponhamos que f é $\mathscr{C}^m$-diferenciável, com $m \in \mathbb{N} \smallsetminus \{0\}$. Isso implica que $\mathrm{D}f$ é contínua. O conjunto $\overset{\leftrightarrow}{\mathscr{L}}(\boldsymbol{E}, \boldsymbol{E}')$ de isomorfismos lineares contínuos é um subconjunto aberto de $\mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$, portanto dada uma vizinhança aberta $U \subseteq \overset{\leftrightarrow}{\mathscr{L}}(\boldsymbol{E}, \boldsymbol{E}')$ de $\mathrm{D}f|_x$, o conjunto $V := (\mathrm{D}f)^{-1}(U) \subseteq E$ é aberto e, para todo $x' \in V$, $\mathrm{D}f|_{x'}$ é invertível. Pela primeira parte da proposição, segue que f^{-1} é diferenciável em todo $y \in f(V)$, portanto $f|_V$ é um difeomorfismo. Como f é $\mathscr{C}^m$-diferenciável, de 19.87 segue que f é um $\mathscr{C}^m$-difeomorfismo local em x. ■

Usando essa proposição, demonstraremos uma das proposições mais importantes da teoria de diferenciabilidade, o Teorema da Função Inversa. Esse teorema é a base de muitos resultados da teoria de diferenciabilidade em espaços normados e também da teoria de diferenciabilidade em variedades, que será tratada mais à frente. Ele implica, em particular, o Teorema da Função Implícita, e o Teorema do Posto, que é válido em dimensão finita.

⊢ **Teorema 19.89** (Função Inversa)**.** *Sejam $\boldsymbol{E}$ e $\boldsymbol{E}'$ espaços normados completos, $A \subseteq E$ um aberto, $x \in A$, $m \in \mathbb{N} \smallsetminus \{0\}$ e $f\colon A \longrightarrow E'$ uma função $\mathscr{C}^m$-diferenciável. Se $\mathrm{D}f|_x$ é invertível, então f é um $\mathscr{C}^m$-difeomorfismo local em x e*

$$\mathrm{D}f^{-1}|_{f(x)} = (\mathrm{D}f|_x)^{-1}.$$

□ *Demonstração.* Começaremos considerando o caso simplificado em que $x = 0$ e $f(x) = 0$. Como f é $\mathscr{C}^1$-diferenciável, a função $\mathrm{D}f\colon A \longrightarrow \mathscr{L}(\boldsymbol{E}, \boldsymbol{E}')$ é contínua. Consideremos a função resto $r(x' - x) = f(x') - f(x) - \mathrm{D}f|_x(x' - x)$. No caso simplificado, isso se resume a

$$r(x') = f(x') - f(x) - \mathrm{D}f|_x(x' - x) = f(x') - \mathrm{D}f|_x(x'),$$

logo

$$f = r + \mathrm{D}f|_x.$$

Segue f é uma distorção do isomorfismo linear $\mathrm{D}f|_x$ numa vizinhança de $x = 0$, portanto de 19.45 concluímos que, para $\varepsilon < \|(\mathrm{D}f|_x)^{-1}\|^{-1}$, algum $\delta \in \mathbb{R}_{>0}$ satisfaz que r restrita a $V := \mathrm{B}\,(0; \delta) \cap A$ é controlada e a distorção de r satisfaz $\langle\!\langle r \rangle\!\rangle \leq \varepsilon$, portanto $\|(\mathrm{D}f|_x)^{-1}\|\,\langle\!\langle r \rangle\!\rangle < 1$. Assim, por 19.34 concluímos que f restrita a V é um mergulho métrico, portanto um mergulho topológico (16.39), ou seja, um homeomorfismo com sua imagem. Finalmente, como f é $\mathscr{C}^m$-diferenciável, segue de 19.88 que f é um $\mathscr{C}^m$-difeomorfismo local em x.

Devemos agora considerar o caso geral. Para isso, definimos a função

$$\phi(v) := f(x + v) - f(x),$$

que é $\mathscr{C}^m$-diferenciável e satisfaz $\phi(0) = 0$ e $\mathrm{D}\phi|_0 = \mathrm{D}f|_x$, portanto invertível. Do caso simplificado existem vizinhanças $U \subseteq E$ de $v = 0$ e $U' \subseteq E'$ de $f(v) = 0$ tais que $\phi|_U\colon U \longrightarrow U'$ é um $\mathscr{C}^m$-difeomorfismo. Definindo $V := x + U$, segue que

$$f|_V(x') = f(x) + \phi(x' - x)$$

e $f|_V$ é uma bijeção sobre $V' := f(x) + U'$. A inversa de $f|_V$ é

$$f|_V^{-1}(y) = x + \phi^{-1}(y - f(x)),$$

também $\mathscr{C}^m$-diferenciável, portanto f é um $\mathscr{C}^m$-difeomorfismo local em x. ■

19.7.9.3 Teorema da Função Implícita

⊢ **Teorema 19.90** (Função Implícita)**.** *Sejam $\boldsymbol{E}_0$, $\boldsymbol{E}_1$ e $\boldsymbol{E}'$ espaços normados completos, $A \subseteq E_0 \times E_1$ um aberto, $a = (a_0, a_1) \in A$, $m \in \mathbb{N} \smallsetminus \{0\}$ e $f\colon A \longrightarrow E'$ uma função $\mathscr{C}^m$-diferenciável. Se a diferencial parcial $\mathrm{D}_1 f|_a\colon E_1 \longrightarrow E'$ é invertível,*

então alguma vizinhança aberta $V \subseteq A$ de a, alguma vizinhança $U \subseteq E_0$ de a_0 e alguma função $\mathscr{C}^m$-diferenciável $\phi\colon U \longrightarrow E_1$ satisfazem

$$\operatorname{graf}(\phi) = V \cap f^{-1}(f(a))$$

e, para todo $x = (x_0, x_1) \in V$,

$$\mathrm{D}\phi|_{x_0} = -(\mathrm{D}_1 f|_x)^{-1} \circ \mathrm{D}_0 f|_x.$$

□ *Demonstração.* Definamos a função

$$\begin{aligned} g\colon A &\longrightarrow E_0 \times E' \\ (x_0, x_1) &\longmapsto (x_0, f(x_0, x_1)). \end{aligned}$$

A função g é $\mathscr{C}^m$-diferenciável e vale, para todo $x = (x_0, x_1) \in A$ e todo $u = (u_0, u_1) \in E_0 \times E_1$,

$$\mathrm{D}g|_x(u) = (u_0, \mathrm{D}f|_x(u)) = (u_0, \mathrm{D}_0 f|_x(u_0) + \mathrm{D}_1 f|_x(u_1)).$$

Em particular, para $x = a$,

$$\mathrm{D}g|_a(u) = (u_0, \mathrm{D}_0 f|_x(u_0) + \mathrm{D}_1 f|_x(u_1)),$$

o que implica que $\mathrm{D}g|_a$ é invertível, pois sua inversa é

$$(\mathrm{D}g|_a)^{-1}(u) = (u_0, (\mathrm{D}_1 f|_a)^{-1}(u_1 - \mathrm{D}_0 f|_a(u_0))).$$

Segue então do teorema da função inversa (19.89) que, para alguma vizinhança $V' \subseteq A$ de a, a função $g|_{V'}\colon V' \longrightarrow g(V')$ é um $\mathscr{C}^m$-difeomorfismo. Vamos restringir um pouco mais o conjunto V'. Como f é $\mathscr{C}^m$-diferenciável, a diferencial parcial $\mathrm{D}_1 f\colon A \longrightarrow \mathscr{L}(E_1, E')$ é uma função contínua. Ainda, como o conjunto $\overleftrightarrow{\mathscr{L}}(E_1, E')$ dos elementos invertíveis de $\mathscr{L}(E_1, E')$ é aberto e $\mathrm{D}_1 f|_a \in \overleftrightarrow{\mathscr{L}}(E_1, E')$, alguma vizinhança aberta $V'' \subseteq A$ de a satisfaz $\mathrm{D}_1 f(V'') \subseteq \overleftrightarrow{\mathscr{L}}(E_1, E')$ (basta tomar a imagem inversa pela função $\mathrm{D}_1 f$ de uma vizinhança aberta de $\mathrm{D}_1 f|_a$ em $\overleftrightarrow{\mathscr{L}}(E_1, E')$). Definindo $V := V' \cap V''$, segue que $V \subseteq A$ é aberto e $a \in V$. A restrição $g|_V\colon V \longrightarrow g(V)$ é um $\mathscr{C}^m$-difeomorfismo de $V \subseteq E_0 \times E_1$ para $g(V) \subseteq E_0 \times E'$.

Definamos $h := \mathrm{p}_1 \circ (g|_V)^{-1}\colon g(V) \longrightarrow E_1$, que é $\mathscr{C}^m$-diferenciável pois é composição de funções $\mathscr{C}^m$-diferenciáveis, e o aberto

$$U := \{x_0 \in E_0 \mid (x_0, f(a)) \in g(V)\}.$$

Definimos a inclusão (na notação de inclusão transladada)

$$\begin{aligned} \iota := \iota_{0,(0,f(a))}\colon E_0 &\longrightarrow E_0 \times E' \\ x_0 &\longmapsto (x_0, f(a)). \end{aligned}$$

Vale que $U = \iota^{-1}(g(V))$, o que mostra que $U \subseteq E_0$ é aberto, e que $a_0 \in U$. Finalmente, definimos a função $\phi := h \circ \iota \colon U \longrightarrow E_1$, que é $\mathscr{C}^m$-diferenciável pois é composição de funções $\mathscr{C}^m$-diferenciáveis. Mostremos que essa função satisfaz

$$\operatorname{graf}(\phi) = V \cap f^{-1}(f(a)).$$

- ($\subseteq$) Seja $x = (x_0, x_1) \in \operatorname{graf}(\phi)$. Então $x_0 \in U$ e $x_1 = \phi(x_0)$. A primeira relação implica que $\iota(x_0) = (x_0, f(a)) \in g(V)$, então, para algum $x' = (x_0', x_1') \in V$, vale que

  $$(x_0', f(x')) = g(x') = (x_0, f(a)) = \iota(x_0).$$

  Disso segue que $x_0' = x_0$ e $f(x') = f(a)$; segue portanto da segunda relação que

  $$x_1 = \phi(x_0) = h \circ \iota(x_0) = \mathrm{p}_1 \circ (g|_V)^{-1} \circ g(x') = x_1'.$$

  Assim, concluímos que $x = x' \in V$ e, finalmente, que

  $$f(a) = f(x') = f(x),$$

  o que mostra que $x \in V \cap f^{-1}(f(a))$.
- ($\supseteq$) Seja $x = (x_0, x_1) \in V \cap f^{-1}(f(a))$. Então $x \in V$ e $f(x) = f(a)$. Da primeira relação, segue que $g(x) \in g(V)$; da segunda relação, que

  $$\iota(x_0) = (x_0, f(a)) = (x_0, f(x)) = g(x).$$

  Isso implica que $x_0 \in \iota^{-1}(g(V)) = U$. Sendo assim, podemos calcular a imagem de x_0 sob ϕ, obtendo

  $$\phi(x_0) = h \circ \iota(x_0) = h \circ g(x) = \mathrm{p}_1(x) = x_1,$$

  portanto $x \in \operatorname{graf}(\phi)$.

Agora que está demonstrado que o gráfico da função ϕ é o conjunto de nível $f^{-1}(f(a))$ restrito a uma vizinhança de a, calculemos a diferencial de ϕ a partir das diferenciais parciais de f. Como para todo $x_0 \in U$ vale $f(x_0, \phi(x_0)) = f(a)$, e $f(a)$ é constante, diferenciando essa igualdade segue que

$$\mathrm{D}_0 f|_x + \mathrm{D}_1 f|_x \circ \mathrm{D}\phi|_{x_0} = 0.$$

Como para $x \in V$ a diferencial $\mathrm{D}_1 f|_x$ é invertível, segue que

$$\mathrm{D}\phi|_{x_0} = -(\mathrm{D}_1 f|_x)^{-1} \circ \mathrm{D}_0 f|_x. \qquad \blacksquare$$

19.7.9.4 Teorema do Posto

19.7.9.5 Extremos locais com restrições

19.8 Integração em espaço normados

19.8.1 Integração de trajetórias

19.8.1.1 Integração em intervalos

Consideraremos a partir de agora um intervalo compacto $I \subseteq \mathbb{R}$ e um espaço normado $\boldsymbol{E}$. O conjunto das trajetórias $\gamma\colon I \longrightarrow E$ limitadas, denotado $\mathscr{B}(I, E)$, é um espaço normado com a norma do supremo definida para $\gamma \in \mathscr{B}(I, E)$ por

$$\|\gamma\|_\infty = \sup_{t \in I} \|\gamma(t)\| .$$

Vamos considerar um subconjunto das funções limitadas, as funções constantes por partes. Para isso relembramos que uma n-partição por intervalos de $I = [a, a']$ é uma lista $(t_i)_{i \in [n+1]} \in \mathbb{R}^{n+1}$ tal que

$$a = t_0 < \cdots < t_n = a'$$

e que o conjunto dessas partições é denotado $\mathsf{P}_{[,]}([a, a'])$.

:⊢ **Definição 19.50.** Seja $I \subseteq \mathbb{R}$ um intervalo limitado com extremos inferior a e superior a'. O *comprimento* de I é

$$\mathrm{vol}^1(I) = a' - a.$$

Isso vale para qualquer um dos intervalos $]a, a'[$, $[a, a'[$, $]a, a']$ e $[a, a']$, com $a \leq a'$.

Note que dada uma partição por intervalos de I como acima, vale

$$\mathrm{vol}^1(I) = \mathop{\mathbf{+}}_{i \in [n]} (t_{i+1} - t_i).$$

:⊢ **Definição 19.51.** Sejam $I \subseteq \mathbb{R}$ um intervalo compacto e $\boldsymbol{E}$ um espaço normado. Uma *trajetória constante por partes*[19] de I para $\boldsymbol{E}$ é uma trajetória $\gamma\colon I \longrightarrow E$ para a qual alguma n-partição por intervalos $(t_i)_{i \in [n+1]} \in \mathsf{P}_{[,]}(I)$ satisfaz que, para todo $i \in [n]$, $\gamma|_{]t_i, t_{i+1}[}$ é constante. O conjunto das trajetórias constantes por partes de I para E é denotado $\mathscr{I}_c(I, E)$.

[19] Também chamada de função 'escada' ou função 'degrau'.

Claramente uma trajetória constante por partes definida em um intervalo compacto é limitada, pois sua imagem tem finitos valores.

Para qualquer trajetória constante por partes γ, existem $x_0, \ldots, x_{n-1} \in E$ e partição por intervalos $(t_i)_{i\in[n+1]} \in \mathsf{P}_{[,]}(I)$ tais que, para todo $i \in [n]$ e todo $t \in \,]t_i, t_{i+1}[$ vale $\gamma(t) = x_i$.

▷ **Exercício 19.18.** *Sejam $I \subseteq \mathbb{R}$ um intervalo compacto e $\boldsymbol{E}$ um espaço normado. O espaço das trajetórias constantes por partes $\mathscr{I}_c(I, E)$ é um subespaço linear de $\mathscr{B}(I, E)$.*

Sendo assim, podemos considerar o fecho de $\mathscr{I}_c(I, E)$ em $\mathscr{B}(I, E)$. As trajetórias γ pertencentes a esse fecho satisfazem que, para alguma sequência $(\gamma_n)_{n\in\mathbb{N}}$ de trajetórias constantes por partes,

$$\gamma = \lim_{n\to\infty} \gamma_n.$$

Nesse caso, consideramos um espaço normado completo.

:⊢ **Definição 19.52.** Seja $\boldsymbol{E}$ um espaço normado completo. Uma *trajetória regulada* em $\boldsymbol{E}$ é uma trajetória $\gamma\colon I \longrightarrow E$ que é limite de alguma sequência de trajetórias constantes por partes. O conjunto das trajetórias reguladas de I para E é denotado $\bar{\mathscr{I}}_c(I, E)$.

As trajetórias reguladas podem ser caracterizadas como as funções que têm pontos limites à esquerda e à direita.

▷ **Exercício 19.19.** *Sejam $I = [a, a'] \subseteq \mathbb{R}$ um intervalo compacto e $\boldsymbol{E}$ um espaço normado completo. Uma trajetória γ é regulada se, e somente se, para todo $t \in \,]a, a']$, γ tem ponto limite à esquerda (inferior) em t e, para todo $t \in [a, a'[$, γ tem ponto limite à direita (superior) em t.*

Como consequência, toda trajetória contínua é regulada. Agora vamos definir a integral de uma trajetória regulada. Antes definimos a integral para funções constantes por partes.

:⊢ **Definição 19.53.** Sejam $I \subseteq \mathbb{R}$ um intervalo compacto, $\boldsymbol{E}$ um espaço normado e $\gamma\colon I \longrightarrow E$ uma trajetória constante por partes com n-partição por por intervalos $(t_i)_{i\in[n+1]} \in \mathsf{P}_{[,]}(I)$ e $x_0, \ldots, x_{n-1} \in E$ tais que, para todo $i \in [n]$ e todo $t \in \,]t_i, t_{i+1}[$ vale $\gamma(t) = x_i$. A *integral de γ em I* é

$$\int_I \gamma := \mathop{+}_{i\in[n]} (t_{i+1} - t_i)x_i.$$

A *integração* em $\mathscr{I}_c(I, E)$ é a função

$$\begin{aligned} \textstyle\int_I \colon \mathscr{I}_c(I, E) &\longrightarrow E \\ \gamma &\longmapsto \int_I \gamma. \end{aligned}$$

Vale notar que a definição é independente da escolha de partição por intervalos para γ.

$\vdash$ **Proposição 19.91.** *Sejam $I \subseteq \mathbb{R}$ um intervalo compacto e $\boldsymbol{E}$ um espaço normado. A integração*

$$\textstyle\int_I \colon \mathscr{I}_c(I, E) \longrightarrow E$$

é linear e contínua e

$$\|\textstyle\int_I(\gamma)\| \leq \mathrm{vol}^1(I)\, \|\gamma\|_\infty .$$

$\square$ *Demonstração.* A linearidade é exercício. Para mostrar a continuidade, basta notar que

$$\|\textstyle\int_I(\gamma)\| \leq \displaystyle\mathop{\mathbin{+}}_{i\in[n]} (t_{i+1} - t_i)\, \|x_i\| \leq \mathrm{vol}^1(I)\, \|\gamma\|_\infty . \qquad \blacksquare$$

A desigualdade que prova a continuidade equivale a dizer que $\|\int_I\| \leq \mathrm{vol}^1(I)$ e se analisamos trajetórias constantes concluímos que de fato

$$\|\textstyle\int_I\| = \mathrm{vol}^1(I).$$

Agora, estendemos a definição para trajetórias reguladas. Consideramos um espaço normado completo $\boldsymbol{E}$. Seja $\gamma \colon I \longrightarrow E$ uma trajetória regulada e $(\gamma_n)_{n\in\mathbb{N}}$ uma sequência de trajetórias constantes por partes tal que $\gamma = \lim_{n\to\infty} \gamma_n$. Como $\int_I$ é linear e contínua, a sequência $(\int_I \gamma_n)_{n\in\mathbb{N}}$ é uma sequência aproximante em E e portanto, da completude do espaço, tem limite. Além disso, se $(\gamma'_n)_{n\in\mathbb{N}}$ é uma sequência de trajetórias constantes por partes tal que $\gamma = \lim_{n\to\infty} \gamma'_n$, então

$$\|\gamma'_n - \gamma_n\| \leq \|\gamma'_n - \gamma\| + \|\gamma - \gamma_n\| ,$$

portanto a sequência $(\gamma'_n - \gamma_n)$ converge para 0, o que implica que

$$\lim_{n\to\infty} \int_I \gamma_n = \lim_{n\to\infty} \int_I \gamma'_n.$$

Definimos então a integral de uma trajetória regulada como esse limite.

$:\vdash$ **Definição 19.54** (Integração)**.** Sejam $I \subseteq \mathbb{R}$ um intervalo compacto, $\boldsymbol{E}$ um espaço normado completo e $\gamma \colon I \longrightarrow E$ uma trajetória regulada e $(\gamma_n)_{n\in\mathbb{N}}$ uma sequência de trajetórias constantes por partes tais que $\gamma = \lim_{n\to\infty} \gamma_n$. A *integral de γ em I* é

$$\int_I \gamma := \lim_{n\to\infty} \int_I \gamma_n.$$

Caso seja necessário explicitar a variável da trajetória γ, denota-se

$$\int_I \gamma(t)\mathrm{d}t := \int_I \gamma.$$

A *integração* em $\bar{\mathscr{I}}_c(I,E)$ é a função

$$\begin{aligned}\textstyle\int_I \colon \bar{\mathscr{I}}_c(I,E) &\longrightarrow E\\ \gamma &\longmapsto \int_I \gamma.\end{aligned}$$

A integração $\int_I \colon \bar{\mathscr{I}}_c(I,E) \longrightarrow E$ é linear e contínua e

$$\left\|\textstyle\int_I(\gamma)\right\| \leq \operatorname{vol}^1(I)\,\|\gamma\|_\infty\,.$$

Isso ocorre pela unicidade da extensão de uma função em um subespaço linear para o fecho desse subespaço.

▷ **Exercício 19.20.** *Sejam $I \subseteq \mathbb{R}$ um intervalo compacto e $\boldsymbol{E}$ um espaço normado completo. Para toda trajetória regulada $\gamma \colon I \longrightarrow E$,*

$$\left\|\int_I \gamma(t)\mathrm{d}t\right\| \leq \int_I \|\gamma(t)\|\,\mathrm{d}t.$$

19.8.1.2 Trajetórias primitivas

Definimos uma primitiva de uma trajetória como uma trajetória cuja derivada é a trajetória original. Em seguida, mostraremos como usar a ferramenta da integração para calcular primitivas de uma trajetória.

:⊢ **Definição 19.55.** Sejam $I \subseteq \mathbb{R}$ um intervalo, $\boldsymbol{E}$ um espaço normado e $\gamma \colon I \longrightarrow E$ uma trajetória. Uma trajetória *primitiva de* γ é uma trajetória contínua $\Gamma \colon I \longrightarrow E$, diferenciável em I°, tal que $\dot{\Gamma} = \gamma$.

É fácil notar que somar um vetor constante a uma trajetória primitiva resulta em uma trajetória primitiva. É relevante notar que a recíproca também é válida. Isso ocorre pois as derivadas de ambas primitivas coincidem em um aberto conexo.

⊢ **Proposição 19.92.** *Sejam $I \subseteq \mathbb{R}$ um intervalo, $\boldsymbol{E}$ um espaço normado, $\gamma \colon I \longrightarrow E$ uma trajetória e $\Gamma_0 \colon I \longrightarrow E$ uma trajetória primitiva de γ. Uma trajetória $\Gamma \colon I \longrightarrow E$ é primitiva de γ se, e somente se, para algum $c \in E$, vale*

$$\Gamma = \Gamma_0 + c.$$

□ *Demonstração.*

- (⇒) Como Γ_0 e Γ são primitivas de γ, então, para todo $t \in I^\circ$

$$\dot{\partial}(\Gamma - \Gamma_0)(t) = \dot{\Gamma}(t) - \dot{\Gamma}_0(t) = \gamma(t) - \gamma(t) = 0.$$

  Assim, a diferencial de $\Gamma - \Gamma_0$ deve ser nula, portanto segue da proposição 19.67 que, para algum $c \in E$, $\Gamma - \Gamma_0 = c$ em I°. Como Γ_0 e Γ são contínuas, segue que vale $\Gamma - \Gamma_0 = c$ em I.
- (⇐) Basta notar que, para todo $t \in I^\circ$,

$$\dot{\Gamma}(t) = \dot{\partial}(\Gamma_0(t) + c) = \dot{\Gamma}_0(t) = \gamma(t),$$

  o que mostra que Γ é primitiva de γ. ■

19.8.1.3 Integração orientada

Para relacionar a integração com a trajetória primitiva e, portanto, com a derivação, devemos definir a integral orientada.

:⊢ **Definição 19.56** (Integração orientada)**.** Sejam $I \subseteq \mathbb{R}$ um intervalo compacto, $\boldsymbol{E}$ um espaço normado completo, $\gamma\colon I \longrightarrow E$ uma trajetória regulada e $t, t' \in I$ tais que $t \leq t'$. A *integral orientada de* γ *de* t *a* t' é

$$\int_t^{t'} \gamma := \int_{[t,t']} \gamma$$

e a *integral orientada de* γ *de* t' *a* t é

$$\int_{t'}^{t} \gamma := -\int_{[t,t']} \gamma.$$

Em particular, isso implica as seguintes propriedades.

▷ **Exercício 19.21.** *Sejam* $I \subseteq \mathbb{R}$ *um intervalo compacto,* $\boldsymbol{E}$ *um espaço normado completo e* $\gamma\colon I \longrightarrow E$ *uma trajetória regulada.*

1. *Para todo* $t \in I$*,*
$$\int_t^t \gamma = 0.$$
2. *Para todos* $t, t', t'' \in I$*,*
$$\int_t^{t''} \gamma = \int_t^{t'} \gamma + \int_{t'}^{t''} \gamma.$$

Usando a definição de integral orientada, podemos relacionar a noção de trajetória primitiva com a ferramenta de integração. Para isso, começamos com a seguinte proposição.

⊢ **Proposição 19.93.** *Sejam* $I \subseteq \mathbb{R}$ *um intervalo compacto,* $\boldsymbol{E}$ *um espaço normado completo,* $\gamma\colon I \longrightarrow E$ *uma trajetória regulada e* $t_0 \in I$*. A trajetória*

$$\begin{aligned} \Gamma\colon I &\longrightarrow E \\ t &\longmapsto \int_{t_0}^{t} \gamma \end{aligned}$$

é contínua. Para todo $t \in I^\circ$ *tal que* γ *é contínua em* t*, a trajetória* Γ *é derivável em* t *e*

$$\dot{\Gamma}(t) = \gamma(t).$$

□ *Demonstração.* Note que, para todo $t \in I$ e todo $\delta \in \mathbb{R}$ tal que $t + \delta \in I$, vale

$$\Gamma(t+\delta) - \Gamma(t) = \int_{t_0}^{t+\delta} \gamma - \int_{t_0}^{t} \gamma = \int_{t}^{t+\delta} \gamma,$$

portanto

$$\|\Gamma(t+\delta) - \Gamma(t)\| \leq |\delta| \sup_{t' \in [t, t+\delta]} \|\gamma(t')\| \leq |\delta| \, \|\gamma\|_\infty \,,$$

o que mostra que Γ é contínua.

Agora, seja $t \in I^\circ$ em que γ é contínua. Como

$$\frac{\Gamma(t+\delta) - \Gamma(t)}{\delta} - \gamma(t) = \frac{1}{\delta} \int_t^{t+\delta} \gamma(t') \mathrm{d}t' - \gamma(t) = \frac{1}{\delta} \int_t^{t+\delta} (\gamma(t') - \gamma(t)) \mathrm{d}t',$$

então

$$\left\| \frac{\Gamma(t+\delta) - \Gamma(t)}{\delta} - \gamma(t) \right\| \leq \sup_{t' \in [t, t+\delta]} \|\gamma(t') - \gamma(t)\| \,,$$

portanto segue da continuidade de γ que $\dot{\Gamma}(t) = \gamma(t)$. ■

A última proposição tem a seguinte consequência.

⊢ **Proposição 19.94** (Teorema Fundamental do Cálculo)**.** *Sejam $I \subseteq \mathbb{R}$ um intervalo compacto, $\boldsymbol{E}$ um espaço normado completo, $\gamma \colon I \longrightarrow E$ uma trajetória contínua.*

1. *Para todo t_0, a trajetória*

$$\begin{aligned} \Gamma_0 \colon I &\longrightarrow E \\ t &\longmapsto \int_{t_0}^{t} \gamma \end{aligned}$$

   é uma primitiva de γ.
2. *Para todos $a, a' \in I$ e toda primitiva $\Gamma \colon I \longrightarrow E$ de γ,*

$$\int_a^{a'} \gamma = \Gamma(a') - \Gamma(a).$$

□ *Demonstração.*
1. Segue direto de 19.93 pois toda trajetória contínua é regulada.
2. Seja $t_0 \in I$. Pelo item anterior, a trajetória

$$\begin{aligned} \Gamma_0 \colon I &\longrightarrow E \\ t &\longmapsto \int_{t_0}^{t} \gamma \end{aligned}$$

é uma primitiva de γ e, pela proposição 19.92, para algum $c \in E$ vale $\Gamma = \Gamma_0 + c$. Disso segue que

$$\Gamma_0(a') - \Gamma_0(a) = \Gamma_0(a') + c - \Gamma_0(a) - c = \Gamma(a') - \Gamma(a),$$

portanto

$$\int_a^{a'} \gamma = \int_{t_0}^{a'} \gamma - \int_{t_0}^{a} \gamma = \Gamma_0(a') - \Gamma_0(a) = \Gamma(a') - \Gamma(a). \qquad \blacksquare$$

19.8.2 Integração de campos vetoriais

Nesta seção estudaremos campos vetoriais, que devem ser imaginados como um conjunto de vetores, cada um preso a um ponto de um aberto do espaço normado. A definição é dada a seguir.

:⊢ **Definição 19.57.** Sejam $\boldsymbol{E}$ um espaço normado e $A \subseteq E$ um aberto. Um *campo vetorial* em A é uma função $v\colon A \longrightarrow E$. Para todo $x \in A$, o *vetor de v tangente a A em x* é

$$v|_x := v(x).$$

Na definição, não exigimos que a função seja contínua, mas em geral consideraremos campos vetoriais que são contínuos ou com alguma regularidade ainda maior, como controlado ou diferenciável.

Além dessa definição, poderíamos considerar também campos vetoriais que variam no tempo, ou seja, uma família de campos vetoriais indexados por um intervalo de números reais.

:⊢ **Definição 19.58.** Sejam $\boldsymbol{E}$ um espaço normado e $A \subseteq E$ um aberto e $I \subseteq \mathbb{R}$ um intervalo aberto. Um *campo vetorial temporal* em $I \times A$ é uma função $v\colon I \times A \longrightarrow E$. Para todo $(t, x) \in I \times A$, o *vetor de v tangente a A em (t, x)* é

$$v_t|_x := v(t, x).$$

Focaremos no estudos dos campos vetoriais e não nos campos vetoriais temporais.

19.8.2.1 Trajetórias integrais

:⊢ **Definição 19.59.** Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto e $v\colon A \longrightarrow E$ um campo vetorial em A. Uma *trajetória integral* de v é uma trajetória derivável $\phi\colon I \longrightarrow A$ tal que I é aberto e, para todo $t \in I$,

$$\dot{\phi}(t) = v|_{\phi(t)}.$$

Uma *trajetória integral maximal* de v é uma trajetória integral $\phi\colon I \longrightarrow A$ de v tal que, para todo intervalo aberto $I' \subseteq \mathbb{R}$ tal que $I \subset I'$, não existe trajetória integral $\phi'\colon I' \longrightarrow A$ de v tal que $\phi'|_I = \phi$.

Isso significa que a trajetória ϕ satisfaz $\dot{\phi} = v \circ \phi$. Isso também pode ser compreendido com a linguagem de equações diferenciais. Uma trajetória integral é uma solução da equação diferencial ordinária

$$\dot{x} = v(x).$$

▷ **Exercício 19.22.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $v\colon A \longrightarrow E$ *um campo vetorial em* A *e* $\phi\colon I \longrightarrow A$ *uma trajetória integral de* v*. Se* v *é* $\mathscr{C}^m$*-diferenciável, então* ϕ *é* $\mathscr{C}^{m+1}$*-diferenciável.*

Podemos também descrever uma trajetória integral em termos de uma integral, o que justifica seu nome.

⊢ **Proposição 19.95.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $I \subseteq \mathbb{R}$ *um intervalo aberto,* $A \subseteq E$ *um aberto,* $v\colon A \longrightarrow E$ *um campo vetorial contínuo,* $(t_0, x_0) \in I \times A$ *e* $\phi\colon I \longrightarrow A$ *uma trajetória diferenciável. A trajetória* ϕ *é uma trajetória integral de* v *tal que* $\phi(t_0) = x_0$ *se, e somente se, para todo* $t \in I$*,*

$$\phi(t) = x_0 + \int_{t_0}^{t} v|_{\phi(s)} \mathrm{d}s.$$

□ *Demonstração.* Notemos primeiro que, como o campo v e a trajetória ϕ são contínuos, a composição $v \circ \phi$ é uma trajetória contínua e portanto integrável.

- (⇒) Suponhamos que ϕ é uma trajetória integral de v tal que $\phi(t_0) = x_0$. Nesse caso, para todo $t \in I$ vale que $\dot{\phi}(t) = v|_{\phi(t)}$, o que significa que ϕ é uma primitiva de $v \circ \phi$, portanto segue da parte 2 do teorema fundamental do cálculo (19.94) que

$$\phi(t) = \phi(t_0) + \phi(t) - \phi(t_0) = x_0 + \int_{t_0}^{t} v|_{\phi(s)} \mathrm{d}s.$$

- (⇐) Suponhamos que ϕ satisfaz a equação do enunciado. Então, pela parte 1 do teorema fundamental do cálculo (19.94), ϕ é uma primitiva de $v \circ \phi$, portanto segue que, para todo $t \in I$,

$$\dot{\phi}(t) = v|_{\phi(t)}.$$

  Além disso, vale que

$$\phi(t_0) = x_0 + \int_{t_0}^{t_0} v|_{\phi(s)} \mathrm{d}s = x_0. \qquad \blacksquare$$

A partir de uma trajetória integral, podemos achar outras trajetórias através de uma mudança de variável por translação.

⊢ **Proposição 19.96** (Reparametrização de trajetória)**.**

Sejam $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $v\colon A \longrightarrow E$ *um campo vetorial em* A *e* $\phi\colon I \longrightarrow A$ *uma trajetória integral de* v*.*

1. *(Translação) Para todo* $c \in \mathbb{R}$*, a trajetória*

$$\begin{aligned} \phi \circ T_c\colon I - c &\longrightarrow A \\ t &\longmapsto \phi(t+c) \end{aligned}$$

   é uma trajetória integral de v*. Se* ϕ *é maximal, então* $\phi \circ T_c$ *é maximal.*
2. *(Dilatação) Para todo* $c \in \mathbb{R} \smallsetminus \{0\}$*, a trajetória*

$$\begin{aligned} \phi \circ E_c\colon c^{-1}I &\longrightarrow A \\ t &\longmapsto \phi(ct) \end{aligned}$$

   é uma trajetória integral de cv*. Se* ϕ *é maximal, então* $\phi \circ E_c$ *é maximal. A trajetória*

$$\begin{aligned} \phi \circ E_0\colon \mathbb{R} &\longrightarrow A \\ t &\longmapsto \phi(0t) \end{aligned}$$

   é uma trajetória integral de $0v$*. Se* ϕ *é maximal, então* $\phi \circ E_0$ *é maximal.*

□ *Demonstração.* 1. Como $\dot{T}_c(t) = 1$, segue da regra da cadeia para trajetórias reparametrizadas (19.12) que

$$\dot{\partial}(\phi \circ T_c)(t) = \dot{T}_c(t)\dot{\phi}(T_c(t)) = \dot{\phi}(t+c) = v|_{\phi(t+c)} = v|_{\phi \circ T_c(t)}.$$

Suponhamos agora que ϕ é maximal. Por absurdo, suponhamos que $\phi \circ T_c$ não fosse maximal. Nesse caso, alguma trajetória integral $\phi'\colon I' \longrightarrow E$ do campo v satisfaria que $(I - c) \subset I'$ e $\phi'|_{I-c} = \phi \circ T_c$. Nesse caso, valeria que $I \subset (I' + c)$ e, para todo $t \in I$,

$$\phi' \circ T_{-c}(t) = \phi \circ T_c \circ T_{-c}(t) = \phi(t),$$

o que contradiria a maximalidade de ϕ.

2. Para $c = 0$, o campo $0v$ é nulo e a trajetória $\phi \circ E_0$, é constante, portanto a afirmação é válida. Consideremos o caso em que $c \in \mathbb{R} \smallsetminus \{0\}$.
Como $\dot{E}_c(t) = c$, segue da regra da cadeia para trajetórias reparametrizadas (19.12) que

$$\dot{\partial}(\phi \circ E_c)(t) = \dot{E}_c(t)\dot{\phi}(E_c(t)) = c\dot{\phi}(ct) = cv|_{\phi(ct)} = (cv)|_{\phi \circ E_c(t)}.$$

Suponhamos agora que ϕ é maximal. Por absurdo, suponhamos que $\phi \circ E_c$ não fosse maximal. Nesse caso, alguma trajetória integral $\phi' \colon I' \longrightarrow E$ do campo cv satisfaria que $(c^{-1}I) \subset I'$ e $\phi'|_{c^{-1}I} = \phi \circ E_c$. Nesse caso, valeria que $I \subset (cI')$ e, para todo $t \in I$,

$$\phi' \circ E_{c^{-1}}(t) = \phi \circ E_c \circ E_{c^{-1}}(t) = \phi(t),$$

o que contradiria a maximalidade de ϕ. ■

19.8.2.2 Existência e unicidade de trajetórias integrais

Para mostrar a existência e unicidade de trajetórias integrais de um campo vetorial $v \colon A \longrightarrow E$, vamos supor que ele seja localmente controlado, ou seja, que para todo ponto $x_0 \in A$, o campo v restrito a alguma vizinhança $V \subseteq A$ de x é controlado: para algum $c \in \mathbb{R}_{\geq 0}$, todos $x, x' \in V$ satisfazem

$$\|v|_{x'} - v|_x\| \leq c \|x' - x\| .$$

Isso em particular vale para campos $\mathscr{C}^1$-diferenciáveis pois, como $\mathrm{D}v$ é contínua, é limitada em uma bola aberta, então pela proposição 19.65 segue que v é controlado na bola aberta, portanto localmente controlado.

Para a proposição seguinte, lembremos que $\mathrm{B}(x_0; r)$ denota a bola aberta de centro $x_0 \in A$ e raio $r \in \mathbb{R}_{>0}$ e que $\mathrm{I}(t_0; \varepsilon) =]t_0 - \varepsilon, t_0 + \varepsilon[$ denota o intervalo aberto de centro $t_0 \in \mathbb{R}$ e raio $\varepsilon \in \mathbb{R}_{>0}$.

⊢ **Proposição 19.97** (Existência e unicidade local de trajetória integral)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado completo,* $A \subseteq E$ *um aberto,* $v \colon A \longrightarrow E$ *um campo vetorial localmente controlado e* $(t_0, x_0) \in \mathbb{R} \times A$*. Para algum* $r \in \mathbb{R}_{>0}$ *tal que* $\mathrm{B}(x_0; r) \subseteq A$ *e algum* $\varepsilon \in \mathbb{R}_{>0}$*,*

1. *Alguma trajetória integral* $\phi \colon \mathrm{I}(t_0; \varepsilon) \longrightarrow \mathrm{B}(x_0; r)$ *de* v *satisfaz* $\phi(t_0) = x_0$*;*
2. *Se* $\phi' \colon \mathrm{I}(t_0; \varepsilon) \longrightarrow \mathrm{B}(x_0; r)$ *é uma trajetória integral de* v *tal que* $\phi'(t_0) = x_0$*, então* $\phi' = \phi$*.*

□ *Demonstração.* Como v é localmente controlado, para algum $r \in \mathbb{R}_{>0}$ o campo v é controlado em $\bar{\mathrm{B}}(x_0; r)$. Denotemos $B := \mathrm{B}(x_0; r)$. Podemos então considerar a constante de distorção $c := \langle\!\langle v|_{\overline{B}} \rangle\!\rangle$, que satisfaz, para todos $x, x' \in \overline{B}$,

$$\|v|_{x'} - v|_x\| \leq c \|x' - x\| .$$

Como a bola $\overline{B}$ é limitada, isso implica que v é limitado em $\overline{B}$ e podemos considerar a constante $l := \|v|_{\overline{B}}\|_\infty$, que satisfaz, para todo $x \in \overline{B}$,

$$\|v|_x\| \leq l.$$

Seja $\varepsilon \in \mathbb{R}_{>0}$ tal que $l\varepsilon < r$. Denotemos $I := \mathrm{I}(t_0; \varepsilon)$.

Consideremos o espaço linear $\mathscr{C}(\overline{I}, \boldsymbol{E})$ das trajetórias contínuas de $\overline{I}$ para E e a norma do supremo $\|\gamma\|_\infty = \sup_{t\in\overline{I}} \|\gamma(t)\|$. Esse espaço é um espaço normado completo (19.11). Definamos o conjunto

$$X := \left\{\gamma \in \mathscr{C}(\overline{I}, \boldsymbol{E}) \;\middle|\; \gamma(\overline{I}) \subseteq \overline{B},\ \gamma(t_0) = x_0\right\}.$$

Esse conjunto é fechado, não-vazio e é um subespaço métrico completo de $\mathscr{C}(\overline{I}, \boldsymbol{E})$ no qual a distância entre duas trajetórias é dada pela norma da diferença das trajetórias. Definimos a função

$$\begin{aligned} F\colon X &\longrightarrow \mathscr{C}(\overline{I}, \boldsymbol{E}) \\ \gamma &\longmapsto F(\gamma)\colon \overline{I} \longrightarrow E \\ &\qquad\qquad t \longmapsto F(\gamma)(t) := x_0 + \int_{t_0}^{t} v|_{\gamma(s)} \mathrm{d}s. \end{aligned}$$

Note que $F(X) \subseteq X$, pois, para toda $\gamma \in X$, vale $F(\gamma)(t_0) = x_0$ e, além disso, para todo $t \in \overline{I}$,

$$\|F(\gamma)(t) - x_0\| = \left\|\int_{t_0}^{t} v|_{\gamma(s)} \mathrm{d}s\right\| \leq \left|\int_{t_0}^{t} \left\|v|_{\gamma(s)}\right\| \mathrm{d}s\right| \leq l\,|t - t_0| \leq l\varepsilon < r,$$

o que mostra que $F(\gamma)(\overline{I}) \subseteq B$. Assim concluímos que $F(\gamma) \in X$, portanto $F\colon X \longrightarrow X$.

Seja $n \in \mathbb{N}$. Mostraremos por indução em n que, para todos $\gamma, \gamma' \in X$ e todo $t \in \overline{I}$,

$$\|F^n(\gamma')(t) - F^n(\gamma)(t)\| \leq \frac{c^n\,|t - t_0|^n}{n!} \|\gamma' - \gamma\|_\infty.$$

Para $n = 0$,

$$\left\|F^0(\gamma')(t) - F^0(\gamma)(t)\right\| = \|\gamma'(t) - \gamma(t)\| \leq \|\gamma' - \gamma\|_\infty = \frac{c^0\,|t - t_0|^0}{0!} \|\gamma' - \gamma\|_\infty.$$

Agora, suponhamos para o passo indutivo que a inequação vale para $n \in \mathbb{N}$. Então,usando a hipótese de indução, segue que

$$\begin{aligned} \left\|F^{n+1}(\gamma')(t) - F^{n+1}(\gamma)(t)\right\| &= \left\|\int_{t_0}^{t} v|_{F^n(\gamma')(s)} - v|_{F^n(\gamma)(s)} \mathrm{d}s\right\| \\ &\leq \left|\int_{t_0}^{t} \left\|v|_{F^n(\gamma')(s)} - v|_{F^n(\gamma)(s)}\right\| \mathrm{d}s\right| \\ &\leq c\left|\int_{t_0}^{t} \|F^n(\gamma')(s) - F^n(\gamma)(s)\| \,\mathrm{d}s\right| \\ &\leq c\left|\int_{t_0}^{t} \frac{c^n\,|s - t_0|^n}{n!} \|\gamma' - \gamma\|_\infty \,\mathrm{d}s\right| \\ &= \frac{c^{n+1}\,|t - t_0|^{n+1}}{(n+1)!} \|\gamma' - \gamma\|_\infty. \end{aligned}$$

Disso segue que, para todos $\gamma, \gamma' \in X$,

$$\|F^n(\gamma') - F^n(\gamma)\|_\infty \leq \frac{(c\varepsilon)^n}{n!} \|\gamma' - \gamma\|_\infty .$$

Isso mostra que F é contínua, pois é controlada. Além disso, como $\frac{(c\varepsilon)^n}{n!}$ é um termo da série exponencial de $\mathrm{e}^{c\varepsilon}$, deve valer que, para algum $n \in \mathbb{N} \smallsetminus \{0\}$,

$$\frac{(c\varepsilon)^n}{n!} < 1,$$

o que implica que F^n é uma contração. Segue então da proposição 16.46 que F tem único ponto fixo $\bar{\phi}$. Por definição de F, essa trajetória satisfaz, para todo $t \in \overline{I}$,

$$\bar{\phi}(t) = F(\bar{\phi})(t) = x_0 + \int_{t_0}^{t} v|_{\bar{\phi}(s)} \mathrm{d}s.$$

Definimos a restrição $\phi := \bar{\phi}|_I$. A trajetória ϕ então é uma trajetória integral de v tal que $\phi(t_0) = x_0$ (19.95) e cuja imagem está contida em B, pois

$$\phi(I) \subseteq \bar{\phi}(\overline{I}) = F(\bar{\phi})(\overline{I}) \subseteq B.$$

Devemos agora mostrar a unicidade de ϕ, pois temos somente a unicidade de $\bar{\phi}$. Suponha que $\phi' \colon I \longrightarrow B$ seja uma trajetória integral de v tal que $\phi'(t_0) = x_0$. Para todo $t \in I$, vale que $\dot{\phi}'(t) = v|_{\phi'(t)}$, logo pela desigualdade do valor médio (19.64) segue que, para todos $t, t' \in I$ tais que $t < t'$,

$$\|\phi'(t') - \phi'(t)\| \leq \sup_{s \in]t,t'[} \left\|\dot{\phi}'(s)\right\| |t' - t| = \sup_{s \in]t,t'[} \left\|v|_{\phi'(s)}\right\| |t' - t| \leq l(t' - t).$$

Isso implica que ϕ' pode ser estendida para uma trajetória contínua $\bar{\phi}' \colon \overline{I} \longrightarrow A$. Ainda, pelo critério da proposição 19.95 segue que

$$\phi'(t) = x_0 + \int_{t_0}^{t} v|_{\phi'(s)} \mathrm{d}s.$$

Então, pela continuidade de $\bar{\phi}'$ segue que

$$\begin{aligned}
\bar{\phi}'(t_0 + \varepsilon) &= \lim_{t \to \varepsilon} \phi(t_0 + t) \\
&= x_0 + \lim_{t \to \varepsilon} \int_{t_0}^{t_0 + t} v|_{\phi'(s)} \mathrm{d}s \\
&= x_0 + \int_{t_0}^{t_0 + \varepsilon} v|_{\phi'(s)} \mathrm{d}s \\
&= F(\bar{\phi}')(t_0 + \varepsilon)
\end{aligned}$$

e, analogamente, que

$$\bar{\phi}'(t_0 - \varepsilon) = x_0 + \int_{t_0}^{t_0-\varepsilon} v|_{\phi'(s)} \mathrm{d}s = F(\bar{\phi}')(t_0 - \varepsilon),$$

o que implica que para todo $t \in \bar{I}$ vale $\bar{\phi}'(t) = F(\bar{\phi}')(t)$, ou seja, que $\bar{\phi}'$ é ponto fixo de F. Assim concluímos que $\bar{\phi}' = \bar{\phi}$, logo $\phi' = \phi$. ■

Uma implicação direta é a unicidade das trajetórias integrais.

⊢ **Proposição 19.98** (Unicidade de trajetória integral)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado completo,* $A \subseteq E$ *um aberto,* $v\colon A \longrightarrow E$ *um campo vetorial localmente controlado e* $\phi, \phi'\colon I \longrightarrow A$ *trajetórias integrais de* v*. Se para algum* $t_0 \in I$ *vale que* $\phi(t_0) = \phi'(t_0)$*, então* $\phi = \phi'$*.*

□ *Demonstração.* Seja $I' := \{t \in I \mid \phi(t) = \phi'(t)\}$. Por hipótese $t_0 \in I'$. Mostraremos que I' é aberto e fechado em I, o que implica que $I' = I$, pois I é conexo e $I' \neq$. Primeiro notemos que $I' = (\phi' - \phi)^{-1}(0)$. Como ϕ e ϕ' são contínuas e $\{0\}$ é fechado, segue que I' é fechado. Agora, seja $t \in I'$. Pela existência e unicidade locais da proposição 19.97 vale que, para algum $\varepsilon \in \mathbb{R}_{>0}$ as trajetórias ϕ e ϕ' coincidem em $J := I \cap \mathrm{I}\,(t; \varepsilon)$; isto é, $\phi|_J = \phi'|_J$, ou seja, $J \subseteq I'$. Isso mostra que I' é aberto em I, portanto que $I = I'$. ■

⊢ **Proposição 19.99** (Existência e unicidade de trajetória integral maximal)**.** *Sejam* $\boldsymbol{E}$ *um espaço normado completo,* $A \subseteq E$ *um aberto,* $v\colon A \longrightarrow E$ *um campo vetorial localmente controlado e* $(t_0, x_0) \in \mathbb{R} \times A$*. Uma única trajetória integral maximal* $\phi\colon I \longrightarrow E$ *de* v *satisfaz* $\phi(t_0) = x_0$*.*

□ *Demonstração.* Seja $\mathcal{I}$ o conjunto dos intervalos abertos $I' \subseteq \mathbb{R}$ tais que $t_0 \in I'$ e uma única trajetória integral $\phi'\colon I' \longrightarrow E$ de v satisfaz $\phi'(t_0) = x_0$. Pela proposição 19.97, esse conjunto não é vazio. Seja I a união de todos os intervalos de $\mathcal{I}$. O conjunto I é um intervalo aberto não vazio tal que $t_0 \in I$. Definamos a trajetória $\phi\colon I \longrightarrow E$, para $t \in I'$, por $\phi(t) := \phi'(t)$. Pela proposição 19.98 a trajetória ϕ está bem definida em cada I'. Por definição de I, a trajetória ϕ é maximal e $\phi(t_0) = x_0$. Seja $\psi\colon J \longrightarrow E$ uma trajetória integral maximal de v tal que $\psi(t_0) = x_0$. Por definição de I, vale $J \subseteq I$ e, como ψ é maximal, $J = I$. Pela proposição 19.98, segue que $\phi = \psi$. ■

19.8.2.3 Fluxo de um campo vetorial

Pela proposição 19.99, dado um campo vetorial localmente controlado $v\colon A \longrightarrow E$ e $(t_0, x_0) \in \mathbb{R} \times A$, uma única trajetória integral maximal $\phi\colon I \longrightarrow A$ de v satisfaz $\phi(t_0) = x_0$. Podemos portanto definir trajetórias maximais e intervalos para cada par (t_0, x_0).

:⊢ **Definição 19.60.** Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto, $v\colon A \longrightarrow E$ um campo vetorial localmente controlado e $(t_0, x_0) \in \mathbb{R} \times A$. A *trajetória integral maximal de v passando por x_0 em t_0* é a trajetória integral maximal $\phi\colon I \longrightarrow E$ de v que satisfaz $\phi(t_0) = x_0$. Denota-se $I^v_{(t_0,x_0)} := I$ e, para todo $t \in I^v_{(t_0,x_0)}$,

$$\Phi^v(t, t_0, x_0) := \phi(t).$$

O *domínio do fluxo total de v* é o conjunto

$$\operatorname{dom}(\Phi^v) := \bigcup_{(t_0,x_0)\in\mathbb{R}\times A} I^v_{(t_0,x_0)} \times \{t_0\} \times \{x_0\} \subseteq \mathbb{R} \times \mathbb{R} \times A.$$

e o *fluxo total de v* é a função

$$\begin{aligned} \Phi^v\colon \operatorname{dom}(\Phi^v) &\longrightarrow A \\ (t, t_0, x_0) &\longmapsto \Phi^v(t, t_0, x_0). \end{aligned}$$

Consideremos para a proposição seguinte um espaço normado completo $\boldsymbol{E}$ e uma transformação linear contínua $L \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E})$. Essa transformação é um campo vetorial em E. Esse campo vetorial é controlado, pois a continuidade significa que, para todos $x, x' \in E$,

$$\|L(x') - L(x)\| \leq \|L\| \, \|x' - x\| \, .$$

⊢ **Proposição 19.100.** *Sejam $\boldsymbol{E}$ um espaço normado completo e $L \in \mathscr{L}(\boldsymbol{E}, \boldsymbol{E})$ uma transformação linear. O fluxo total de L é*

$$\begin{aligned} \Phi^L\colon \mathbb{R} \times \mathbb{R} \times E &\longrightarrow E \\ (t, t_0, x_0) &\longmapsto \mathrm{e}^{(t-t_0)L}(x_0). \end{aligned}$$

□ *Demonstração.* Definamos, para todo $t \in \mathbb{R}$, $\gamma(t) := \mathrm{e}^{tL}$. Pela proposição 19.50,

$$\dot{\gamma}(t) = \mathrm{e}^{tL} \circ L = L \circ \mathrm{e}^{tL}.$$

Agora, seja $x_0 \in E$ e considere

$$\begin{aligned} \phi\colon \mathbb{R} &\longrightarrow E \\ t &\longmapsto \mathrm{e}^{tL}(x_0). \end{aligned}$$

Essa trajetória é a trajetória maximal de L passando por x_0 em $t_0 = 0$. Claramente $\phi(0) = \mathrm{e}^{0L}x_0 = x_0$. Além disso, $\phi(t) = \gamma(t)(x_0)$, portanto segue do resultado anterior que

$$\dot{\phi}(t) = \dot{\partial}(\gamma(t)(x_0)) = \dot{\gamma}(t)(x_0) = L \circ \mathrm{e}^{tL}(x_0) = L(\phi(t)),$$

o que mostra que ϕ é trajetória integral de L. A trajetória ϕ é maximal pois está definida em $\mathbb{R}$.

Por fim, para todo $t_0 \in \mathbb{R}$, segue da proposição 19.96 que a trajetória maximal de L passando por x_0 em t_0 é

$$\phi \circ T_{-t_0}(t) = \mathrm{e}^{(t-t_0)L}(x_0). \qquad \blacksquare$$

Algumas propriedades do fluxo total seguem facilmente.

$\vdash$ **Proposição 19.101.** *Sejam* $\boldsymbol{E}$ *um espaço normado completo,* $A \subseteq E$ *um aberto e* $v\colon A \longrightarrow E$ *um campo vetorial localmente controlado.*

1. *Para todo* $(t_0, x_0) \in \mathbb{R} \times A$, $(t_0, t_0, x_0) \in \operatorname{dom}(\Phi^v)$ *e*

$$\Phi^v(t_0, t_0, x_0) = x_0;$$

2. *Para todo* $(t_1, t_0, x_0) \in \operatorname{dom}(\Phi^v)$ *e todo* $t \in \mathbb{R}$, *vale que* $(t, t_0, x_0) \in \operatorname{dom}(\Phi^v)$ *se, e somente se,* $(t, t_1, \Phi^v(t_1, t_0, x_0)) \in \operatorname{dom}(\Phi^v)$*; nesse caso,*

$$\Phi^v(t, t_0, x_0) = \Phi^v(t, t_1, \Phi^v(t_1, t_0, x_0));$$

3. *Para todo* $(t_1, t_0, x_0) \in \operatorname{dom}(\Phi^v)$, *vale que* $(t_0, t_1, \Phi^v(t_1, t_0, x_0)) \in \operatorname{dom}(\Phi^v)$ *e*

$$\Phi^v(t_0, t_1, \Phi^v(t_1, t_0, x_0)) = x_0;$$

4. *Para todo* $x_0 \in A$ *e todos* $t_0, t_1, t \in \mathbb{R}$, *vale que* $(t, t_1, x_0) \in \operatorname{dom}(\Phi^v)$ *se, e somente se,* $(t - (t_1 - t_0), t_0, x_0) \in \operatorname{dom}(\Phi^v)$*; nesse caso,*

$$\Phi^v(t, t_1, x_0) = \Phi^v(t - (t_1 - t_0), t_0, x_0);$$

5. *Para todos* $c \in \mathbb{R}$, $(t_0, x_0) \in \mathbb{R} \times A$ *e* $t \in \mathbb{R}$, *vale que* $(t, t_0, x_0) \in \operatorname{dom}(\Phi^{cv})$ *se, e somente se,* $(ct, ct_0, x_0) \in \operatorname{dom}(\Phi^v)$*; nesse caso,*

$$\Phi^{cv}(t, t_0, x_0) = \Phi^v(ct, ct_0, x_0).$$

6. *Para todos* $(t_0, x_0) \in \mathbb{R} \times A$ *e todo* $t \in \mathbb{R}$, *vale que* $(t, t_0, x_0) \in \operatorname{dom}(\Phi^v)$ *se, e somente se,* $(1, 0, x_0) \in \operatorname{dom}(\Phi^{(t-t_0)v})$*; nesse caso,*

$$\Phi^{(t-t_0)v}(1, 0, x_0) = \Phi^v(t, t_0, x_0).$$

$\square$ *Demonstração.* Segue direto da definição de fluxo pela unicidade das trajetórias maximais (19.99) e das proposições de reparametrização de trajetórias integrais por translação e dilatação (19.96).

$\blacksquare$

Pelo item 6 da proposição anterior, podemos definir o fluxo (local) de v.

:⊢ **Definição 19.61.** Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto e $v\colon A \longrightarrow E$ um campo vetorial localmente controlado. O *domínio do fluxo* de v é o conjunto

$$\operatorname{dom}(\phi^v) := \bigcup_{x \in A} I^v_{(0,x)} \times \{x\} \subseteq \mathbb{R} \times A.$$

O *fluxo* de v é a função

$$\begin{aligned} \phi^v\colon \operatorname{dom}(\phi^v) &\longrightarrow E \\ (t,x) &\longmapsto \phi^{(t)v}(x) := \Phi^{tv}(1,0,x). \end{aligned}$$

Para todo $x \in A$, o *intervalo maximal da trajetória integral* de v por x é $I^v_x := I^v_{(0,x)}$ a *trajetória integral maximal de v por x* é a função

$$\begin{aligned} \phi^v_x\colon I^v_x &\longrightarrow E \\ t &\longmapsto \phi^{(t)v}_x := \phi^{(t)v}(x). \end{aligned}$$

Denota-se $\phi^{tv}(x)$ por simplicidade. Pelos comentários anteriores, essa definição nos dá as seguintes relações.

$$\begin{aligned} \phi^{(t-t_0)v}(x_0) &= \Phi^{(t-t_0)v}(1,0,x_0) \\ &= \Phi^v(t-t_0,0,x_0) \\ &= \Phi^v(t,t_0,x_0). \end{aligned}$$

Valem propriedades análogas às do fluxo total de v.

⊢ **Proposição 19.102.** *Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto e $v\colon A \longrightarrow E$ um campo vetorial localmente controlado.*

1. *Para todo $x \in A$, vale que $0 \in I^v_x$ e*

$$\phi^{0v}(x) = x.$$

2. *Para todo $x \in A$, todo $t \in I^v_x$ e todo $t' \in I^v_{\phi^{tv}(x)}$, vale que $(t'+t) \in I^v_x$ e*

$$\phi^{(t'+t)v}(x) = \phi^{t'v}(\phi^{tv}(x)).$$

3. *Para todo $x \in A$ e todo $t \in I^v_x$, vale que $-t \in I^v_{\phi^{tv}(x)}$ e*

$$\phi^{-tv}(\phi^{tv}(x)) = x.$$

4. *Para todo $x \in A$ e todos $t_0, t, t'_0, t' \in \mathbb{R}$ tais que $t - t_0 = t' - t'_0$ e $t - t_0 \in I^v_x$,*

$$\phi^{(t-t_0)v}(x) = \phi^{(t'-t'_0)v}(x).$$

5. *Para todo $x \in A$ e todos $t, t' \in \mathbb{R}$ tais que $t, t' \in I_x^v$, vale que $t' + t \in I_x^v$ se, e somente se, $1 \in I_x^{t'v+tv}$; nesse caso,*
$$\phi^{t'v+tv}(x) = \phi^{(t'+t)v}(x);$$
6. *Para todo $x \in A$ e todos $t, t' \in \mathbb{R}$, vale que $t' \in I_x^{tv}$ se, e somente se, $t't \in I_x^v$; nesse caso,*
$$\phi^{t'(tv)}(x) = \phi^{(t't)v}(x).$$

□ *Demonstração.* Segue direto de 19.101. ■

A notação ϕ_x^{tv} é boa pois ela compartilha com a exponencial o seguinte comportamento:
$$\dot{\phi}_x^{tv} = v|_{\phi_x^{tv}} = v \circ \phi_x^{tv},$$
ou seja, quando derivamos o fluxo de um campo v, a "constante" v cai "multiplicando" o fluxo.

19.8.2.4 Propriedades das condições iniciais

A primeira proposição que apresentamos diz como estimar a distância entre dois pontos transportados pelo fluxo com respeito à distância original dos pontos. Ela afirma que o fluxo tem um comportamento controlado e sua distorção é majorada pela exponencial da distorção do campo.

⊢ **Proposição 19.103.** *Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto, $v\colon A \longrightarrow E$ um campo vetorial controlado. Para todos $x, x' \in A$ e todo $t \in I_x^v \cap I_{x'}^v$*
$$\left\|\phi^{tv}(x') - \phi^{tv}(x)\right\| \leq \mathrm{e}^{|t|\langle\langle v\rangle\rangle} \left\|x' - x\right\|.$$

Podemos pensar que, como comentado anteriormente, esse resultado mostra que o fluxo é controlado e satisfaz
$$\langle\langle \phi^{tv} \rangle\rangle \leq \mathrm{e}^{\langle\langle tv \rangle\rangle}.$$
No entanto, isso não é tecnicamente correto porque o fluxo ϕ^{tv} não necessariamente está definido para todos $x, x' \in A$ quando mantemos um t fixo.

⊢ **Proposição 19.104.** *Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto, $v\colon A \longrightarrow E$ um campo vetorial controlado. Sejam $x_0 \in A$ e $t_0 \in \mathbb{R}$ e $I \subseteq I_{(t_0,x_0)}^v$ um intervalo fechado. Para alguma bola aberta $\mathrm{B}(x_0; r) \subseteq A$ vale que, para todos $x \in \mathrm{B}(x_0; r)$, $I \subseteq I_{(t_0,x)}^v$.*

⊢ **Proposição 19.105.** *Sejam $\boldsymbol{E}$ um espaço normado completo, $A \subseteq E$ um aberto, $v\colon A \longrightarrow E$ um campo vetorial controlado, $x \in A$ e $I \subseteq I_x^v$ um intervalo fechado. Para algum $r \in \mathbb{R}_{>0}$ tal que $\mathrm{B}(x; r) \subseteq A$, todo $x' \in \mathrm{B}(x; r)$ satisfaz que $I \subseteq I_{x'}^v$.*

19.8.2.5 Propriedades geométricas do fluxo

19.8.2.6 Continuidade e diferenciabilidade do fluxo

19.9 Campos tensoriais

Campos tensoriais devem ser imaginados como um conjunto de tensores, cada um preso a um ponto de um aberto do espaço normado. A definição é dada a seguir.

:⊢ **Definição 19.62.** Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto e $k, k' \in \mathbb{N}$. Um *campo tensorial de tipo* (k, k') (ou *campo* (k, k')*-tensorial*) em A é uma função $T\colon A \longrightarrow E^{\otimes(k,k')}$. Para todo $x \in A$, o *tensor de* T *tangente a* A *em* x é

$$T|_x := T(x).$$

Um *campo escalar* é um campo tensorial de tipo $(0, 0)$, um *campo* k*-vetorial* é um campo tensorial de tipo $(0, k)$ (para $k = 1$, *campo vetorial*) e um *campo* k*-covetorial* é um campo tensorial de tipo $(k, 0)$ (para $k = 1$, *campo covetorial*).

O conjunto dos campos (k, k')-tensoriais $\mathscr{C}^r$-diferenciáveis é denotado $\mathfrak{T}^{(k,k')}_{\mathscr{C}^r}(A)$ e, em particular, $\mathfrak{T}^{(k,k')}(A) := \mathfrak{T}^{(k,k')}_{\mathscr{C}^\infty}(A)$.

Na definição, não exigimos que a função seja contínua, mas em geral consideraremos campos tensoriais que são contínuos ou com alguma regularidade ainda maior.

Além dessa definição, poderíamos considerar também campos tensoriais que variam no tempo, ou seja, uma família de campos tensoriais indexados por um intervalo de números reais.

:⊢ **Definição 19.63.** Sejam $\boldsymbol{E}$ um espaço normado e $A \subseteq E$ um aberto, $I \subseteq \mathbb{R}$ um intervalo aberto e $k, k' \in \mathbb{N}$. Um *campo tensorial temporal de tipo* (k, k') (ou *campo temporal* (k, l)*-tensorial*) em $I \times A$ é uma função $T\colon I \times A \longrightarrow E^{\otimes(k,k')}$. Para todo $(t, x) \in I \times A$, o *tensor de* T *tangente a* A *em* (t, x) é

$$T_t|_x := T(t, x).$$

Focaremos no estudo dos campos tensoriais e não nos campos tensoriais temporais.

O conjunto dos campos tensoriais $\mathfrak{T}^{(k,k')}_{\mathscr{C}^r}(A)$ é um subconjunto das funções $\mathcal{F}(A, E^{\otimes(k,k')})$, que é um espaço linear com as operações puxadas do espaço linear $E^{\otimes(k,k')}$ (definidas ponto a ponto). Para definir uma estrutura de espaço linear em $\mathfrak{T}^{(k,k')}_{\mathscr{C}^r}(A)$, basta mostrar que ele é um subespaço. Mas isso ocorre pois todas as operações algébricas preservam o tipo do tensor e a classe de diferenciabilidade dos campos.

Estudaremos a seguir campos escalares, vetoriais e covetoriais, mas antes ressaltamos um resultado geral sobre a diferenciabilidade da avaliação de uma transformação bilinear em um par de funções diferenciáveis, que segue diretamente da diferenciabilidade da transformação bilinear e da regra da cadeia. Nas seções seguintes, aplicaremos esse resultado em casos específicos para campos vetoriais, covetoriais, entre outros.

⊢ **Proposição 19.106.** *Sejam $\boldsymbol{E}$, $\boldsymbol{E}'$, $\boldsymbol{F}$ e $\boldsymbol{F}'$ espaços normados, $A \subseteq E$ um aberto, $B\colon F \times F' \longrightarrow E'$ uma função bilinear e $f\colon A \longrightarrow F$ e $f'\colon A \longrightarrow F'$ funções diferenciáveis. A função*

$$\begin{aligned} B(f,f')\colon A &\longrightarrow E' \\ x &\longmapsto B(f,f')(x) := B(f(x),f'(x)) \end{aligned}$$

é diferenciável e, para todo $u \in E$,

$$\mathrm{D}B(f,f')|_x(u) = B(\mathrm{D}f|_x(u), f'(x)) + B(f(x), \mathrm{D}f'|_x(u)).$$

Se f é $\mathscr{C}^r$-diferenciável e f' é $\mathscr{C}^{r'}$-diferenciável, então $B(f,f')$ é $\mathscr{C}^{r\wedge r'}$-diferenciável.

□ *Demonstração.* Note que $B(f,f') = B \circ (\phi, v)$, em que

$$\begin{aligned} (f,f')\colon A &\longrightarrow E^* \times E \\ x &\longmapsto (f(x), f'(x)). \end{aligned}$$

Como B é uma transformação bilinear, segue de 19.73 que ela é diferenciável e, para todos $(x,x'), (v,v') \in F \times F'$,

$$\mathrm{D}B|_{(x,x')}(v,v') = B(v,x') + B(x,v').$$

A função (f,f') é diferenciável pela proposição 19.68, pois f e f' o são, e

$$\mathrm{D}(f,f')|_x = (\mathrm{D}f|_x, \mathrm{D}f'|_x).$$

Portanto, pela regra da cadeia segue que, para todo $u \in E$,

$$\begin{aligned} \mathrm{D}B(f,f')|_x(u) &= \mathrm{D}B|_{(f,f')(x)} \circ \mathrm{D}(f,f')|_x(u) \\ &= \mathrm{D}B|_{(f(x),f'(x))} \circ (\mathrm{D}f|_x, \mathrm{D}f'|_x)(u) \\ &= \mathrm{D}B|_{(f(x),f'(x))}(\mathrm{D}f|_x(u), \mathrm{D}f'|_x(u)) \\ &= B(\mathrm{D}f|_x(u), f'(x)) + B(\mathrm{D}f|_x, \mathrm{D}f'|_x(u)). \end{aligned}$$

Finalmente, como B é bilinear, é $\mathscr{C}^\infty$-diferenciável, portanto se f é $\mathscr{C}^r$-diferenciável e f' é $\mathscr{C}^{r'}$-diferenciável, a função (f,f') é tão diferenciável quanto o mínimo entre r e r', e então a composição $B \circ (f,f')$ é $\mathscr{C}^{r\wedge r'}$-diferenciável. ■

19.9.1 Campos escalares

19.9.1.1 Álgebra dos campos escalares

Consideraremos aqui o conjunto dos campos escalares $\mathscr{C}^r(A, \mathbb{R})$.

19.9.2 Campos vetoriais

19.9.2.1 Álgebra dos campos vetoriais

Além das operações da estrutura de espaço linear em $\mathscr{C}^r(A, E)$, podemos multiplicar um campo escalar $f\colon A \longrightarrow E$ por um campo vetorial $v\colon A \longrightarrow E$ pontualmente, resultando em um novo campo vetorial

$$\begin{aligned} fv\colon A &\longrightarrow E \\ x &\longmapsto f(x)v(x). \end{aligned}$$

A diferenciabilidade desse novo campo vetorial segue do resultado geral sobre a diferenciabilidade da aplicação de uma transformação bilinear em um par de funções diferenciáveis.

⊢ **Proposição 19.107.** *Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto, $f\colon A \longrightarrow \mathbb{R}$ um campo escalar diferenciável e $v\colon A \longrightarrow E$ um campo vetorial diferenciável. O campo vetorial $fv\colon A \longrightarrow E$ é diferenciável e, para todo $u \in E$,*

$$\mathrm{D}(fv)|_x(u) = \mathrm{D}f|_x(u)v(x) + f(x)\mathrm{D}v|_x(u).$$

Se f é $\mathscr{C}^r$-diferenciável e v' é $\mathscr{C}^{r'}$-diferenciável, então fv é $\mathscr{C}^{r\wedge r'}$-diferenciável.

□ *Demonstração.* Consideremos a multiplicação por escalar dada por

$$\begin{aligned} \cdot\ : \mathbb{R} \times E &\longrightarrow E \\ (c, v) &\longmapsto cv. \end{aligned}$$

Essa função é bilinear, e $f(v) = \cdot \circ (f, v)$. Portanto segue da proposição 19.106 que, para todo $u \in E$,

$$\mathrm{D}(fv)|_x(u) = \mathrm{D}f|_x(u) \cdot v(x) + f(x) \cdot \mathrm{D}v|_x(u). \qquad \blacksquare$$

19.9.2.2 Derivada ao longo de um campo vetorial

Definiremos agora uma operação de um campo vetorial no espaço dos campos escalares. Para isso, notemos primeiro que, se temos um campo escalar $f\colon A \longrightarrow \mathbb{R}$, sua diferencial $\mathrm{D}f\colon A \longrightarrow \mathscr{L}(E, \mathbb{R})$ é um campo covetorial, pois $\mathscr{L}(E, \mathbb{R}) = E^*$. A

diferencial de um campo escalar é uma função que, para cada ponto $x \in A$, retorna um covetor $\mathrm{D}f|_x \colon E \longrightarrow \mathbb{R}$. Sendo assim, dado um campo vetorial $v \colon A \longrightarrow E$, podemos avaliar pontualmente a diferencial $\mathrm{D}f$ no campo v, de modo a termos um novo campo escalar

$$\begin{aligned} \mathrm{D}f(v) \colon A &\longrightarrow \mathbb{R} \\ x &\longmapsto \mathrm{D}f|_x(v(x)). \end{aligned}$$

A diferenciabilidade desse novo campo escalar segue do resultado geral sobre a diferenciabilidade da aplicação de uma transformação bilinear em um par de funções diferenciáveis.

⊢ **Proposição 19.108.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $f \colon A \longrightarrow \mathbb{R}$ *um campo escalar* 2*-diferenciável e* $v \colon A \longrightarrow E$ *um campo vetorial diferenciável. O campo escalar* $\mathrm{D}f(v) \colon A \longrightarrow \mathbb{R}$ *é diferenciável e, para todo* $u \in E$*,*

$$\mathrm{D}(\mathrm{D}f(v))|_x(u) = \mathrm{D}^{(2)}f|_x(u, v(x)) + \mathrm{D}f|_x \circ \mathrm{D}v|_x(u).$$

Se f *é* $\mathscr{C}^{1+r}$*-diferenciável e* v *é* $\mathscr{C}^{r'}$*-diferenciável, então* $\mathrm{D}f(v)$ *é* $\mathscr{C}^{r \wedge r'}$*-diferenciável.*

□ *Demonstração.* Consideremos a avaliação de um covetor em um vetor, dada por

$$\begin{aligned} B \colon E^* \times E &\longrightarrow E \\ (\phi, v) &\longmapsto \phi(v). \end{aligned}$$

Essa função é bilinear, e $\mathrm{D}f(v) = B \circ (\mathrm{D}f, v)$. Portanto segue da proposição 19.106 que, para todo $u \in E$,

$$\begin{aligned} \mathrm{D}(\mathrm{D}f(v))|_x(u) &= \mathrm{D}\mathrm{D}f|_x(u)(v(x)) + \mathrm{D}f|_x(\mathrm{D}v|_x(u)) \\ &= \mathrm{D}^2 f|_x(u)(v(x)) + \mathrm{D}f|_x \circ \mathrm{D}v|_x(u) \\ &= \mathrm{D}^{(2)}f|_x(u, v(x)) + \mathrm{D}f|_x \circ \mathrm{D}v|_x(u). \end{aligned}$$

Se f é $\mathscr{C}^{1+r}$-diferenciável, então $\mathrm{D}f$ é $\mathscr{C}^r$-diferenciável, portanto segue que $\mathrm{D}f(v)$ é $\mathscr{C}^{r \wedge r'}$-diferenciável. ■

Dadas tais considerações, podemos definir a derivada ao longo de um campo vetorial.

:⊢ **Definição 19.64.** Sejam $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto, $v \colon A \longrightarrow E$ um campo vetorial e $f \colon A \longrightarrow \mathbb{R}$ um campo escalar diferenciável. A *derivada direcional de* f *ao longo de* v é o campo escalar

$$\partial_v f := \mathrm{D}f(v),$$

dado pontualmente, para todo $x \in A$, por

$$\partial_v f(x) := \mathrm{D}f|_x(v(x)).$$

Para um campo vetorial v ($\mathscr{C}^r$-diferenciável), a *derivação ao longo de* v é a função

$$\begin{aligned} \partial_v \colon \mathscr{C}^{1+r}(A, \mathbb{R}) &\longrightarrow \mathscr{C}^r(A, \mathbb{R}) \\ f &\longmapsto \partial_v f. \end{aligned}$$

⊢ **Proposição 19.109.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto e* $v \colon A \longrightarrow E$ *um campo vetorial* $\mathscr{C}^r$*-diferenciável.*

1. *Se* $\partial_v = 0$*, então* $v = 0$*.*
2. *(Derivação) Para todos campos escalares* $f, f' \in \mathscr{C}^{1+r}(A, \mathbb{R})$*,*

$$\partial_v(ff') = (\partial_v f)f' + f(\partial_v f').$$

□ *Demonstração.* 1. Se $\partial_v = 0$, então, para todo campo escalar $f \in \mathscr{C}^{1+r}$ vale $\partial_v f = 0$, o que por sua vez significa que, para todo $x \in A$,

$$0 = \partial_v f(x) = \mathrm{D}f|_x(v(x)).$$

Suponhamos que $v \neq 0$. Então, para algum $x \in A$, $v(x) \neq 0$. Como $v(x) \in E$, segue do teorema de Hahn-Banach que existe funcional linear contínuo $f \in \mathscr{L}(E, \mathbb{R})$ tal que $f(v(x)) \neq 0$. Como f é linear, $\mathrm{D}f|_x = f$, portanto

$$\partial_v f(x) = \mathrm{D}f|_x(v(x)) = f(v(x)) \neq 0,$$

o que mostra que $\partial_v f \neq 0$. Como $f \neq 0$, segue que $\partial_v \neq 0$.

2. Para todo $x \in A$, segue da regra do produto para a diferencial que

$$\begin{aligned} \partial_v(ff')(x) &= \mathrm{D}(ff')|_x(v(x)) \\ &= \mathrm{D}f|_x(v(x))f'(x) + f(x)\mathrm{D}f'|_x(v(x)) \\ &= (\partial_v f(x))f'(x) + f(x)(\partial_v f'(x)). \end{aligned}$$

■

19.9.2.3 Colchete de campos vetoriais

Definimos agora uma operação entre dois campos que vetoriais que dará ao espaço linear dos campos vetoriais suaves uma estrutura de álgebra. Para isso, notemos que, dado um campo vetorial diferenciável $v \colon A \longrightarrow E$, sua diferencial $\mathrm{D}v \colon A \longrightarrow \mathscr{L}(E, E)$ pode ser entendida como um campo $(1,1)$-tensorial através do isomorfismo $\mathscr{L}(E, E) \simeq E^* \otimes E$. A diferencial é uma função que, para cada

ponto $x \in A$, retorna uma transformação linear contínua $\mathrm{D}v|_x \colon E \longrightarrow E$. Essa transformação linear pode, portanto, ser avaliada em um vetor. Sendo assim, dado outro campo vetorial $v' \colon A \longrightarrow E$, podemos avaliar pontualmente a diferencial $\mathrm{D}v$ no campo v', de modo a termos um novo campo vetorial

$$\begin{aligned} \mathrm{D}v(v') \colon A &\longrightarrow E \\ x &\longmapsto \mathrm{D}v(v')(x) = \mathrm{D}v|_x(v'(x)). \end{aligned}$$

A diferenciabilidade desse novo campo vetorial segue do resultado sobre a diferenciabilidade da aplicação de uma função bilinear em um par de funções diferenciáveis.

⊢ **Proposição 19.110.** *Sejam* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto,* $v \colon A \longrightarrow E$ *um campo vetorial* 2*-diferenciável e* $v' \colon A \longrightarrow E$ *um campo vetorial diferenciável. O campo vetorial* $\mathrm{D}v(v') \colon A \longrightarrow E$ *é diferenciável e, para todo* $u \in E$,

$$\mathrm{D}(\mathrm{D}v(v'))|_x(u) = \mathrm{D}^{(2)}v|_x(u, v'(x)) + \mathrm{D}v|_x \circ \mathrm{D}v'|_x(u).$$

Se v *é* $\mathscr{C}^{1+r}$*-diferenciável e* v' *é* $\mathscr{C}^{r'}$*-diferenciável, então* $\mathrm{D}v(v')$ *é* $\mathscr{C}^{r \wedge r'}$*-diferenciável.*

□ *Demonstração.* Consideremos a avaliação de uma transformação linear em um vetor, dada por

$$\begin{aligned} B \colon \mathscr{L}(E, E) \times E &\longrightarrow E \\ (l, v) &\longmapsto l(v). \end{aligned}$$

Essa função é bilinear, e $\mathrm{D}v(v') = B \circ (\mathrm{D}v, v')$. Portanto segue da proposição 19.106 que, para todo $u \in E$,

$$\begin{aligned} \mathrm{D}(\mathrm{D}v(v'))|_x(u) &= \mathrm{D}\mathrm{D}v|_x(u)(v'(x)) + \mathrm{D}v|_x(\mathrm{D}v'|_x(u)) \\ &= \mathrm{D}^{(2)}v|_x(u, v'(x)) + \mathrm{D}v|_x \circ \mathrm{D}v'|_x(u). \end{aligned}$$

Se v é $\mathscr{C}^{1+r}$-diferenciável, então $\mathrm{D}v$ é $\mathscr{C}^r$-diferenciável, portanto segue que $\mathrm{D}v(v')$ é $\mathscr{C}^{r \wedge r'}$-diferenciável. ■

:⊢ **Definição 19.65.** Seja $\boldsymbol{E}$ um espaço normado, $A \subseteq E$ um aberto e $v, v' \colon A \longrightarrow E$ campos vetoriais diferenciáveis. O *colchete entre* v *e* v' é o campo vetorial

$$[v, v'] := \mathrm{D}v'(v) - \mathrm{D}v(v''),$$

dado pontualmente, para todo $x \in A$, por

$$[v, v'](x) := \mathrm{D}v'|_x(v(x)) - \mathrm{D}v|_x(v'(x)).$$

Apresentamos a seguir algumas propriedades básicas do colchete de campos vetoriais.

⊢ **Proposição 19.111.** *Seja* $\boldsymbol{E}$ *um espaço normado e* $A \subseteq E$ *um aberto.*

1. *(Bilinearidade) Para todos campos vetoriais* $v, v' \in \mathscr{C}^{1+r}(A, E)$ *e todo* $c \in \mathbb{R}$,

$$\begin{aligned}[cv + v', v''] &= c\,[v, v'] + [v', v'']\,,\\ [v, cv' + v''] &= c\,[v, v'] + [v, v'']\,.\end{aligned}$$

2. *(Alternatividade) Para todo campo vetorial* $v \in \mathscr{C}^{1+r}(A, E)$,

$$[v, v] = 0.$$

3. *(Derivação adjunta) Para todos campos vetoriais* $v, v', v'' \in \mathscr{C}^{2+r}(A, E)$,

$$[v, [v', v'']] = [[v, v'], v''] + [v', [v, v'']]\,.$$

4. *Para todos campos vetoriais* $v, v' \in \mathscr{C}^{1+r}(A, E)$ *e todo campo escalar* $f \in \mathscr{C}^{1+r}(A, \mathbb{R})$,

$$\begin{aligned}[v, fv'] &= \partial_v f v' + f\,[v, v]\,,\\ [fv, v'] &= f\,[v, v] - \partial_{v'} f v.\end{aligned}$$

□ *Demonstração.* 1. Segue da linearidade da diferencial e da bilinearidade da avaliação de uma transformação linear em um vetor.

2. Segue direto da definição que, para todo $x \in A$,

$$[v, v]\,(x) = \mathrm{D}v|_x(v(x)) - \mathrm{D}v|_x(v(x)) = 0.$$

3. Da definição do colchete de campos vetoriais,

$$\begin{aligned}[v, [v', v'']]\,(x) &= \mathrm{D}\,[v', v'']\,|_x(v(x)) - \mathrm{D}v|_x([v', v'']\,(x))\\ &= \mathrm{D}\,[v', v'']\,|_x(v(x)) - \mathrm{D}v|_x(\mathrm{D}v''|_x(v'(x)) - \mathrm{D}v'|_x(v''(x)))\\ &= \mathrm{D}\,[v', v'']\,|_x(v(x)) - \mathrm{D}v|_x \circ \mathrm{D}v''|_x(v'(x)) + \mathrm{D}v|_x \circ \mathrm{D}v'|_x(v''(x)).\end{aligned}$$

Da proposição 19.110 segue que, para todo $u \in E$,

$$\mathrm{D}(\mathrm{D}v''(v'))|_x(u) = \mathrm{D}^{(2)}v''|_x(u, v'(x)) + \mathrm{D}v''|_x \circ \mathrm{D}v'|_x(u),$$

portanto, como $[v', v''] = \mathrm{D}v''(v) - \mathrm{D}v'(v'')$,

$$\begin{aligned}\mathrm{D}\,[v', v'']\,|_x(v(x)) &= \mathrm{D}(\mathrm{D}v''(v'))|_x(v(x)) - \mathrm{D}(\mathrm{D}v'(v''))|_x(v(x))\\ &= \mathrm{D}^{(2)}v''|_x(v(x), v'(x)) + \mathrm{D}v''|_x \circ \mathrm{D}v'|_x(v(x))\\ &\quad - \mathrm{D}^{(2)}v'|_x(v(x), v''(x)) - \mathrm{D}v'|_x \circ \mathrm{D}v''|_x(v(x)).\end{aligned}$$

Segue então que

$$\begin{aligned}[v,[v',v'']](x) &= \mathrm{D}^{(2)}v''|_x(v(x),v'(x)) - \mathrm{D}^{(2)}v'|_x(v(x),v''(x)) \\ &+ \mathrm{D}v''|_x \circ \mathrm{D}v'|_x(v(x)) - \mathrm{D}v'|_x \circ \mathrm{D}v''|_x(v(x)) \\ &- \mathrm{D}v|_x \circ \mathrm{D}v''|_x(v'(x)) + \mathrm{D}v|_x \circ \mathrm{D}v'|_x(v''(x)).\end{aligned}$$

Como consequência dos itens anteriores, o colchete é antissimétrico, portanto $[[v,v''],v''] = [v'',[v',v]]$. Permutando v, v' e v'' na fórmula anterior, obtêm-se

$$\begin{aligned}[[v,v''],v''](x) &= \mathrm{D}^{(2)}v|_x(v''(x),v'(x)) - \mathrm{D}^{(2)}v'|_x(v''(x),v(x)) \\ &+ \mathrm{D}v|_x \circ \mathrm{D}v'|_x(v''(x)) - \mathrm{D}v'|_x \circ \mathrm{D}v|_x(v'(x)) \\ &- \mathrm{D}v''|_x \circ \mathrm{D}v|_x(v'(x)) + \mathrm{D}v''|_x \circ \mathrm{D}v'|_x(v(x))\end{aligned}$$

e

$$\begin{aligned}[v',[v,v'']](x) &= \mathrm{D}^{(2)}v''|_x(v'(x),v(x)) - \mathrm{D}^{(2)}v|_x(v'(x),v''(x)) \\ &+ \mathrm{D}v''|_x \circ \mathrm{D}v|_x(v'(x)) - \mathrm{D}v|_x \circ \mathrm{D}v''|_x(v'(x)) \\ &- \mathrm{D}v'|_x \circ \mathrm{D}v''|_x(v(x)) + \mathrm{D}v'|_x \circ \mathrm{D}v|_x(v''(x)).\end{aligned}$$

Finalmente, como v, v' e v'' são $\mathscr{C}^2$-diferenciáveis, as diferenciais segundas $\mathrm{D}^{(2)}v|_x$, $\mathrm{D}^{(2)}v'|_x$ e $\mathrm{D}^{(2)}v''|_x$ são simétricas, portanto somando as fórmulas anteriores e cancelando os termos repetidos,

$$[v,[v',v'']](x) = [[v,v'],v''](x) + [v',[v,v'']](x).$$

4. Segue da proposição 19.107 que

$$\begin{aligned}[v,fv'](x) &= \mathrm{D}(fv')|_x(v(x)) - \mathrm{D}v|_x(f(x)v'(x)) \\ &= \mathrm{D}f|_x(v(x))v'(x) + f(x)\mathrm{D}v'|_x(v(x)) - f(x)\mathrm{D}v|_x(v'(x)) \\ &= \partial_v f(x)v'(x) + f(x)\,[v,v'](x).\end{aligned}$$

■

Agora relacionamos o colchete de campos vetoriais com a derivada ao longo de um campo vetorial.

⊢ **Proposição 19.112.** *Seja* $\boldsymbol{E}$ *um espaço normado,* $A \subseteq E$ *um aberto e* $v, v' \colon A \longrightarrow E$ *campos vetoriais* $\mathscr{C}^{2+r}$*-diferenciáveis. Então*

$$\partial_{[v,v']} = \partial_v \circ \partial_{v'} - \partial_{v'} \circ \partial_v.$$

□ *Demonstração.* Sejam $f \in \mathscr{C}^{2+r}(A,\mathbb{R})$ e $x \in A$. Por definição, $\partial_{v'}f = \mathrm{D}f(v')$. Segue então da proposição 19.108 que, para todo $u \in E$,

$$\mathrm{D}(\mathrm{D}f(v'))|_x(u) = \mathrm{D}^{(2)}f|_x(u,v'(x)) + \mathrm{D}f|_x \circ \mathrm{D}v'|_x(u),$$

portanto,

$$\begin{aligned}\partial_v \circ \partial_{v'} f(x) &= \partial_v(\mathrm{D}f(v'))(x)\\ &= \mathrm{D}(\mathrm{D}f(v'))|_x(v(x))\\ &= \mathrm{D}^{(2)}f|_x(v(x), v'(x)) + \mathrm{D}f|_x \circ \mathrm{D}v'|_x(v(x)).\end{aligned}$$

Permutando v e v', obtemos

$$\partial_{v'} \circ \partial_v f(x) = \mathrm{D}^{(2)}f|_x(v'(x), v(x)) + \mathrm{D}f|_x \circ \mathrm{D}v|_x(v'(x)).$$

Como f é $\mathscr{C}^{2+r}$-diferenciável, $\mathrm{D}^{(2)}f|_x$ é simétrica, logo

$$\partial_v \circ \partial_{v'} f(x) - \partial_{v'} \circ \partial_v f(x) = \mathrm{D}f|_x \circ \mathrm{D}v'|_x(v(x)) - \mathrm{D}f|_x \circ \mathrm{D}v|_x(v'(x)).$$

Assim concluímos que

$$\begin{aligned}\partial_{[v,v']} f(x) &= \mathrm{D}f|_x([v, v'](x))\\ &= \mathrm{D}f|_x(\mathrm{D}v'|_x(v(x)) - \mathrm{D}v|_x(v'(x)))\\ &= \mathrm{D}f|_x \circ \mathrm{D}v'|_x(v(x)) - \mathrm{D}f|_x \circ \mathrm{D}v|_x(v'(x))\\ &= \partial_v \circ \partial_{v'} f(x) - \partial_{v'} \circ \partial_v f(x).\end{aligned}$$

■

19.9.3 Álgebra dos campos covetoriais alternados

Capítulo 20

Espaços lineares com produto interno

O foco dessa seção é definir o produto interno em um espaço vetorial e estudar suas propriedades geométricas. O produto interno será definido para qualquer espaço vetorial sobre o corpo dos números complexos, que denotaremos por $\mathbb{R}^2$. De fato, ele será definido para qualquer subcorpo de $\mathbb{R}^2$ invariante por conjugação complexa, o que inclui o corpo dos números reais. O produto interno poderia ser definido sobre os números quaterniônicos $\mathbb{R}^4$ da mesma forma, mas notando que esse anel não é comutativo, embora seja invertível no sentido de que todo elemento não 0 tem inversa, tratando-se portanto de um anel de divisão. Muitas das propriedades do produto interno ainda valeriam para coeficientes quaterniônicos, mas cuidado deve ser tomado devido à não-comutatividade.

20.1 Produto interno

:⊢ **Definição 20.1.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ invariante por conjugação complexa[1]. Um *produto interno* em $\boldsymbol{V}$ é uma função $\langle \cdot, \cdot \rangle : V \times V \to C$ que satisfaz

1. (Linearidade na primeira entrada)

   1.1. Para todos $v_0, v_1, v \in V$,

   $$\langle v_0 + v_1, v \rangle = \langle v_0, v \rangle + \langle v_1, v \rangle ;$$

   1.2. Para todos $v, v' \in V$ e $c \in C$,

   $$\langle cv, v' \rangle = c \langle v, v' \rangle ;$$

[1] Um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ tal que, para todo $c \in C$, temos $c^* \in C$.

2. (Simetria conjugada) Para todos $v, v' \in V$,

$$\langle v, v' \rangle = \langle v', v \rangle^*;$$

3. (Positividade) Para todo $v \in V$, $\langle v, v \rangle \in [0, \infty[$;
4. (Definição positiva) Para todo $v \in V$, se $\langle v, v \rangle = 0$, então $v = 0$.

:⊢ **Definição 20.2.** Um *espaço com produto interno* é um par $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ em que $\boldsymbol{V}$ é um espaço linear sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ invariante por conjugação complexa e $\langle \cdot, \cdot \rangle : \boldsymbol{V} \times \boldsymbol{V} \to C$ é um produto interno em $\boldsymbol{V}$.

⊢ **Proposição 20.1** (Propriedades de Produto Interno)**.** *Seja* $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ *um espaço com produto interno sobre um corpo* $\boldsymbol{C} \subseteq \mathbb{R}^2$*. Então*

1. *(Linearidade conjugada na segunda entrada)*

   1.1. *Para todos* $v_0, v_1, v \in V$*,*

   $$\langle v, v_0 + v_1 \rangle = \langle v, v_0 \rangle + \langle v, v_1 \rangle;$$

   1.2. *Para todos* $v, v' \in V$ *e* $c \in C$*,*

   $$\langle v', cv \rangle = c^* \langle v', v \rangle;$$

2. *Para todos* $v_0, \ldots, v_n, v \in V$ *e* $c_0, \ldots, c_n \in C$*,*

   $$\left\langle \mathop{+}_{i=0}^{n} c_i v_i, v \right\rangle = \mathop{+}_{i=0}^{n} c_i \langle v_i, v \rangle$$

   e

   $$\left\langle v, \mathop{+}_{i=0}^{n} c_i v_i \right\rangle = \mathop{+}_{i=0}^{n} c_i{}^* \langle v, v_i \rangle.$$

3. *(Desigualdade do produto interno*[2]*) Para todos* $v, v' \in V$*,*

   $$|\langle v, v' \rangle|^2 \leq \langle v, v \rangle \langle v', v' \rangle$$

   e a igualdade ocorre se, e somente se, um vetor é múltiplo do outro.

□ *Demonstração.* 1. Exercício simples.
2. Exercício simples.

[2] Conhecida por 'desigualdade de Cauchy–Schwarz'.

3. Se existe $c \in C$ tal que $v' = cv$, então

$$\begin{aligned}
|\langle v, v'\rangle|^2 &= |\langle v, cv\rangle|^2 \\
&= |c|^2 \, |\langle v, v\rangle|^2 \\
&= cc^* \, |\langle v, v\rangle|^2 \\
&= \langle v, v\rangle \, \langle cv, cv\rangle \\
&= \langle v, v\rangle \, \langle v', v'\rangle \, .
\end{aligned}$$

Caso contrário, se $v' - cv \neq 0$ para todo $c \in C$, então para $c = \frac{\langle v', v\rangle}{\langle v, v\rangle}$ segue que

$$\begin{aligned}
0 &< \langle v' - cv, v' - cv\rangle \\
&= \langle v', v'\rangle - c \, \langle v, v'\rangle - c^* \, \langle v', v\rangle + |c|^2 \, \langle v, v\rangle \\
&= \langle v', v'\rangle - c\langle v', v\rangle^* - c^* \, \langle v', v\rangle + |c|^2 \, \langle v, v\rangle \\
&= \langle v', v'\rangle - \frac{|\langle v, v'\rangle|^2}{\langle v, v\rangle} - \frac{|\langle v, v'\rangle|^2}{\langle v, v\rangle} + \frac{\langle v, v'\rangle^2}{\langle v, v\rangle} \\
&= \langle v', v'\rangle - \frac{|\langle v, v'\rangle|^2}{\langle v, v\rangle},
\end{aligned}$$

o que implica

$$|\langle v, v'\rangle|^2 < \langle v, v\rangle \, \langle v', v'\rangle \, .$$

■

⊢ **Proposição 20.2.** *Sejam $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ invariante por conjugação complexa e $(b_i)_{i \in I}$ uma base ordenada de $\boldsymbol{V}$. Existe único produto interno $\langle \cdot, \cdot \rangle$ em $\boldsymbol{V}$ tal que, para todos $i, j \in I$, $\langle b_i, b_j \rangle = \delta_{i,j}$.*

20.2 Norma induzida, ortogonalidade e ângulo

20.2.1 Norma

:⊢ **Definição 20.3.** Seja $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$. A *norma* (induzida pelo produto interno) de $\boldsymbol{V}$ é a função

$$\begin{aligned}
\|\cdot\| \colon V &\longrightarrow \mathbb{R} \\
v &\longmapsto \langle v, v\rangle^{\frac{1}{2}} \, .
\end{aligned}$$

Em termos da norma, a desigualdade do produto interno fica

$$|\langle v, v'\rangle| \leq \|v\| \, \|v'\| \, .$$

⊢ **Proposição 20.3.** *Seja* $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ *um espaço com produto interno sobre um corpo* $\boldsymbol{C} \subseteq \mathbb{R}^2$. *A função* $\|\cdot\|$ *é uma norma em* $\boldsymbol{V}$.

□ *Demonstração.* 1. (Separação) Seja $v \in V$ tal que $\|v\| = 0$. Então $\langle v, v\rangle^{\frac{1}{2}} = 0$, portanto $\langle v, v\rangle = 0$, o que implica $v = 0$.

2. (Homogeneidade absoluta) Sejam $c \in C$ e $v \in V$. Então

$$\|cv\| = \langle cv, cv\rangle^{\frac{1}{2}} = (cc^* \langle v, v\rangle)^{\frac{1}{2}} = |c| \, \|v\| \, .$$

3. (Subaditividade) Para todos $v, v' \in V$, da desigualdade do produto interno segue que

$$\begin{aligned}
\|v + v'\|^2 &= \langle v + v', v + v'\rangle \\
&= \langle v, v\rangle + \langle v, v'\rangle + \langle v', v\rangle + \langle v', v'\rangle \\
&= \|v\|^2 + \langle v, v'\rangle + \langle v, v'\rangle^* + \|v'\|^2 \\
&= \|v\|^2 + 2\Re(\langle v, v'\rangle) + \|v'\|^2 \\
&\leq \|v\|^2 + 2\,|\langle v, v'\rangle| + \|v'\|^2 \\
&\leq \|v\|^2 + 2\,\|v\|\,\|v'\| + \|v'\|^2 \\
&= (\|v\| + \|v'\|)^2.
\end{aligned}$$

■

▷ **Exercício 20.1.** *Seja* $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ *um espaço com produto interno.*

1. *(Regra do Paralelogramo) Para todos* $v, v' \in V$,

$$2\left(\|v\|^2 + \|v'\|^2\right) = \|v + v'\|^2 + \|v - v'\|^2 \, ;$$

2. *(Polarização) Para todos* $v, v' \in V$, *valem*

$$\Re(\langle v, v'\rangle) = \frac{1}{4}(\|v + v'\|^2 - \|v - v'\|^2),$$

$$\Im(\langle v, v'\rangle) = \frac{1}{4}(\|v + \mathrm{i}v'\|^2 - \|v - \mathrm{i}v'\|^2)\mathrm{i},$$

e

$$\langle v, v'\rangle = \frac{(\|v + v'\|^2 - \|v - v'\|^2) + (\|v + \mathrm{i}v'\|^2 - \|v - \mathrm{i}v'\|^2)\mathrm{i}}{4}.$$

De certa forma, vale uma recíproca da polarização quando uma norma tem a propriedade de paralelogramo. De fato, a igualdade do paralelogramo pode ser enfraquecida para uma desigualdade.

$\vdash$ **Proposição 20.4.** *Seja* $(\boldsymbol{V}, \|\cdot\|)$ *um espaço normado tal que, para todos* $v, v' \in V$,

$$2(\|v\|^2 + \|v'\|^2) \leq \|v + v'\|^2 + \|v - v'\|^2 .$$

A função

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon V \times V &\longrightarrow C \\ (v, v') &\longmapsto \langle v, v' \rangle := \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \end{aligned}$$

é um produto interno em $\boldsymbol{V}$ *tal que, para todo* $v \in V$, $\|v\| = \langle v, v \rangle^{\frac{1}{2}}$. *Para um corpo real,*

$$\langle v, v' \rangle = \frac{\|v + v'\|^2 - \|v - v'\|^2}{4}$$

e, para um corpo complexo,

$$\langle v, v' \rangle = \frac{(\|v + v'\|^2 - \|v - v'\|^2) + (\|v + \mathrm{i}v'\|^2 - \|v - \mathrm{i}v'\|^2)\mathrm{i}}{4}.$$

$\square$ *Demonstração.* O primeiro passo é mostrar que vale a regra do paralelogramo. Segue da desigualdade que

$$2\left(\left\| \frac{v + v'}{2} \right\|^2 + \left\| \frac{v - v'}{2} \right\|^2 \right) \leq \left\| \frac{v + v'}{2} + \frac{v - v'}{2} \right\|^2 + \left\| \frac{v + v'}{2} - \frac{v - v'}{2} \right\|^2$$

e disso segue que

$$\begin{aligned} \|v + v'\|^2 + \|v - v'\|^2 &= 4\left(\left\| \frac{v + v'}{2} \right\|^2 + \left\| \frac{v - v'}{2} \right\|^2 \right) \\ &\leq 2\left(\left\| \frac{v + v'}{2} + \frac{v - v'}{2} \right\|^2 + \left\| \frac{v + v'}{2} - \frac{v - v'}{2} \right\|^2 \right) \\ &= 2(\|v\|^2 + \|v'\|^2). \end{aligned}$$

Assim, temos a desigualdade oposta e segue a regra do paralelogramo.

A definição do produto interno é equivalente a

$$\langle v, v' \rangle = \frac{(\|v + v'\|^2 - \|v - v'\|^2) + (\|v + \mathrm{i}v'\|^2 - \|v - \mathrm{i}v'\|^2)\mathrm{i}}{4}.$$

Como $\mathrm{i}^{-1} = -\mathrm{i}$, para todo $v \in V$ vale

$$\|v + \mathrm{i}v\| = \|\mathrm{i}(-\mathrm{i}v + v)\| = |\mathrm{i}| \, \|v - \mathrm{i}v\| = \|v - \mathrm{i}v\| ,$$

portanto segue da definição do produto interno que

$$\begin{aligned}\langle v, v\rangle &= \frac{1}{4}\mathop{+}_{k=0}^{3}\left\|v+\mathrm{i}^k v\right\|^2 \mathrm{i}^k \\ &= \frac{(\|v+v\|^2-\|v-v\|^2)+(\|v+\mathrm{i}v\|^2-\|v-\mathrm{i}v\|^2)\mathrm{i}}{4} \\ &= \frac{\|2v\|^2}{4} \\ &= \|v\|^2 .\end{aligned}$$

Para um corpo real, como $\|v+\mathrm{i}v'\| = \|v-\mathrm{i}v'\|$, a fórmula se reduz a

$$\langle v, v'\rangle = \frac{\|v+v'\|^2-\|v-v'\|^2}{4}.$$

Mostremos que $\langle\cdot,\cdot\rangle$ é um produto interno.

1. (Linearidade na primeira entrada)
   1.1. (Aditividade) Sejam $v, v', v'' \in V$ e $k \in [4]$. Pela propriedade do paralelogramo,

$$\left\|v+\mathrm{i}^k v''+v'\right\|^2 = 2\left\|v+\mathrm{i}^k v''\right\|^2 + 2\|v'\|^2 - \left\|v+\mathrm{i}^k v''-v'\right\|^2$$

e

$$\left\|v'+\mathrm{i}^k v''+v\right\|^2 = 2\left\|v'+\mathrm{i}^k v''\right\|^2 + 2\|v\|^2 - \left\|v'+\mathrm{i}^k v''-v\right\|^2 .$$

Então

$$\begin{aligned}\left\|v+v'+\mathrm{i}^k v''\right\|^2 &= \frac{\left\|v+\mathrm{i}^k v''+v'\right\|^2+\left\|v'+\mathrm{i}^k v''+v\right\|^2}{2} \\ &= \left\|v+\mathrm{i}^k v''\right\|^2+\left\|v'+\mathrm{i}^k v''\right\|^2 \\ &+ \|v\|^2+\|v'\|^2 \\ &- \frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2+\left\|v'+\mathrm{i}^k v''-v\right\|^2}{2}.\end{aligned}$$

Como

$$\left\|v'+\mathrm{i}^k v''-v\right\|^2 = \left\|-(v-\mathrm{i}^k v''-v')\right\|^2 = \left\|v+\mathrm{i}^{k+2}v''-v'\right\|^2 ,$$

segue que

$$
\begin{aligned}
&\mathop{+}_{k=0}^{3}\left(\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2+\left\|v'+\mathrm{i}^k v''-v\right\|^2}{2}\right)\mathrm{i}^k \\
&\quad=\mathop{+}_{k=0}^{3}\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2}{2}\mathrm{i}^k+\mathop{+}_{k=0}^{3}\frac{\left\|v'+\mathrm{i}^k v''-v\right\|^2}{2}\mathrm{i}^k \\
&\quad=\mathop{+}_{k=0}^{3}\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2}{2}\mathrm{i}^k+\mathop{+}_{k=0}^{3}\frac{\left\|v+\mathrm{i}^{k+2} v''-v'\right\|^2}{2}\mathrm{i}^k \\
&\quad=\mathop{+}_{k=0}^{3}\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2}{2}\mathrm{i}^k-\mathop{+}_{k=0}^{3}\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2}{2}\mathrm{i}^k \\
&\quad=0,
\end{aligned}
$$

portanto

$$
\begin{aligned}
\langle v+v',v''\rangle &= \frac{1}{4}\mathop{+}_{k=0}^{3}\left\|v+v'+\mathrm{i}^k v''\right\|^2\mathrm{i}^k \\
&= \frac{1}{4}\mathop{+}_{k=0}^{3}\left(\left\|v+\mathrm{i}^k v''\right\|^2+\left\|v'+\mathrm{i}^k v''\right\|^2\right)\mathrm{i}^k \\
&+\frac{1}{4}\mathop{+}_{k=0}^{3}\left(\|v\|^2+\|v'\|^2\right)\mathrm{i}^k \\
&-\frac{1}{4}\mathop{+}_{k=0}^{3}\left(\frac{\left\|v+\mathrm{i}^k v''-v'\right\|^2+\left\|v'+\mathrm{i}^k v''-v\right\|^2}{2}\right)\mathrm{i}^k \\
&= \frac{1}{4}\mathop{+}_{k=0}^{3}\left\|v+\mathrm{i}^k v''\right\|^2\mathrm{i}^k+\frac{1}{4}\mathop{+}_{k=0}^{3}\left\|v'+\mathrm{i}^k v''\right\|^2\mathrm{i}^k \\
&= \langle v,v''\rangle+\langle v',v''\rangle .
\end{aligned}
$$

1.2. (Homogeneidade) Sejam $v, v' \in V$. Mostraremos a homogeneidade por partes: para os números naturais, inteiros, racionais, reais e então complexos.
(Números naturais) Mostremos por indução. Para $n = 0$, basta notar

que

$$\begin{aligned}
\langle 0v, v' \rangle &= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| 0v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} \left| \mathrm{i}^k \right| \left\| v' \right\|^2 \mathrm{i}^k \\
&= \left\| v' \right\|^2 \frac{1}{4} \mathop{+}_{k=0}^{3} \mathrm{i}^k \\
&= 0 \\
&= 0 \langle v, v' \rangle .
\end{aligned}$$

Supondo que vale para n, segue da aditividade que

$$\begin{aligned}
\langle (n+1)v, v' \rangle &= \langle nv + v, v' \rangle \\
&= \langle nv, v' \rangle + \langle v, v' \rangle \\
&= n \langle v, v' \rangle + \langle v, v' \rangle \\
&= (n+1) \langle v, v' \rangle .
\end{aligned}$$

(Números inteiros) Seja $n \in \mathbb{N}$. A homogeneidade vale para -1 pois, como $\mathrm{i}^2 = \mathrm{i}^{-2} = -1$,

$$\begin{aligned}
\langle -v, v' \rangle &= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| -v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} |-1| \left\| v - \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| v + \mathrm{i}^{k-2} v' \right\|^2 \mathrm{i}^k \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^{k+2} \\
&= \frac{1}{4} \mathop{+}_{k=0}^{3} \left\| v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \mathrm{i}^2 \\
&= -\frac{1}{4} \mathop{+}_{k=0}^{3} \left\| v + \mathrm{i}^k v' \right\|^2 \mathrm{i}^k \\
&= - \langle v, v' \rangle .
\end{aligned}$$

Segue da homogeneidade para -1 que

$$\langle(-n)v, v'\rangle = \langle-(nv), n'\rangle = -\langle nv, v'\rangle = -n\langle v, v'\rangle.$$

(Números racionais) Seja $\frac{n}{d} \in \mathbb{Q}$. Segue da homogeneidade para números inteiros que

$$\left\langle \frac{n}{d}v, v' \right\rangle = \frac{d}{d}\left\langle \frac{n}{d}v, v' \right\rangle = \frac{n}{d}\left\langle \frac{d}{d}v, v' \right\rangle = \frac{n}{d}\langle v, v'\rangle.$$

(Números reais) Sejam $c \in \mathbb{R}$ e $(c_n)_{n\in\mathbb{N}}$ uma sequência de números racionais tal que $c = \lim_{n\to\infty} c_n$. A função $\langle\cdot,\cdot\rangle$ é contínua, pois é composição de adição, subtração, multiplicação por escalar e norma, que são contínuas. Da continuidade de $\langle\cdot,\cdot\rangle$ e da homogeneidade para números racionais, segue que

$$\langle cv, v'\rangle = \left\langle \lim_{n\to\infty} c_n v, v' \right\rangle = \lim_{n\to\infty} c_n \langle v, v'\rangle = c\langle v, v'\rangle.$$

(Números complexos) Seja $c \in \mathbb{R}^2$, $c = c_0 + c_1\mathrm{i}$. Como $\mathrm{i}^{-1} = -\mathrm{i}$,

$$\begin{aligned}
\langle \mathrm{i}v, v'\rangle &= \frac{1}{4}\mathop{+}_{k=0}^{3} \left\|\mathrm{i}v + \mathrm{i}^k v'\right\|^2 \mathrm{i}^k \\
&= \frac{1}{4}\mathop{+}_{k=0}^{3} \left\|\mathrm{i}(v + \mathrm{i}^{k-1}v')\right\|^2 \mathrm{i}^k \\
&= \frac{1}{4}\mathop{+}_{k=0}^{3} |\mathrm{i}|^2 \left\|v + \mathrm{i}^{k-1}v'\right\|^2 \mathrm{i}^k \\
&= \mathrm{i}\frac{1}{4}\mathop{+}_{k=0}^{3} \left\|v + \mathrm{i}^{k-1}v'\right\|^2 \mathrm{i}^{k-1} \\
&= \mathrm{i}\langle v, v'\rangle.
\end{aligned}$$

Segue da aditividade, da homogeneidade para números reais e da homogeneidade para i que

$$\begin{aligned}
\langle cv, v'\rangle &= \langle (c_0 + c_1\mathrm{i})v, v'\rangle \\
&= \langle c_0 v + c_1\mathrm{i}v, v'\rangle \\
&= c_0\langle v, v'\rangle + c_1\mathrm{i}\langle v, v'\rangle \\
&= (c_0 + c_1\mathrm{i})\langle v, v'\rangle \\
&= c\langle v, v'\rangle.
\end{aligned}$$

2. (Simetria conjugada) Sejam $v, v' \in V$. Segue direto da simetria da adição e de $\mathrm{i}^{-1} = -\mathrm{i}$ que

$$\begin{aligned}
\langle v', v\rangle^* &= \frac{1}{4}\overset{3}{\underset{k=0}{+}} \left\|v' + \mathrm{i}^k v\right\|^2 \mathrm{i}^{k^*} \\
&= \frac{1}{4}\overset{3}{\underset{k=0}{+}} \left\|v' + \mathrm{i}^k v\right\|^2 \mathrm{i}^{k*} \\
&= \frac{1}{4}(\|v' + v\|^2 (1) + \|v' + \mathrm{i}v\|^2 (-\mathrm{i}) + \|v' - v\|^2 (-1) + \|v' - \mathrm{i}v\|^2 \mathrm{i}) \\
&= \frac{1}{4}(\|v + v'\|^2 - |\mathrm{i}|^2 \|v - \mathrm{i}v'\|^2 \mathrm{i} - |-1|^2 \|v - v'\|^2 + |-\mathrm{i}|^2 \|v + \mathrm{i}v'\|^2 \mathrm{i}) \\
&= \langle v, v'\rangle .
\end{aligned}$$

3. (Positividade) Seja $v \in V$. Segue da positividade da norma que

$$\langle v, v\rangle = \|v\|^2 \geq 0.$$

4. (Definição positiva) Seja $v \in V$ tal que $\langle v, v\rangle = 0$. Então

$$0 = \langle v, v\rangle = \|v\|^2 ,$$

e segue da separação da norma que $v = 0$. ■

20.2.2 Perpendicularidade e paralelismo

:⊢ **Definição 20.4.** Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$. Vetores *paralelos* são vetores $v, v' \in V$ para os quais existe $c \in C \setminus \{0\}$ satisfazendo $v' = cv$. Denota-se $v \parallel v'$.

Vetores *perpendiculares* (ou *ortogonais*) são vetores $v, v' \in V$ que satisfazem $\langle v, v'\rangle = 0$. Denota-se $v \perp v'$.

Um conjunto *perpendicular* (ou *ortogonal*) é um conjunto $U \subseteq V$ tal que, para todos $u, u' \in U$, $u \perp u'$, e um conjunto *ortonormal* é um conjunto ortogonal U tal que, para todo $u \in U$, $\|v\| = 1$. O *complemento perpendicular* (ou *ortogonal*) de um conjunto $U \subseteq V$ é o conjunto

$$U^\perp := \{v \in V \mid \forall_{u\in U}\ v \perp u\} .$$

⊢ **Proposição 20.5** (Propriedades de $\parallel$ e $\perp$)**.** *Seja* $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ *um espaço com produto interno sobre um corpo* $\boldsymbol{C} \subseteq \mathbb{R}^2$.

1. *A relação de paralelismo* $\parallel$ *é uma equivalência;*
2. *A relação de perpendicularidade* $\perp$ *é simétrica;*

3. *Para todos* $u, v, v' \in V \smallsetminus \{0\}$*, se* $v \parallel v'$ *e* $v \perp u$*, então* $v' \perp u$*.*
4. *Para todos* $v, v' \in V \smallsetminus \{0\}$*,* $v \parallel v'$ *se, e somente se,* $\{v, v'\}$ *é linearmente dependente.*
5. *Para todos* $v, v' \in V \smallsetminus \{0\}$*, se* $v \perp v'$*, então* $\{v, v'\}$ *é linearmente independente.*
6. *Para todo* $v \in V$*, se* $v \parallel 0$ *então* $v = 0$*; para todo* $v \in V$*,* $v \perp 0$*.*

⊢ **Proposição 20.6.** *Seja* $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ *um espaço com produto interno sobre um corpo* $\boldsymbol{C} \subseteq \mathbb{R}^2$*.*

1. *Para todo* $U \subseteq V$*,* $U^\perp$ *é um subespaço linear de* V*.*
2. $V^\perp = \{0\}$*.*
3. *Para todo subespaço linear* $U \subseteq V$*,* $(U^\perp)^\perp = U$*.*

□ *Demonstração.* Primeiro notamos que $0 \in U^\perp$, pois para todo $u \in U$, $0 \perp u$. Segundo, sejam $v, v' \in U^\perp$ e $c \in C$. Então, para todo $u \in U$, $\langle v, u \rangle = \langle v', u \rangle = 0$, logo

$$\langle cv + v', u \rangle = c \langle v, u \rangle + \langle v', u \rangle = 0,$$

o que mostra que $cv + v' \in U^\perp$. ■

:⊢ **Definição 20.5.** Seja $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno. *Subespaços perpendiculares* em $\boldsymbol{V}$ são subespaços lineares $U, U' \subseteq V$ tais, para todos $u \in U$ e $u' \in U'$, $u \perp u'$. Denota-se $U \perp U'$.

20.2.3 Projeções paralela e perpendicular

:⊢ **Definição 20.6.** Sejam $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ e $u \in V$. A *projeção paralela*[3] de V sobre u é

$$\begin{aligned} \mathrm{p}_{\parallel u} \colon V &\longrightarrow V \\ v &\longmapsto \begin{cases} \dfrac{\langle v, u \rangle}{\|u\|^2} u, & u \neq 0 \\ 0, & u = 0. \end{cases} \end{aligned}$$

A *projeção perpendicular* de V sobre u é

$$\begin{aligned} \mathrm{p}_{\perp u} \colon V &\longrightarrow V \\ v &\longmapsto v - \mathrm{p}_{\parallel u}(v). \end{aligned}$$

[3]Essa projeção é conhecida como *projeção ortogonal* de V sobre u, mas aqui adotaremos as nomenclatura de paralela, já que definimos também a projeção perpendicular, e essa pode ser confundida com a ortogonal.

⊢ **Proposição 20.7** (Propriedades das projeções)**.** *Seja* $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ *um espaço com produto interno sobre um corpo* $\boldsymbol{C} \subseteq \mathbb{R}^2$*.*

1. $\mathrm{p}_{\|0} = 0$ *e* $\mathrm{p}_{\perp 0} = \mathrm{I}$*;*
2. *Para todo* $u \in V$*, as projeções* $\mathrm{p}_{\|u} \colon V \to V$ *e* $\mathrm{p}_{\perp u} \colon V \to V$ *são projeções lineares.*
3. *Para todos* $u, v \in V$*,*
   3.1. *se* $u \not\perp v$*, então* $\mathrm{p}_{\|u}(v) \parallel u$*;*
   3.2. $\mathrm{p}_{\perp u}(v) \perp u$*;*
   3.3. *se* $v \parallel u$*, então* $\mathrm{p}_{\|u}(v) = v$ *(ou seja,* $\mathrm{p}_{\|u}|_{\langle u\rangle} = \mathrm{I}$*);*
   3.4. *se* $v \perp u$*, então* $\mathrm{p}_{\perp u}(v) = v$ *(ou seja,* $\mathrm{p}_{\perp u}|_{\langle u\rangle^\perp} = \mathrm{I}$*);*
4. *Para todos* $u, v \in V \smallsetminus \{0\}$*,* $\{u, v\}$ *é linearmente independente se, e somente se,* $\mathrm{p}_{\perp u}(v) \neq 0$*.*

□ *Demonstração.* 1. Direto da definição.

2. O caso em que $u = 0$ é consequência do item anterior, pois 0 e I são lineares e idempotentes. Consideremos $u \in V \smallsetminus \{0\}$.
   (Linearidade) Sejam $v, v' \in V$ e $c \in C$.

$$\mathrm{p}_{\|u}(cv + v') = \frac{\langle cv + v', u\rangle}{\|u\|^2} u = \frac{c\langle v, u\rangle + \langle v', u\rangle}{\|u\|^2} u = c\mathrm{p}_{\|u}(v) + \mathrm{p}_{\|u}(v').$$

   (Idempotência) Seja $v \in V$.

$$\mathrm{p}_{\|u}\left(\mathrm{p}_{\|u}(v)\right) = \frac{\left\langle \frac{\langle v,u\rangle}{\|u\|^2} u, u\right\rangle}{\|u\|^2} u = \frac{\langle v, u\rangle}{\|u\|^2} \frac{\langle u, u\rangle}{\|u\|^2} u = \frac{\langle v, u\rangle}{\|u\|^2} u = \mathrm{p}_{\|u}(v).$$

   Segue direto da definição $\mathrm{p}_{\perp u} = \mathrm{I} - \mathrm{p}_{\|u}$ e de $\mathrm{p}_{\|u}$ ser projeção linear.

3. 3.1. Se $u \not\perp v$, $\langle u, v\rangle \neq 0$, o que implica que $u \neq 0$ e $\frac{\langle v,u\rangle}{\|u\|^2} \neq 0$. Como por definição $\mathrm{p}_{\|u}(v) = \frac{\langle v,u\rangle}{\|u\|^2} u$, segue que $\mathrm{p}_{\|u}(v) \parallel u$.
   3.2. Se $u = 0$, $\mathrm{p}_{\perp u}(v) \perp u$. Se $u \neq 0$,

$$\left\langle \mathrm{p}_{\|u}(v), u\right\rangle = \left\langle \frac{\langle v, u\rangle}{\|u\|^2} u, u\right\rangle = \frac{\langle v, u\rangle}{\|u\|^2} \langle u, u\rangle = \langle v, u\rangle,$$

   portanto

$$\langle \mathrm{p}_{\perp u}(v), u\rangle = \left\langle v - \mathrm{p}_{\|u}(v), u\right\rangle = \langle v, u\rangle - \left\langle \mathrm{p}_{\|u}(v), u\right\rangle = 0,$$

   o que mostra que $\mathrm{p}_{\perp u}(v) \perp u$.

3.3. Se $v \parallel u$, existe $c \in C \smallsetminus \{0\}$ tal que $v = cu$. Se $u = 0$, então $v = 0$, logo $\mathrm{p}_{\parallel u}(v) = \mathrm{p}_{\parallel 0}(0) = 0 = v$. Se $u \neq 0$, então

$$\mathrm{p}_{\parallel u}(v) = \frac{\langle v, u\rangle}{\|u\|^2} u = \frac{\langle cu, u\rangle}{\|u\|^2} u = cu = v.$$

3.4. Se $v \perp u$, então $\langle u, v\rangle = 0$. Se $u = 0$, então $\mathrm{p}_{\perp u} = \mathrm{I}$, logo $\mathrm{p}_{\perp u}(v) = v$. Se $u \neq 0$,

$$\mathrm{p}_{\perp u}(v) = v - \frac{\langle v, u\rangle}{\|u\|^2} = v.$$

■

Todo v pode ser decomposto como $v = \mathrm{p}_{\parallel u}(v) + \mathrm{p}_{\perp u}(v)$.

20.2.3.1 Processo de ortogonalização

Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno sobre um corpo $\boldsymbol{C} \subseteq \mathbb{R}^2$ e $B = (v_i)_{i\in[d]}$ uma base finita de $\boldsymbol{V}$. A base $(u_i)_{i\in[d]}$, definida recursivamente por

$$u_i := \mathrm{p}_{\perp u_{i-1}} \circ \cdots \circ \mathrm{p}_{\perp u_0}(v_i) = v_i - \mathop{\bigplus}_{j\in[i]} \mathrm{p}_{\parallel u_j}(v_i)$$

é uma base ortogonal (perpendicular) de $\boldsymbol{V}$. A base $(e_i)_{i\in[d]}$ definida por $e_i := \frac{u_i}{\|u_i\|}$ é uma base ortonormal (ou perpendicular unitária) de $\boldsymbol{V}$.

20.2.4 Projeções ortogonais

Lembremos que, para toda projeção linear $p\colon V \to V$, vale que $(\mathrm{I} - p)$ é projeção linear, $p^{-1}(0) = (\mathrm{I} - p)(V)$ e

$$V = p(V) \oplus (\mathrm{I} - p)(V).$$

:⊢ **Definição 20.7.** Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno. Uma *projeção ortogonal* em $\boldsymbol{V}$ é uma projeção linear $p\colon V \to V$ tal que $p(V) \perp (\mathrm{I} - p)(V)$.

⊢ **Proposição 20.8.** *Sejam $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno (completo) e $p\colon V \to V$ uma projeção ortogonal em $\boldsymbol{V}$.*

1. *Para todos $v, v' \in V$,*

$$\langle p(v), v' - p(v')\rangle = \langle v - p(v), p(v')\rangle = 0;$$

2. *Para todos $v, v' \in V$,*

$$\langle p(v), v'\rangle = \langle p(v), p(v')\rangle = \langle v, p(v')\rangle\,;$$

⊢ **Proposição 20.9.** *Sejam* $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ *um espaço com produto interno completo e* $U \subseteq V$ *um subespaço linear fechado. Existe projeção ortogonal* $p\colon V \to V$ *tal que* $p(V) = U$.

□ *Demonstração.* Como U é fechado, é um subespaço completo. Seja $v \in V$. O conjunto

$$N := \{\|v - u\| \mid u \in U\}$$

tem ínfimo e, como U é completo, N tem mínimo. Definimos $p(v)$ como o ponto $u \in U$ tal que $\|v - u\| = \min N$. Mostremos que p é uma projeção ortogonal.

(Idempotência) Seja $v \in V$. Então, como $p(v) \in U$, segue que $p(p(v)) = p(v) \in U$, pois $\|p(v) - p(v)\| = 0$, logo $p^2 = p$.

(Ortogonalidade) Sejam $u \in U = p(V)$ e $v \in (\mathrm{I} - p)(V)$. Se $u = 0$, então $\langle v, u\rangle = \langle v, 0\rangle = 0$. Suponhamos $u \neq 0$. Temos que

$$\begin{aligned}
\left\|v - \mathrm{p}_{\|u}(v)\right\|^2 &= \|v\|^2 - 2\left\langle v, \mathrm{p}_{\|}(v)u\right\rangle + \left\|\mathrm{p}_{\|u}(v)\right\|^2 \\
&= \|v\|^2 - 2\frac{\langle v, u\rangle}{\|u\|^2}\langle v, u\rangle + \frac{\langle v, u\rangle^2}{\|u\|^4}\|u\|^2 \\
&= \|v\|^2 - 2\frac{\langle v, u\rangle^2}{\|u\|^2} + \frac{\langle v, u\rangle^2}{\|u\|^2} \\
&= \|v\|^2 - \frac{\langle v, u\rangle^2}{\|u\|^2}.
\end{aligned}$$

Seja $v' \in V$ tal que $(\mathrm{I} - p)(v') = v$. Então $(p(v') + \mathrm{p}_{\|u}(v)) \in U$, logo pela minimalidade de $p(v')$ segue que

$$\begin{aligned}
\|v\|^2 &= \|v' - p(v')\|^2 \\
&\leq \left\|v' - (p(v') + \mathrm{p}_{\|u}(v))\right\|^2 \\
&= \left\|v - \mathrm{p}_{\|u}(v)\right\|^2 \\
&= \|v\|^2 - \frac{\langle v, u\rangle^2}{\|u\|^2},
\end{aligned}$$

logo $\frac{\langle v,u\rangle^2}{\|u\|^2} = 0$, o que é equivalente a $\langle v, u\rangle = 0$.

(Linearidade) Sejam $v, v' \in V$ e $c \in C$. Então $(cv + v') - p(cv + v')$, $cv - p(cv)$ e $v' - p(v') \in (\mathrm{I} - p)(V)$. Da ortogonalidade de p, para todo $u \in U = p(V)$ temos

$$0 = \langle (cv + v') - p(cv + v'), u\rangle = \langle cv - cp(v), u\rangle = \langle v' - p(v'), u\rangle\,,$$

portanto

$$\begin{aligned}
0 &= \langle (cv + v') - p(cv + v'), u\rangle - \langle cv - cp(v), u\rangle - \langle v' - p(v'), u\rangle \\
&= \langle cp(v) + p(v') - p(cv + v'), u\rangle\,.
\end{aligned}$$

Tomando $u = cp(v) + p(v') - p(cv + v')$, temos que $u \in U$ e então

$$0 = \langle p(cv) + p(v') - p(cv + v'), cp(v) + p(v') - p(cv + v') \rangle ,$$

o que implica $p(cv + v') = cp(v) + p(v')$. ■

⊢ **Proposição 20.10.** *Seja* $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ *um espaço com produto interno. Toda projeção ortogonal* $p \colon V \to V$ *em* $\boldsymbol{V}$ *é contínua.*

□ *Demonstração.* Basta notar que, pela desigualdade do produto interno, para todo $v \in V$,

$$\|p(v)\|^2 = \langle p(v), p(v) \rangle = \langle p(v), v \rangle \leq \|p(v)\| \, \|v\| ,$$

logo $\|p(v)\| \leq \|v\|$. Isso mostra que p é limitada, o que é equivalente a ser contínua. ■

20.2.5 Ângulo

Definiremos agora a noção de ângulo induzida pelo produto interno. Pela desigualdade de Cauchy-Schwarz, sabemos que, para todos $v, v' \in V$,

$$|\langle v, v' \rangle| \leq \|v\| \, \|v'\| .$$

Disso segue que, para todos $v, v' \in V \smallsetminus \{0\}$,

$$\frac{|\langle v, v' \rangle|}{\|v\| \, \|v'\|} \leq 1.$$

Se o produto interno for real, então

$$-1 \leq \frac{\langle v, v' \rangle}{\|v\| \, \|v'\|} \leq 1.$$

Isso significa que a função $\cos^{-1} \colon [-1, 1] \to [0, \tau \diagup 2]$ está definida para esses valores, portanto podemos definir a função ângulo como a seguir. No caso em que o produto interno não é real, ainda se pode definir o ângulo considerando o valor de $\cos^{-1}$ para $\frac{|\langle v, v' \rangle|}{\|v\| \|v'\|} \in [0, 1]$, e com imagem $[0, \tau \diagup 4]$, mas não estudaremos esse caso aqui.

FIGURA 20.1: Gráfico da função $\cos^{-1}\colon [-1,1] \to [0, \frac{\tau}{2}]$.

:⊢ **Definição 20.8.** Sejam $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno real e $v, v' \in V \smallsetminus \{0\}$. O *ângulo* entre v e v' é

$$\sphericalangle(v, v') := \cos^{-1}\left(\frac{\langle v, v' \rangle}{\|v\| \, \|v'\|}\right).$$

A função *ângulo* de $\boldsymbol{V}$ é a função

$$\begin{aligned} \sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} &\longrightarrow \left[0, \frac{\tau}{2}\right] \\ (v, v') &\longmapsto \sphericalangle(v, v'). \end{aligned}$$

FIGURA 20.2: Representação de vetores e o ângulo entre eles.

⊢ **Proposição 20.11** (Propriedades de Ângulo)**.** *Seja* $(\boldsymbol{V}, \langle \cdot, \cdot \rangle)$ *um espaço com produto interno real.*

1. *Para todos* $v, v' \in V \smallsetminus \{0\}$,

$$\langle v, v' \rangle = \|v\| \, \|v'\| \cos(\sphericalangle (v, v'));$$

2. *Para todos* $v, v' \in V \smallsetminus \{0\}$ *e* $cc' \in \mathbb{R} \smallsetminus \{0\}$,

$$\sphericalangle (cv, c'v') = \begin{cases} \sphericalangle (v, v'), & cc' > 0 \\ \dfrac{\tau}{2} - \sphericalangle (v, v'), & cc' < 0; \end{cases}$$

3. *Para todos* $v, v' \in V \smallsetminus \{0\}$,

$$v \parallel v' \quad \Longleftrightarrow \quad \sphericalangle (v, v') \in \{0, \tau \diagup 2\}.$$

Se existe $c \in \left]0, \infty\right[$ *tal que* $v' = cv$, *então* $\sphericalangle (v, v') = 0$, *e se existe* $c \in \left]-\infty, 0\right[$ *tal que* $v' = cv$, *então* $\sphericalangle (v, v') = \frac{\tau}{2}$;

4. *Para todos* $v, v' \in V \smallsetminus \{0\}$,

$$v \perp v' \quad \Longleftrightarrow \quad \sphericalangle (v, v') = \frac{\tau}{4}.$$

□ *Demonstração.* 1. Segue diretamente da definição.

2. Da igualdade

$$\sphericalangle (v, v') = \cos^{-1}\left(\frac{\langle v, v' \rangle}{\|v\| \, \|v'\|}\right)$$

segue que, para todos $c, c' \in \mathbb{R} \smallsetminus \{0\}$

$$\begin{aligned} \sphericalangle (cv, c'v') &= \cos^{-1}\left(\frac{\langle cv, c'v' \rangle}{\|cv\| \, \|c'v'\|}\right) \\ &= \cos^{-1}\left(\frac{cc' \, \langle v, v' \rangle}{|c| \, |c'| \, \|v\| \, \|v'\|}\right) \\ &= \cos^{-1}\left(\frac{cc'}{|cc'|} \frac{\langle v, v' \rangle}{\|v\| \, \|v'\|}\right). \end{aligned}$$

Nesse caso, se $cc' > 0$, então

$$\sphericalangle (cv, c'v') = \cos^{-1}\left(1 \frac{\langle v, v' \rangle}{\|v\| \, \|v'\|}\right) = \sphericalangle (v, v'),$$

e, se $cc' < 0$, então

$$\sphericalangle (cv, c'v') = \cos^{-1}\left(-1 \frac{\langle v, v' \rangle}{\|v\| \, \|v'\|}\right) = \frac{\tau}{2} - \sphericalangle (v, v').$$

3. Notemos que, para todo $v \in V \setminus \{0\}$,

$$\sphericalangle(v,v) = \cos^{-1}\left(\frac{\langle v,v\rangle}{\|v\|\,\|v\|}\right) = \cos^{-1}(1) = 0.$$

Notemos também que $\cos^{-1}\colon [-1,1] \to [0,\tau/2]$ é uma bijeção tal que $\cos^{-1}(1) = 0$ e $\cos^{-1}(-1) = \tau/2$.
Suponhamos agora que existe $c \in \mathbb{R} \setminus \{0\}$ tal que $v' = cv$. Nesse caso, pelo item anterior segue que, relacionando $\sphericalangle(v,v')$ com $\sphericalangle(v,v)$,

$$\sphericalangle(v,v') = \sphericalangle(v,cv) = \begin{cases} 0, & c > 0 \\ \dfrac{\tau}{2}, & c < 0. \end{cases}$$

Reciprocamente, suponhamos que $\sphericalangle(v,v') = 0$. Da bijetividade de $\cos^{-1}$ segue que $\frac{\langle v,v'\rangle}{\|v\|\|v'\|} = 1$, logo $|\langle v,v'\rangle| = \|v\|\,\|v'\|$ e pela desigualdade do produto interno existe $c \in \mathbb{R} \setminus \{0\}$ tal que $v' = cv$. Como $0 = \sphericalangle(v,v') = \sphericalangle(v,cv)$, $c \in]0,\infty[$. Suponhamos então que $\sphericalangle(v,v') = \frac{\tau}{2}$. Da bijetividade de $\cos^{-1}$ segue que $\frac{\langle v,v'\rangle}{\|v\|\|v'\|} = -1$, logo $|\langle v,v'\rangle| = \|v\|\,\|v'\|$ e pela desigualdade do produto interno existe $c \in \mathbb{R} \setminus \{0\}$ tal que $v' = cv$. Como $\frac{\tau}{2} = \sphericalangle(v,v') = \sphericalangle(v,cv)$, $c \in]-\infty,0[$.
4. Segue diretamente da bijetividade de $\cos^{-1}$, pois $\sphericalangle(v,v') = \frac{\tau}{4}$ se, e somente se, $\frac{\langle v,v\rangle}{\|v\|\|v\|} = 0$, o que ocorre se, e somente se, $\langle v,v'\rangle = 0$. ■

:⊢ **Definição 20.9.** Seja $d \in \mathbb{N}$. A *métrica esférica* sobre $\mathbb{S}^d$ é a função

$$\begin{aligned} |\cdot,\cdot|_{\mathbb{S}^d}\colon \mathbb{S}^d &\longrightarrow \mathbb{S}^d \\ (x,x') &\longmapsto \sphericalangle(x,x') = \cos^{-1}(\langle x,x'\rangle). \end{aligned}$$

⊢ **Proposição 20.12.** *A métrica esférica é uma métrica sobre* $\mathbb{S}^d$.

□ *Demonstração.* 1. (Separação) Para todo $x \in \mathbb{S}^d$, $\langle x,x\rangle = \|x\|^2 = 1$, portanto $\sphericalangle(x,x) = 0$.
2. (Simetria) Para todos $x,x' \in \mathbb{S}^d$, segue da simetria de $\langle\cdot,\cdot\rangle$ que

$$\sphericalangle(x,x') = \cos^{-1}(\langle x,x'\rangle) = \cos^{-1}(\langle x',x\rangle) = \sphericalangle(x',x).$$

3. (Desigualdade triangular) Sejam $x,x',x'' \in \mathbb{S}^d$. Para simplificar a notação, denotemos

$$\begin{aligned} \theta_{xx'} &:= \sphericalangle(x,x') \\ \theta_{x'x''} &:= \sphericalangle(x',x'') \\ \theta_{xx''} &:= \sphericalangle(x,x'') \end{aligned}$$

A matriz de produtos internos

$$\begin{aligned} G(x,x',x'') &= \begin{pmatrix} \langle x,x\rangle & \langle x,x'\rangle & \langle x,x''\rangle \\ \langle x',x\rangle & \langle x',x'\rangle & \langle x',x''\rangle \\ \langle x'',x\rangle & \langle x'',x'\rangle & \langle x'',x''\rangle \end{pmatrix} \\ &= \begin{pmatrix} 1 & \cos(\theta_{xx'}) & \cos(\theta_{xx''}) \\ \cos(\theta_{xx'}) & 1 & \cos(\theta_{x'x''}) \\ \cos(\theta_{xx''}) & \cos(\theta_{x'x''}) & 1 \end{pmatrix} \end{aligned}$$

tem determinante positivo, portanto

$$1-\cos(\theta_{xx'})^2-\cos(\theta_{x'x''})^2-\cos(\theta_{xx''})^2+2\cos(\theta_{xx'})\cos(\theta_{x'x''})\cos(\theta_{xx''})\geq 0.$$

Reorganizando os ternos e somando $(\cos(\theta_{xx'})\cos(\theta_{x'x''}))^2$ a ambos os lados, obtém-se

$$\begin{aligned} &1-\cos(\theta_{xx'})^2-\cos(\theta_{x'x''})^2+(\cos(\theta_{xx'})\cos(\theta_{x'x''}))^2 \\ &\quad\geq (\cos(\theta_{xx'})\cos(\theta_{x'x''}))^2-2\cos(\theta_{xx'})\cos(\theta_{x'x''})\cos(\theta_{xx''})+\cos(\theta_{xx''})^2. \end{aligned}$$

Por um lado, vale que

$$\begin{aligned} (\operatorname{sen}(\theta_{xx'})\operatorname{sen}(\theta_{x'x''}))^2 &= \operatorname{sen}(\theta_{xx'})^2\operatorname{sen}(\theta_{x'x''})^2 \\ &= (1-\cos(\theta_{xx'})^2)(1-\cos(\theta_{x'x''})^2) \\ &= 1-\cos(\theta_{xx'})^2-\cos(\theta_{x'x''})^2+(\cos(\theta_{xx'})\cos(\theta_{x'x''}))^2 \end{aligned}$$

e, por outro,

$$\begin{aligned} &(\cos(\theta_{xx'})\cos(\theta_{x'x''})-\cos(\theta_{xx''}))^2 \\ &\quad= (\cos(\theta_{xx'})\cos(\theta_{x'x''}))^2-2\cos(\theta_{xx'})\cos(\theta_{x'x''})\cos(\theta_{xx''})+\cos(\theta_{xx''})^2. \end{aligned}$$

Dessas igualdades, segue que

$$(\operatorname{sen}(\theta_{xx'})\operatorname{sen}(\theta_{x'x''}))^2\geq(\cos(\theta_{xx'})\cos(\theta_{x'x''})-\cos(\theta_{xx''}))^2.$$

Como $\theta_{xx'},\theta_{x'x''}\in[0,\tau\diagup 2]$, então $\operatorname{sen}(\theta_{xx'})\geq 0$ e $\operatorname{sen}(\theta_{x'x''})\geq 0$, e portanto

$$\operatorname{sen}(\theta_{xx'})\operatorname{sen}(\theta_{x'x''})\geq\cos(\theta_{xx'})\cos(\theta_{x'x''})-\cos(\theta_{xx''}).$$

Reorganizando essa desigualdade, segue que

$$\begin{aligned} \cos(\theta_{xx''}) &\geq \cos(\theta_{xx'})\cos(\theta_{x'x''})-\operatorname{sen}(\theta_{xx'})\operatorname{sen}(\theta_{x'x''}) \\ &= \cos(\theta_{xx'}+\theta_{x'x''}). \end{aligned}$$

Por fim, como $\cos^{-1}$ é decrescente, conclui-se

$$\begin{aligned} \theta_{xx''} &= \cos^{-1}(\cos(\theta_{xx''})) \\ &\leq \cos^{-1}(\cos(\theta_{xx'}+\theta x'x'')) \\ &= \theta_{xx'}+\theta_{x'x''}, \end{aligned}$$

ou seja,

$$\sphericalangle(x,x'')\leq\sphericalangle(x,x')+\sphericalangle(x',x''). \qquad \blacksquare$$

20.2.5.1 Definição de uma função ângulo

Tentamos, nesta seção, definir uma função ângulo em um espaço normado de forma que possamos, a partir dela e da norma, definir um produto interno. A investigação consiste em destacar as propriedades que uma função ângulo deve satisfazer. A ideia é que ela seja uma função invariante pela semirreta positiva gerada por um vetor, ou seja, que não mude de valor se multiplicarmos um vetor por um número positivo. Além disso, a função deve também ser uma distância nesse espaço de semirretas. Assumiremos aqui que o espaço normado é real.

Primeira definição: uma definição ad hoc

:⊢ **Definição 20.10.** Seja $(\boldsymbol{V}, \|\cdot\|)$ um espaço normado. Uma *função ângulo* em $(\boldsymbol{V}, \|\cdot\|)$ é uma função contínua $\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau \diagup 2]$ tal que

1. (Nulidade) Para todos $v \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v) = 0;$$
2. (Simetria) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v') = \sphericalangle(v', v);$$
3. (Compatibilidade) Para todos $v, v', v'' \in V \smallsetminus \{0\}$ tais que $v' \neq -v$,
$$\|v + v'\| \cos(\sphericalangle(v + v', v'')) = \|v\| \cos(\sphericalangle(v, v'')) + \|v'\| \cos(\sphericalangle(v', v''));$$

Um *espaço angulado* é uma tripla $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ em que $(\boldsymbol{V}, \|\cdot\|)$ é um espaço normado e $\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau \diagup 2]$ uma função ângulo em $(\boldsymbol{V}, \|\cdot\|)$.

⊢ **Proposição 20.13.** *Seja $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ um espaço angulado.*

1. *(Homogeneidade) Para todos $v, v' \in V \smallsetminus \{0\}$ e $c \in]0, \infty[$,*
$$\sphericalangle(cv, v') = \sphericalangle(v, v');$$
2. *(Suplementação) Para todos $v, v' \in V \smallsetminus \{0\}$,*
$$\sphericalangle(v, v') + \sphericalangle(-v, v') = \tau \diagup 2;$$

□ *Demonstração.* 1. A homogeneidade vale para $\mathbb{N} \smallsetminus \{0\}$. Mostremos por indução. Para $n = 1$, claramente vale. Suponhamos que vale para $n \in \mathbb{N} \smallsetminus \{0\}$. Da compatibilidade segue que

$$\begin{aligned}
\|(n+1)v\| \cos(\sphericalangle((n+1)v, v')) &= \|nv\| \cos(\sphericalangle(nv, v')) + \|v\| \cos(\sphericalangle(v, v')) \\
&= n\|v\| \cos(\sphericalangle(v, v')) + \|v\| \cos(\sphericalangle(v, v')) \\
&= (n+1)\|v\| \cos(\sphericalangle(v, v')) \\
&= \|(n+1)v\| \cos(\sphericalangle(v, v')),
\end{aligned}$$

portanto
$$\cos(\sphericalangle((n+1)v, v')) = \cos(\sphericalangle(v, v'));$$
como $\sphericalangle((n+1)v, v'), \sphericalangle(v, v') \in [0, \tau \diagup 2]$, segue que
$$\sphericalangle((n+1)v, v') = \sphericalangle(v, v').$$

A homogeneidade vale para os números racionais estritamente positivos. Seja $\frac{n}{d} \in \mathbb{Q} \cap]0, \infty[$. Da homogeneidade para números naturais positivos segue que
$$\sphericalangle\left(\frac{n}{d}v, v'\right) = \sphericalangle\left(d\frac{n}{d}v, v'\right) = \sphericalangle(nv, v') = \sphericalangle(v, v').$$
A homogeneidade vale para os números reais positivos. Seja $c \in]0, \infty[$ e $(c_n)_{n\in\mathbb{N}}$ uma sequência de números racionais estritamente positivos tais que $\lim_{n\to\infty} c_n = c$. Então, da continuidade do ângulo, segue que
$$\sphericalangle(cv, v') = \sphericalangle\left(\lim_{n\to\infty} c_n v, v'\right) = \lim_{n\to\infty} \sphericalangle(c_n v, v') = \sphericalangle(v, v').$$

2. Da compatibilidade e homogeneidade segue que
$$\begin{aligned}\|v\| \cos(\sphericalangle(v, v')) &= \|2v - v\| \cos(\sphericalangle(2v - v, v')) \\ &= \|2v\| \cos(\sphericalangle(2v, v')) + \|-v\| \cos(\sphericalangle(-v, v')) \\ &= 2\|v\| \cos(\sphericalangle(v, v')) + \|v\| \cos(\sphericalangle(-v, v'));\end{aligned}$$
disso e da fórmula do cosseno de ângulo suplementar segue que
$$\cos(\sphericalangle(-v, v')) = -\cos(\sphericalangle(v, v')) = \cos(\tau \diagup 2 - \sphericalangle(v, v')),$$
portanto
$$\sphericalangle(-v, v') = \tau \diagup 2 - \sphericalangle(v, v').$$

■

⊢ **Proposição 20.14.** *Seja* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado. A função*
$$\begin{aligned}\langle\cdot,\cdot\rangle \colon V \times V &\longrightarrow \mathbb{R} \\ (v, v') &\longmapsto \begin{cases} 0, & v = 0 \text{ ou } v' = 0 \\ \|v\|\,\|v'\| \cos(\sphericalangle(v, v')), & v \neq 0 \text{ e } v' \neq 0. \end{cases}\end{aligned}$$
é um produto interno em $\boldsymbol{V}$.

□ *Demonstração.* Os casos em que $v = 0$ ou $v' = 0$ são todos imediatos, portanto assumiremos que os vetores são sempre não nulos.

1. (Linearidade na primeira entrada)

   1.1. (Aditividade) Sejam $v, v', v'' \in V$. Queremos mostrar que

$$\langle v + v', v'' \rangle = \langle v, v'' \rangle + \langle v', v'' \rangle .$$

   Como supomos que v'' é não nulo, isso é equivalente à compatibilidade

$$\|v + v'\| \cos(\sphericalangle (v + v', v'')) = \|v\| \cos(\sphericalangle (v, v'')) + \|v'\| \cos(\sphericalangle (v', v'')).$$

   1.2. (Homogeneidade) Sejam $v, v' \in V$ e $c \in \mathbb{R}$. Separamos em caos.

   1.2.1. ($c = 0$) Esse caso é trivial.

   1.2.2. ($c > 0$) Segue da homogeneidade do ângulo que

$$\begin{aligned} \langle cv, v' \rangle &= \|cv\| \, \|v'\| \cos(\sphericalangle (cv, v')) \\ &= c \, \|v\| \, \|v'\| \cos(\sphericalangle (v, v')) \\ &= c \, \langle v, v' \rangle . \end{aligned}$$

   1.2.3. ($c < 0$) Segue das propriedades de homogeneidade e suplementação do ângulo e da fórmula do cosseno de ângulo suplementar que

$$\begin{aligned} \langle cv, v' \rangle &= \|cv\| \, \|v'\| \cos(\sphericalangle (cv, v')) \\ &= -c \, \|v\| \, \|v'\| \cos(\sphericalangle (-|c| \, v, v')) \\ &= c \, \|v\| \, \|v'\| \, (-\cos(\sphericalangle (-v, v'))) \\ &= c \, \|v\| \, \|v'\| \, (-\cos(\tau \diagup 2 - \sphericalangle (v, v'))) \\ &= c \, \|v\| \, \|v'\| \cos(\sphericalangle (v, v')) \\ &= c \, \langle v, v' \rangle . \end{aligned}$$

2. (Simetria) Sejam $v, v' \in V$. Segue da simetria do produto e da simetria do ângulo que

$$\langle v, v' \rangle = \|v\| \, \|v'\| \cos(\sphericalangle (v, v')) = \|v'\| \, \|v\| \cos(\sphericalangle (v', v)) = \langle v', v \rangle .$$

3. (Positividade) Seja $v \in V$. Segue da nulidade do ângulo e da positividade da norma que

$$\langle v, v \rangle = \|v\| \, \|v\| \cos(\sphericalangle (v, v)) = \|v\|^2 \cos(0) = \|v\|^2 \in [0, \infty[.$$

4. (Definição positiva) Seja $v \in V$. Se $v \neq 0$, então $\|v\| \neq 0$ e, pelo item anterior, $\langle v, v \rangle = \|v\|^2$, portanto $\langle v, v \rangle \neq 0$.

■

Segunda definição: investigação de mais propriedades Para começar a discussão, listamos algumas propriedades interessantes e esperadas de uma função ângulo, tanto por si própria como em relação à norma. Consideramos um espaço normado $(\boldsymbol{V}, \|\cdot\|)$ e uma função $\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau \diagup 2]$

1. (0-Homogeneidade) Para todos $v, v' \in V$ e $c, c' \in \,]0, \infty[$,
$$\sphericalangle\,(cv, c'v') = \sphericalangle\,(v, v')\,;$$
2. (Separação) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, v') = 0$$
se, e somente se, existe $c \in \,]0, \infty[$ tal que $v' = cv$;
3. (Simetria) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, v') = \sphericalangle\,(v', v)\,;$$
4. (Desigualdade Triangular) Para todos $v, v', v'' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, v'') \leq \sphericalangle\,(v, v') + \sphericalangle\,(v', v'')\,;$$
5. (Aditividade) Para todos $v, v' \in V \smallsetminus \{0\}$ e $c, c' \in [0, \infty[$ tais que $cv + c'v' \neq 0$,
$$\sphericalangle\,(v, v') = \sphericalangle\,(v, cv + c'v') + \sphericalangle\,(cv + c'v', v')\,;$$
6. (Aditividade limite) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, -v) = \sphericalangle\,(v, v') + \sphericalangle\,(v', -v)\,;$$
7. (Isotropia) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, -v) = \sphericalangle\,(v', -v')$$
8. (Radiano) Para todo $v \in V \smallsetminus \{0\}$
$$\sphericalangle\,(v, -v) = \tau \diagup 2;$$
9. (Retificação) Para todos $v, v', w, w' \in V \smallsetminus \{0\}$ tais que $\sphericalangle\,(v, v') = \sphericalangle\,(-v, v')$ e $\sphericalangle\,(w, w') = \sphericalangle\,(-w, w')$,
$$\sphericalangle\,(v, v') = \sphericalangle\,(w, w')\,;$$
10. (Suplementação) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle\,(v, v') + \sphericalangle\,(-v, v') = \tau \diagup 2;$$

11. (Lado-ângulo-lado) Para todos $v, v', w, w' \in V \smallsetminus \{0\}$ tais que $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ e $\sphericalangle(v, v') = \sphericalangle(w, w')$, valem
    11.1. $\|v + v'\| = \|w + w'\|$;
    11.2. $\sphericalangle(v, v + v') = \sphericalangle(w, w + w')$.
12. (Lei dos cossenos) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v') = \cos^{-1}\left(\frac{\|v + v'\|^2 - \|v\|^2 - \|v'\|^2}{2\,\|v\|\,\|v'\|}\right)$$
13. (Pitágoras) Para todos $v, v' \in V \smallsetminus \{0\}$, $\sphericalangle(v, v') = \tau \diagup 4$ se, e somente se,
$$\|v + v'\|^2 = \|v\|^2 + \|v'\|^2;$$
14. (Tales) Para todos $v, v' \in V \smallsetminus \{0\}$ tais que $\|v\| = \|v'\|$,
$$\sphericalangle(v + v', v - v') = \tau \diagup 4.$$
15. (Compatibilidade) Para todos $v, v', v'' \in V \smallsetminus \{0\}$ tais que $v' \neq -v$,
$$\|v + v'\| \cos(\sphericalangle(v + v', v'')) = \|v\| \cos(\sphericalangle(v, v'')) + \|v'\| \cos(\sphericalangle(v', v''));$$

A propriedade de retificação é o postulado de Euclides de que todos ângulos retos são iguais, e a isotropia é similar, pois assume que todo ângulo de meia volta é igual. Como em Euclides, essas propriedades não assumem um valor específico de ângulo. Podemos escolher um valor e *normalizar* o ângulo, ou seja, dividir o ângulo por esse valor para que o valor do ângulo de meia volta seja $\tau \diagup 2$ (sendo τ o período[4] de cos, uma volta completa), para que depois possamos calcular o cosseno desse ângulo. A propriedade do radiano é uma forma de normalização e já garante isso, definindo que para ser uma função ângulo o valor de meia volta deve ser $\tau \diagup 2$. Assumindo a aditividade (limite) e a isotropia, podemos mostrar a propriedade de retificação, ou seja, que todos ângulos retos são iguais entre si. Assumindo a propriedade do radiano, podemos mostrar que todo ângulo reto vale $\tau \diagup 4$, um quarto de volta, pois se os ângulos suplementares são iguais e somam $\tau \diagup 2$, cada um deles vale $\tau \diagup 4$. Ainda, da aditividade (limite) e da propriedade do radiano, segue a suplementação. A aditividade limite segue da aditividade assumindo que a função é contínua e homogênea.

$\vdash$ **Proposição 20.15.** *Sejam* $(\boldsymbol{V}, \|\cdot\|)$ *um espaço normado e*
$$\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \infty[$$
uma função contínua que satisfaz 0*-homogeneidade, aditividade e isotropia.*

[4] O número τ vale aproximadamente 6,28 e é o dobro de π, como geralmente é denotada metade do período de cos.

1. *(Aditividade limite) Para todos* $v, v' \in V \smallsetminus \{0\}$,

$$\sphericalangle(v, -v) = \sphericalangle(v, v') + \sphericalangle(v', -v);$$

2. *(Maximalidade) Existe único* $\pi \in [0, \infty[$ *tal que, para todos* $v, v' \in V \smallsetminus \{0\}$,

$$\sphericalangle(v, v') \leq \sphericalangle(v, -v) = \pi;$$

3. *(Retificação) Para todos* $v, v', w, w' \in V \smallsetminus \{0\}$ *tais que* $\sphericalangle(v, v') = \sphericalangle(v', -v)$ *e* $\sphericalangle(w, w') = \sphericalangle(w', -w)$,

$$\sphericalangle(v, v') = \sphericalangle(w, w').$$

4. *(Suplementação) Para todos* $v, v' \in V \smallsetminus \{0\}$,

$$\sphericalangle(v, v') + \sphericalangle(v', -v) = \pi;$$

□ *Demonstração.* 1. Seja $c \in]0, \infty[$ tal que $cv + v' \neq 0$ e $v' - cv \neq 0$. Como $v' = \frac{1}{2}(cv + v') + \frac{1}{2}(v' - cv)$, segue da aditividade do ângulo que

$$\sphericalangle(cv + v', v' - cv) = \sphericalangle(cv + v', v') + \sphericalangle(v', v' - cv).$$

Como

$$\frac{v}{\|v\|} = \lim_{c \to \infty} \frac{cv + v'}{\|cv + v'\|}$$

e

$$\frac{-v}{\|v\|} = \lim_{c \to \infty} \frac{v' - cv}{\|v' - cv\|},$$

segue da 0-homogeneidade e da continuidade do ângulo que

$$\begin{aligned}
\sphericalangle(v, -v) &= \sphericalangle\left(\frac{v}{\|v\|}, \frac{-v}{\|v\|}\right) \\
&= \sphericalangle\left(\lim_{c \to \infty} \frac{cv + v'}{\|cv + v'\|}, \lim_{c \to \infty} \frac{v' - cv}{\|v' - cv\|}\right) \\
&= \lim_{c \to \infty} \sphericalangle\left(\frac{cv + v'}{\|cv + v'\|}, \frac{v' - cv}{\|v' - cv\|}\right) \\
&= \lim_{c \to \infty} \sphericalangle(cv + v', v' - cv) \\
&= \lim_{c \to \infty} \sphericalangle(cv + v', v') + \sphericalangle(v', v' - cv) \\
&= \lim_{c \to \infty} \sphericalangle\left(\frac{cv + v'}{\|cv + v'\|}, v'\right) + \sphericalangle\left(v', \frac{v' - cv}{\|v' - cv\|}\right) \\
&= \sphericalangle(v, v') + \sphericalangle(v', -v).
\end{aligned}$$

2. Definimos $\pi := \sphericalangle(v, -v)$ e a unicidade segue da isotropia, pois $\sphericalangle(v', -v') = \sphericalangle(v, -v) = \pi$. Agora, da positividade temos $\sphericalangle(v', -v) \geq 0$, portanto da aditividade limite segue que

$$\sphericalangle(v, v') \leq \sphericalangle(v, v') + \sphericalangle(v', -v) = \sphericalangle(v, -v) = \pi.$$

3. Da aditividade limite e da isotropia segue que

$$\begin{aligned}
\sphericalangle(v, v') &= \frac{\sphericalangle(v, v') + \sphericalangle(v, v')}{2} \\
&= \frac{\sphericalangle(v, v') + \sphericalangle(v', -v)}{2} \\
&= \frac{\sphericalangle(v, -v)}{2} \\
&= \frac{\sphericalangle(w, -w)}{2} \\
&= \sphericalangle(w, w').
\end{aligned}$$

Segue direto dos itens anteriores. ■

Com essa definição, a propriedade do radiano seria equivalente a termos

$$\pi = \frac{\tau}{2},$$

ou seja, o valor máximo da função $\sphericalangle$, denotado π, é definido como metade do período da função cos, denotado τ. Isso quer dizer que estamos medindo os ângulos em radianos, que são os valores mais naturais para calcular senos e cossenos.

A propriedade lado-ângulo-lado é equivalente a dizer que $\|v + v'\|$ e $\sphericalangle(v, v + v')$ são funções de $\|v\|$, $\|v'\|$ e $\sphericalangle(v, v')$. Essa propriedade é fundamental para relacionar a norma do espaço com a função ângulo e ela que vai, no final, nos permitir mostrar a lei dos cossenos, que essencialmente é equivalente a termos um produto interno. Sem uma propriedade que relacione o ângulo e a norma, não temos porque esperar que tenhamos um produto interno dado pelo produto das normas e do cosseno do ângulo, pois existem normas que não admitem produtos internos, precisamente aquelas que não satisfazem a lei do paralelogramo, o se tivéssemos uma norma dessa, qualquer função ângulo independente de norma, seja ela qual fosse, não nos daria um produto interno.

Após essa discussão, partimos para uma definição da função ângulo que por simplicidade já considera a propriedade de suplementação em vez das propriedades de isotropia e radiano, pois elas são equivalentes.

:⊢ **Definição 20.11.** Seja $(\boldsymbol{V}, \|\cdot\|)$ um espaço normado. Uma *função ângulo* em $(\boldsymbol{V}, \|\cdot\|)$ é uma função contínua $\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau \diagup 2]$ tal que

1. (0-Homogeneidade) Para todos $v, v' \in V$ e $c, c' \in \,]0, \infty[$,
$$\sphericalangle(cv, c'v') = \sphericalangle(v, v');$$
2. (Separação) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v') = 0$$
se, e somente se, existe $c \in \,]0, \infty[$ tal que $v' = cv$;
3. (Simetria) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v') = \sphericalangle(v', v);$$
4. (Aditividade) Para todos $v, v' \in V \smallsetminus \{0\}$ e $c, c' \in [0, \infty[$ tais que $cv + c'v' \neq 0$,
$$\sphericalangle(v, v') = \sphericalangle(v, cv + c'v') + \sphericalangle(cv + c'v', v');$$
5. (Suplementação) Para todos $v, v' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v') + \sphericalangle(v', -v) = \tau \diagup 2;$$
6. (LAL: lado-ângulo-lado) Para todos $v, v', w, w' \in V \smallsetminus \{0\}$ tais que $v + v' \neq 0$, $w + w' \neq 0$, $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ e $\sphericalangle(v, v') = \sphericalangle(w, w')$, valem
   6.1. $\|v + v'\| = \|w + w'\|$;
   6.2. $\sphericalangle(v, v + v') = \sphericalangle(w, w + w')$;

Um *espaço angulado* é uma tripla $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ em que $(\boldsymbol{V}, \|\cdot\|)$ é um espaço normado e $\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau \diagup 2]$ uma função ângulo em $(\boldsymbol{V}, \|\cdot\|)$.

▷ **Exercício 20.2.** *Seja $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ um espaço angulado.*

1. *(Alternos internos) Para todos $v, v' \in V \smallsetminus \{0\}$,*
$$\sphericalangle(v, v') = \sphericalangle(-v, -v');$$
2. *(Ângulos externos) Para todos $v, v' \in V \smallsetminus \{0\}$,*
   2.1. (Adjacente) $\sphericalangle(v', v' - v) < \sphericalangle(v', -v)$;
   2.2. (Oposto) $\sphericalangle(v, v - v') < \sphericalangle(v', -v)$.

⊢ **Proposição 20.16.** *Seja $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ um espaço angulado.*

1. *(LAL negativo) Para todos $v, v', w, w' \in V \smallsetminus \{0\}$ tais que $v - v' \neq 0$, $w - w' \neq 0$, $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ e $\sphericalangle(v, v') = \sphericalangle(w, w')$, valem*

1.1. $\|v - v'\| = \|w - w'\|$;
1.2. $\sphericalangle(v, v - v') = \sphericalangle(w, w - w')$;

2. *(Triângulo isósceles) Para todos* $v, v' \in V \setminus \{0\}$, *tais que* $\|v\| = \|v'\|$,

   2.1. (Positivo) $\sphericalangle(v, v + v') = \sphericalangle(v', v + v')$;
   2.2. (Negativo) $\sphericalangle(v, v - v') = \sphericalangle(v', v' - v)$.

3. *(Oposição lado-ângulo) Para todos* $v, v' \in V \setminus \{0\}$ *tais que* $v' - v \neq 0$,

   3.1. $\|v\| < \|v'\|$ *se, e somente se,* $\sphericalangle(v', v' - v) < \sphericalangle(v, v - v')$;
   3.2. $\|v\| = \|v'\|$ *se, e somente se,* $\sphericalangle(v', v' - v) = \sphericalangle(v, v - v')$.

4. *(Monotonicidade lado-ângulo) Para todos* $v, v', w, w' \in V \setminus \{0\}$ *tais que* $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ *e* $\sphericalangle(v, v') < \sphericalangle(w, w')$,

$$\|v' - v\| < \|w' - w\|.$$

□ *Demonstração.* 1. Como $\|v'\| = \|w'\|$, temos $\|-v'\| = \|-w'\|$; como

$$\sphericalangle(v, v') = \sphericalangle(w, w'),$$

segue da suplementação e da simetria que

$$\sphericalangle(v, -v') = \tau \diagup 2 - \sphericalangle(v, v') = \tau \diagup 2 - \sphericalangle(w, w') = \sphericalangle(w, -w').$$

Segue então de LAL que

$$\|v - v'\| = \|v + (-v')\| = \|w + (-w')\| = \|w - w'\|$$

e

$$\sphericalangle(v, v - v') = \sphericalangle(v, v + (-v')) = \sphericalangle(w, w + (-w')) = \sphericalangle(w, w - w').$$

2. Como $\|v\| = \|v'\|$, $\|v'\| = \|v\|$ e, por simetria do ângulo, $\sphericalangle(v, v') = \sphericalangle(v', v)$, segue de LAL que

$$\sphericalangle(v, v + v') = \sphericalangle(v', v' + v) = \sphericalangle(v', v + v').$$

A outra igualdade segue analogamente usando LAL (negativo).

3. Assumamos que $\|v\| < \|v'\|$. Seja $c := \|v\| \diagup \|v'\| \in]0, 1[$. Como $(1 - c) > 0$ e $v' - v = (1 - c)v' + (cv' - v)$, da homogeneidade e da aditividade do ângulo segue que

$$\begin{aligned}
\sphericalangle(v', v' - v) &= \sphericalangle((1 - c)v', v' - v) \\
&< \sphericalangle((1 - c)v', v' - v) + \sphericalangle(v' - v, cv' - v) \\
&= \sphericalangle((1 - c)v', cv' - v) \\
&= \sphericalangle(cv', cv' - v);
\end{aligned}$$

como $\|cv'\| = c\,\|v'\| = \|v\|$, da proposição do triângulo isósceles (negativo) segue que

$$\sphericalangle\,(cv', cv' - v) = \sphericalangle\,(v, v - cv')\,;$$

como $c > 0$, $(1 - c) > 0$ e $v - cv' = (1 - c)v + c(v - v')$, da aditividade do ângulo segue que

$$\sphericalangle\,(v, v - cv') < \sphericalangle\,(v, v - cv') + \sphericalangle\,(v - cv', v - v') = \sphericalangle\,(v, v - v')\,.$$

Assim, concluímos que

$$\begin{aligned}\sphericalangle\,(v', v' - v) &< \sphericalangle\,(cv', cv' - v)\\ &= \sphericalangle\,(v, v - cv')\\ &< \sphericalangle\,(v, v - v')\,.\end{aligned}$$

Assumindo que $\|v'\| < \|v\|$, por simetria do enunciado segue que

$$\sphericalangle\,(v, v - v') < \sphericalangle\,(v', v - v')\,.$$

Assim, se $\sphericalangle\,(v, v - v') = \sphericalangle\,(v', v - v')$, então $\|v\| = \|v'\|$ e, da proposição do triângulo isósceles segue a recíproca.

4. Como $\sphericalangle\,(v, v') < \sphericalangle\,(w, w')$, tomamos[5] $c, c' \in\,]0, \infty[$ tais que

$$\|v'\| = \|cw + c'w'\|$$

e $\sphericalangle\,(v, v') = \sphericalangle\,(w, cw + c'w')$ e definimos $u := cw + c'w'$. Por LAL (negativo) segue que $\|v - v'\| = \|w - u\|$.
Para estimar a norma de $w - u$, consideramos 3 casos.

- $(c + c' = 1)$ Nesse caso temos, $0 < c' = 1 - c < 1$, logo

$$\begin{aligned}\|v - v'\| &= \|w - u\|\\ &= \|w - (cw + (1 - c)w')\|\\ &= \|(1 - c)(w - w')\|\\ &= |1 - c|\,\|w - w'\|\\ &< \|w - w'\|\,;\end{aligned}$$

- $(c + c' < 1)$ Como $w' - w = (u - w) + (w' - u)$, segue da positividade e da aditividade que

$$\begin{aligned}\sphericalangle\,(w' - w, w' - u) &< \sphericalangle\,(u - w, w' - w) + \sphericalangle\,(w' - w, w' - u)\\ &= \sphericalangle\,(u - w, w' - u)\,;\end{aligned}$$

[5] Isso segue da continuidade da função ângulo $\sphericalangle\,(w, w + kw')$ para achar k que satisfaça a condição do ângulo, e então por normalização para igualar a norma de v'.

como $(1-c-c') \diagup c > 0$, $c' \diagup c > 0$ e

$$\begin{aligned} u - w &= cw + c'w' - w \\ &= (c-1)w + c'w' \\ &= (-(1-c-c') - c')\, w + \left(-\frac{1-c-c'}{c}c' + \frac{c'}{c}(1-c')\right) w' \\ &= \frac{1-c-c'}{c}(-cw - c'w') + \frac{c'}{c}(w' - cw - c'w') \\ &= \frac{1-c-c'}{c}(-u) + \frac{c'}{c}(w' - u), \end{aligned}$$

segue da positividade e da aditividade do ângulo que

$$\begin{aligned} \sphericalangle\,(u-w, w'-u) &< \sphericalangle\,(-u, u-w) + \sphericalangle\,(u-w, w'-u) \\ &= \sphericalangle\,(-u, w'-u)\,; \end{aligned}$$

como $\|-u\| = \|u\| = \|w'\|$, segue da propriedade do triângulo isósceles (positivo) que

$$\sphericalangle\,(-u, w'-u) = \sphericalangle\,(w', w'-u)\,;$$

definimos $k := c' \diagup (1-c) < 1$ e segue que

$$\begin{aligned} (1-c)(kw' - w) &= (1-c)\frac{c'}{1-c}w' - (1-c)w \\ &= (cw + c'w') - w \\ &= u - w; \end{aligned}$$

como $(1-k)w' = (w'-u) - (kw'-u)$, segue da 0-homogeneidade, da positividade e da aditividade e dos ângulos alternos internos que

$$\begin{aligned} \sphericalangle\,(w'-u, w') &= \sphericalangle\,(w'-u, (1-k)w') \\ &= \sphericalangle\,(w'-u, (w'-u) - (kw'-u)) \\ &< \sphericalangle\,(w'-u, -(kw'-u)) \\ &= \sphericalangle\,(-(u-w'), -(u-w)) \\ &= \sphericalangle\,(u-w', u-w)\,. \end{aligned}$$

Assim, concluímos que

$$\begin{aligned} \sphericalangle\,(w'-w, (w'-w) - (u-w)) &= \sphericalangle\,(w'-w, w'-u) \\ &< \sphericalangle\,(u-w, w'-u) \\ &< \sphericalangle\,(-u, w'-u) \\ &= \sphericalangle\,(w', w'-u) \\ &< \sphericalangle\,(u-w, u-w') \\ &= \sphericalangle\,(u-w, (u-w) - (w'-w)) \end{aligned}$$

e da proposição anterior de oposição lado-ângulo segue que

$$\|v' - v\| = \|u - w\| < \|w' - w\|.$$

- $(c + c' > 1)$ Definimos $k := 1 \diagup (c + c')$ e segue que

$$\begin{aligned} w' - ku &= w' - k(cw + c'w') \\ &= (1 - kc')w' - kcw \\ &= \left(\frac{c + c' - c'}{c + c'}\right) w' - kcw \\ &= \left(\frac{c}{c + c'}\right) w' - kcw \\ &= kcw' - kcw \\ &= kc(w' - w). \end{aligned}$$

Como $k > 0$, $1 - k > 0$ e $w' - ku = k(w' - u) + (1 - k)w'$, segue da 0-homogeneidade, da positividade e da aditividade que

$$\begin{aligned} \sphericalangle(w' - w, w' - u) &= \sphericalangle(w' - u, w' - w) \\ &= \sphericalangle(w' - u, w' - ku) \\ &< \sphericalangle(w' - u, w' - ku) + \sphericalangle(w' - ku, w') \\ &= \sphericalangle(w' - u, w'); \end{aligned}$$

como $\|w'\| = \|u\|$, segue da proposição do triângulo isósceles (negativo) e da simetria que

$$\sphericalangle(w' - u, w') = \sphericalangle(u - w', u);$$

como

$$(c + c' - 1)u = cu + c'w' - cw - c'w' = c'(u - w') + c(u - w),$$

segue da positividade e da aditividade que

$$\begin{aligned} \sphericalangle(u - w', u) &< \sphericalangle(u - w', u) + \sphericalangle(u, u - w) \\ &= \sphericalangle(u - w', u - w). \end{aligned}$$

Assim, concluímos que

$$\begin{aligned} \sphericalangle(w' - w, (w' - w) - (u - w)) &= \sphericalangle(w' - w, w' - u) \\ &< \sphericalangle(w' - u, w') \\ &= \sphericalangle(u - w', u) \\ &< \sphericalangle(u - w', u - w) \\ &= \sphericalangle(u - w, (u - w) - (w' - w)) \end{aligned}$$

e da proposição anterior de oposição lado-ângulo segue que

$$\|v' - v\| = \|u - w\| < \|w' - w\|.$$

■

FIGURA 20.3: Soma e diferença de vetores de mesma norma

⊢ **Proposição 20.17.** *Seja* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado.*

1. *(LLL) Para todos* $v, v', w, w' \in V \smallsetminus \{0\}$ *tais que* $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ *e* $\|v + v'\| = \|w + w'\|$,
$$\sphericalangle(v, v') = \sphericalangle(w, w').$$
2. *(ALA) Para todos* $v, v', w, w' \in V \smallsetminus \{0\}$ *tais que* $\|v\| = \|w\|$, $\sphericalangle(v, v') = \sphericalangle(w, w')$ *e* $\sphericalangle(v, v + v') = \sphericalangle(w, w + w')$,
$$\|v + v'\| = \|w + w'\|$$
3. *(LAA) Para todos* $v, v', w, w' \in V \smallsetminus \{0\}$ *tais que* $\|v + v'\| = \|w + w'\|$, $\sphericalangle(v, v') = \sphericalangle(w, w')$ *e* $\sphericalangle(v, v + v') = \sphericalangle(w, w + w')$,
   3.1. $\|v\| = \|w\|$;
   3.2. $\sphericalangle(v, v + v') = \sphericalangle(w, w + w')$.

Perpendicularidade

:⊢ **Definição 20.12.** Seja $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ um espaço angulado. Vetores *perpendiculares* são vetores $v, v' \in V \smallsetminus \{0\}$ tais que

$$\sphericalangle(v, v') = \sphericalangle(v', -v).$$

Denota-se $v \perp v'$.

⊢ **Proposição 20.18** (Perpendicular)**.** *Sejam* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado e* $v, v' \in V \smallsetminus \{0\}$ *linearmente independentes. A função*

$$\begin{aligned} d\colon \mathbb{R} &\longrightarrow [0, \infty[\\ c &\longmapsto \|v' - cv\| \end{aligned}$$

é contínua, $\lim_{c \to \pm\infty} d(c) = \infty$*, tem único ponto de mínimo* $c_\perp \in \mathbb{R}$ *e*

$$v \perp (v' - c_\perp v).$$

□ *Demonstração.* A função d é contínua pois é composição da norma, subtração e multiplicação por escalar, que são contínuas. Dados $c, k \in \mathbb{R}$, segue da desigualdade triangular que

$$\begin{aligned} d(c) &= \|v' - cv\| \\ &\geq \|cv - kv\| - \|v' - kv\| \\ &= \left(\frac{|c-k| \, \|v\|}{\|v' - kv\|} - 1 \right) \|v' - kv\| . \end{aligned}$$

Como

$$\lim_{c \to \pm\infty} \frac{|c-k| \, \|v\|}{\|v' - kv\|} = \infty,$$

segue que

$$\lim_{c \to \pm\infty} d(c) = \infty.$$

Da continuidade e dos limites serem infinitos segue que d tem um ponto de mínimo $c_0 \in \mathbb{R}$. Ainda, para todo $s \in \,]d(c_0), \infty[$, existem $c_-, c_+ \in \mathbb{R}$ tais que $c_- < c_0 < c_+$ e

$$\|v' - c_+ v\| = d(c_+) = s = d(c_-) = \|v' - c_- v\| .$$

Definindo $c_\perp := \frac{c_- + c_+}{2}$, segue de homogeneidade e da proposição do triângulo isósceles (positivo) que

$$\begin{aligned} \sphericalangle (v' - c_+ v, v' - c_\perp v) &= \sphericalangle (v' - c_+ v, 2v' - (c_- + c_+) v) \\ &= \sphericalangle (v' - c_+ v, (v' - c_- v) + (v' - c_+ v)) \\ &= \sphericalangle (v' - c_- v, (v' - c_- v) + (v' - c_+ v)) \\ &= \sphericalangle (v' - c_- v, 2v' - (c_- + c_+) v) \\ &= \sphericalangle (v' - c_- v, v' - c_\perp v) . \end{aligned}$$

Como $\|v' - c_\perp v\| = \|v' - c_\perp v\|$ e $\|v' - c_+ v\| = \|v' - c_- v\|$, por LAL (negativo) segue que

$$\begin{aligned} \sphericalangle \left(v' - c_\perp v, \frac{c_- - c_+}{2} v \right) &= \sphericalangle (v' - c_\perp v, v' - c_\perp v - (v' - c_+ v)) \\ &= \sphericalangle (v' - c_\perp v, v' - c_\perp v - (v' - c_- v)) \\ &= \sphericalangle \left(v' - c_\perp v, \frac{c_+ - c_-}{2} v \right) . \end{aligned}$$

Por fim, segue da simetria e da homogeneidade que

$$\sphericalangle (v, v' - c_\perp v) = \sphericalangle (-v, v' - c_\perp v) ,$$

logo $v \perp (v' - c_\perp v)$.

Notemos que tal $c_\perp$ é único: suponhamos que existam $c, c' \in \mathbb{R}$, $c < c'$, tais que $v \perp (v' - cv)$ e $v \perp (v' - c'v)$. Então $c' - c > 0$ e

$$v' - cv = v' - c'v + (c' - c)v,$$

portanto da aditividade do ângulo segue que

$$\sphericalangle(v, v' - c'v) = \sphericalangle(v, v' - cv) + \sphericalangle(v' - cv, v' - c'v).$$

Como $v \perp (v' - cv)$ e $v \perp (v' - c'v)$, segue da unicidade do ângulo reto que

$$\sphericalangle(v' - cv, v' - c'v) = \sphericalangle(v, v' - c'v) - \sphericalangle(v, v' - cv) = 0;$$

da separação do ângulo, existe $k \in]0, \infty[$ $v' - c'v = k(v' - cv)$ e, como v, v' são linearmente independentes, segue que $k = 1$ e $c = c'$, o que contradiz $c < c'$.

Por fim, notemos que d é estritamente crescente em $]c_\perp, \infty[$ e estritamente decrescente em $]-\infty, c_\perp[$. Mostraremos que é crescente; a demonstração de que é decrescente é análoga. Suponhamos, por absurdo, que existam $k_-, k_+ \in]c_\perp, \infty[$ tais que $k_- < k_+$ e $d(k_-) = d(k_+)$. Pela mesma construção anterior, $k_\perp := \frac{k_- + k_+}{2}$ satisfaz $v \perp (v' - k_\perp v)$. Como $c_\perp$ é único com essa propriedade, segue que $k_\perp = c_\perp$. Mas $c_\perp \leq k_- < k_\perp$, contradição. Isso mostra que não podem existir tais k_-, k_+ distintos, logo da continuidade de d e de $\lim_{c\to\infty} d(c) = \infty$ segue que d estritamente crescente. Analogamente se mostra que d é estritamente decrescente em $]-\infty, c_\perp[$ e isso implica que $c_\perp$ é único ponto de mínimo de d, portanto $c_0 = c_\perp$ e $v \perp (v' - c_0 v)$. ■

⊢ **Corolário 20.19.** *Para todos $v, v' \in V \setminus \{0\}$ tais que $v \perp v'$,*

$$\|v + v'\| > \|v'\|$$

□ *Demonstração.* Como $(v + v') - v = v'$ e $v \perp ((v + v') - v)$, $c = 1$ é o ponto de mínimo da função $\|(v + v') - cv\|$ e então

$$\|v + v'\| = \|(v + v') - 0v\| > \|(v + v') - 1v\| = \|v'\|.$$

■

A última proposição nos dá para cada par $v, v' \in V \setminus \{0\}$ de vetores linearmente independentes, existe único $c_\perp \in \mathbb{R}$ tal que $v \perp (v' - c_\perp v)$. Isso significa que existe uma função $c_\perp(v, v')$. A seguir, mostraremos que essa função é de fato uma função das variáveis $\|v\|$, $\|v'\|$ e $\sphericalangle(v, v')$.

⊢ **Proposição 20.20.** *Seja $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ um espaço angulado. Para todos $v, v' \in V \setminus \{0\}$ e $w, w' \in V \setminus \{0\}$, sendo $\{v, v'\}$ e $\{w, w'\}$ linearmente independentes, tais que $\|v\| = \|w\|$, $\|v'\| = \|w'\|$ e $\sphericalangle(v, v') = \sphericalangle(w, w')$, vale*

$$c_\perp(v, v') = c_\perp(w, w').$$

□ *Demonstração.* Por simplicidade, denotemos $c := c_\perp(v, v')$ e $k := c_\perp(w, w')$. Como v, v' e w, w' são linearmente independentes, então $(v' - cv), (w' - kv) \neq 0$. Consideramos o caso em que $c = 0$. Nesse caso, $v' - cv = v'$ e portanto

$$\sphericalangle(w, w') = \sphericalangle(v, v') = \sphericalangle(v, v' - cv) = \sphericalangle(w', w' - kw) = \tau \diagup 4,$$

logo $w \perp w'$, e segue da unicidade de k que $k = 0$. Analogamente, supondo $k = 0$ obtemos $c = 0$.

Consideramos agora o caso em que $c \neq 0$ e $k \neq 0$. Mostraremos que c e k têm o mesmo sinal. Consideramos, por absurdo, os dois casos em que isso não ocorre.

1. ($k < 0 < c$) Nesse caso, temos $-k > 0$. Como $v' = cv + (v' - cv)$ e $w' - kw = (-k)w + w'$, segue da positividade e da aditividade que

$$\begin{aligned}
\sphericalangle(v, v') &< \sphericalangle(v, v') + \sphericalangle(v', v' - cv) \\
&= \sphericalangle(v, v' - cv) \\
&= \tau \diagup 4 \\
&= \sphericalangle(w, w' - kw) \\
&= \sphericalangle(w, w' - kw) + \sphericalangle(w' - kw, w') \\
&= \sphericalangle(w, w'),
\end{aligned}$$

   o que é uma contradição.
2. ($c < 0 < k$) Obtemos uma contradição de modo análogo ao caso anterior.

Sendo assim, temos que, se $c > 0$ e $k > 0$, então

$$\sphericalangle(cv, v') = \sphericalangle(v, v') = \sphericalangle(w, w') = \sphericalangle(kw, w'),$$

e, se $c < 0$ e $k < 0$, então

$$\sphericalangle(cv, v') = \tau \diagup 2 - \sphericalangle(v, v') = \tau \diagup 2 - \sphericalangle(w, w') = \sphericalangle(kw, w'),$$

portanto em ambos os casos vale

$$\sphericalangle(cv, v') = \sphericalangle(kw, w').$$

Assim, como $v' = cv + (v' - cv)$, $w' = kw + (w' - kw)$, e $\|v'\| = \|w'\|$, $\sphericalangle(cv, v') = \sphericalangle(kw, w')$ e $\sphericalangle(v, v' - cv) = \tau \diagup 4 = \sphericalangle(w, w' - kw)$, segue do critério LAA que $\|cv\| = \|kw\|$. Como $\|v\| = \|w\|$, segue que

$$|c| = \frac{\|cv\|}{\|v\|} = \frac{\|kw\|}{\|w\|} = |k|,$$

portanto $k = c$ ou $k = -c$. Mas sabemos que eles têm o mesmo sinal, portanto $c = k$. ■

$\vdash$ **Proposição 20.21.** *Seja* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado. Para todos* $v, v' \in V \smallsetminus \{0\}$ *linearmente independentes e* $c, c' \in]0, \infty[$

$$c_\perp(cv, c'v') = \frac{c'}{c} c_\perp(v, v').$$

$\square$ *Demonstração.* Como

$$v' - c_\perp(v, v')v = \frac{1}{c'}\left(c'v' - \frac{c'}{c} c_\perp(v, v')cv\right),$$

segue da 0-homogeneidade que

$$\begin{aligned}
\sphericalangle(c'v', v' - c_\perp(v, v')cv) &= \frac{\tau}{4} \\
&= \sphericalangle(v', v' - c_\perp(v, v')v) \\
&= \sphericalangle\left(c'v', \frac{1}{c'}\left(c'v' - \frac{c'}{c} c_\perp(v, v')cv\right)\right) \\
&= \sphericalangle\left(c'v', c'v' - \frac{c'}{c} c_\perp(v, v')cv\right).
\end{aligned}$$

Da unicidade de $c_\perp(cv, c'v')$ segue que

$$c_\perp(cv, c'v') = \frac{c'}{c} c_\perp(v, v').$$

$\blacksquare$

$\vdash$ **Proposição 20.22.** *Seja* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado. Para todos* $v, v' \in V \smallsetminus \{0\}$ *e* $w, w' \in V \smallsetminus \{0\}$ *tais que* $\sphericalangle(v, v') = \sphericalangle(w, w')$,

$$\frac{\|v\|}{\|v'\|} c_\perp(v, v') = \frac{\|w\|}{\|w'\|} c_\perp(w, w').$$

$\square$ *Demonstração.* Como

$$\left\|\frac{v}{\|v\|}\right\| = 1 = \left\|\frac{w}{\|w\|}\right\|, \qquad \left\|\frac{v'}{\|v'\|}\right\| = 1 = \left\|\frac{w'}{\|w'\|}\right\|$$

e, pela 0-homogeneidade,

$$\sphericalangle\left(\frac{v}{\|v\|}, \left\|\frac{v'}{\|v'\|}\right\|\right) = \sphericalangle(v, v') = \sphericalangle(w, w') = \sphericalangle\left(\frac{w}{\|w\|}, \left\|\frac{w'}{\|w'\|}\right\|\right),$$

segue da proposição 20.20 que

$$c_\perp\left(\frac{v}{\|v\|},\frac{v'}{\|v'\|}\right)=c_\perp\left(\frac{w}{\|w\|},\frac{w'}{\|w'\|}\right).$$

Assim, segue da proposição 20.21 que

$$\begin{aligned}\frac{\|v\|}{\|v'\|}c_\perp(v,v')&=c_\perp\left(\frac{v}{\|v\|},\frac{v'}{\|v'\|}\right)\\&=c_\perp\left(\frac{w}{\|w\|},\frac{w'}{\|w'\|}\right)\\&=\frac{\|w\|}{\|w'\|}c_\perp(w,w').\end{aligned}$$

■

Essa proposição mostra que $\frac{\|v\|}{\|v'\|}c_\perp(v,v')$ só depende de $\sphericalangle(v,v')$, ou seja, que existe uma função

$$\begin{aligned}\angle\colon]0,\tau/2[&\longrightarrow\mathbb{R}\\a&\longmapsto\angle(a)\end{aligned}$$

tal que

$$\angle(\sphericalangle(v,v'))=\frac{\|v\|}{\|v'\|}c_\perp(v,v').$$

Mostraremos que essa função é $\angle=\cos$.

⊢ **Proposição 20.23.** *Seja* $(\boldsymbol{V}, \|\cdot\|, \sphericalangle)$ *um espaço angulado.*

1. *(Retificação) Para todos* $v, v' \in V \setminus \{0\}$, $\sphericalangle(v, v') = \sphericalangle(-v, v')$ *se, e somente se,*
$$\sphericalangle(v, v') = \frac{\tau}{4}.$$
2. *(Alternos internos) Para todos* $v, v' \in V \setminus \{0\}$,
$$\sphericalangle(v, v') = \sphericalangle(-v, -v').$$
3. *(Monotonicidade complementar) Para todos* $v, v' \in V \setminus \{0\}$ *linearmente independentes e* $c, c' \,]0, \infty[$ *tais que* $c < c'$,
$$\sphericalangle(v, v + cv') < \sphericalangle(v, v + c'v').$$
4. *(Monotonicidade suplementar) Para todos* $v, v' \in V \setminus \{0\}$ *linearmente independentes e* $c, c' \,]0, \infty[$ *tais que* $c < c'$,
$$\sphericalangle(v, v' + c(-v)) < \sphericalangle(v, v' + c'(-v)).$$
5. *Para todos* $v, v' \in V \setminus \{0\}$ *e* $c \,]0, \infty[$,
$$\lim_{c \to \infty} \sphericalangle(v, v + cv') = \sphericalangle(v, v');$$
6. *Para todos* $v, v' \in V \setminus \{0\}$ *e* $c \,]0, \infty[$,
$$\lim_{c \to \infty} \sphericalangle(v, v' + c(-v)) = \sphericalangle(v, -v);$$

□ *Demonstração.* 1. Se $\sphericalangle(v, v') = \sphericalangle(-v, v')$, segue da suplementação que
$$\sphericalangle(v, v') = \frac{2\sphericalangle(v, v')}{2} = \frac{\sphericalangle(v, v') + \sphericalangle(-v, v')}{2} = \frac{\tau}{4}.$$
Reciprocamente, se $\sphericalangle(v, v') = \tau \diagup 4$, segue da suplementação que
$$\sphericalangle(v, v') = \frac{\tau}{4} = \frac{\tau}{2} - \sphericalangle(v, v') = \sphericalangle(-v, v').$$

2. Segue da suplementação e da simetria que
$$\sphericalangle(v, v') = \frac{\tau}{2} - \sphericalangle(-v, v') = \frac{\tau}{2} - \left(\frac{\tau}{2} - \sphericalangle(-v, -v')\right) = \sphericalangle(-v, -v').$$

3. Como $c < c'$, então para todo $k \in]0, \infty[$ temos $v + cv' \neq k(v + c'v')$, portanto da separação e da positividade do ângulo segue que $0 < \sphericalangle(v + cv', v + c'v')$. Definindo $k := 1 - c \diagup c'$ e $k' := c \diagup c'$, temos que $k' > 0$, pois $c, c' > 0$, e $k > 0$, pois $c < c'$, portanto
$$v + cv' = (k + k')v + (k'c')v' = kv + k'(v + c'v').$$
Assim da aditividade do ângulo segue que
$$\begin{aligned}\sphericalangle(v, v + cv') &< \sphericalangle(v, v + cv') + \sphericalangle(v + cv', v + c'v') \\ &= \sphericalangle(v, v + c'v')\end{aligned}$$
4. Análogo ao item anterior.
5. Como $\sphericalangle$ é homogênea e contínua,
$$\sphericalangle(v, v + cv') = \sphericalangle\left(v, \frac{v + cv'}{\|v + cv'\|}\right) \to \sphericalangle(v, v').$$
6. Análogo ao item anterior.

■

Aditividade do produto interno Sejam $v, v', v'' \in V$. Queremos mostrar que
$$\langle v + v', v'' \rangle = \langle v, v'' \rangle + \langle v', v'' \rangle.$$
Como supomos que v'' é não nulo, isso é equivalente a
$$\|v + v'\| \cos(\sphericalangle(v + v', v'')) = \|v\| \cos(\sphericalangle(v, v'')) + \|v'\| \cos(\sphericalangle(v', v'')).$$
Separamos em casos.

1. (v e v' são linearmente dependentes) Nesse caso, para algum $c \in \mathbb{R} \setminus \{0\}$, $v' = cv$. O caso em que $c = 0$ não ocorre, pois estamos supondo $v' \neq 0$. Da homogeneidade do produto interno, provada no subitem anterior,
$$\begin{aligned}\langle v + v', v'' \rangle &= \langle v + cv, v'' \rangle \\ &= \langle (1 + c)v, v'' \rangle \\ &= (1 + c) \langle v, v'' \rangle \\ &= \langle v, v'' \rangle + c \langle v, v'' \rangle \\ &= \langle v, v'' \rangle + \langle cv, v'' \rangle \\ &= \langle v, v'' \rangle + \langle v', v'' \rangle.\end{aligned}$$
2. (v e v' são linearmente independentes e v'' é gerado por eles) Nesse caso, existem $c, c' \in \mathbb{R}$ tais que $v'' = cv + c'v' \neq 0$.
TERMINAR
3. (v, v' e v'' são linearmente independentes)
TERMINAR

Uma possível generalização multidimensional

:⊢ **Definição 20.13.** Seja $(\boldsymbol{V}, \|\cdot\|)$ um espaço normado. Uma *função ângulo 3-dimensional* em $(\boldsymbol{V}, \|\cdot\|)$ é uma função contínua

$$\sphericalangle\colon V \smallsetminus \{0\} \times V \smallsetminus \{0\} \times V \smallsetminus \{0\} \longrightarrow [0, \tau]$$

tal que

1. (Homogeneidade) Para todos $v, v', v'' \in V$ e $c, c', c'' \in \,]0, \infty[$,
$$\sphericalangle(cv, c'v', c''v'') = \sphericalangle(v, v', v'');$$
2. (Separação) Para todos $v, v', v'' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v', v'') = 0$$
se, e somente se, $\{v, v', v''\}$ é conicamente dependente (existem $c, c', c'' \in \,]0, \infty[$ tais que $cv + c'v' + c''v'' = 0$);
3. (Simetria) Para todos $v, v', v'' \in V \smallsetminus \{0\}$ e toda bijeção $f\colon \{v, v', v''\} \longrightarrow \{v, v', v''\}$,
$$\sphericalangle(f(v), f(v'), f(v'')) = \sphericalangle(v, v', v'');$$
4. (Suplementação) Para todos $v, v', v'' \in V \smallsetminus \{0\}$,
$$\sphericalangle(v, v', v'') + \sphericalangle(-v, v', v'') + \sphericalangle(-v, -v', v'') + \sphericalangle(v, -v', v'') = \tau;$$
5. (Aditividade) Para todos $v, v', v'' \in V \smallsetminus \{0\}$ e $c, c', c'' \in \,]0, \infty[$ tais que $cv + c'v' + c''v'' \neq 0$,
$$\begin{aligned}\sphericalangle(v, v', v'') = &\sphericalangle(v, v', cv + c'v' + c''v'')\\ &+ \sphericalangle(v, cv + c'v' + c''v'', v'')\\ &+ \sphericalangle(cv + c'v' + c''v'', v', v'');\end{aligned}$$
6. (Ângulo interno) Para todos $v, v', v'', w, w', w'' \in V \smallsetminus \{0\}$ tais que $v + v' \neq 0$, $w + w' \neq 0$, $\|v\| = \|w\|$, $\|v'\| = \|w'\|$, $\|v''\| = \|w''\|$, $\|v + v'\| = \|w + w'\|$, $\|v' + v''\| = \|w' + w''\|$ e $\sphericalangle(v, v', v'') = \sphericalangle(w, w', w'')$, então
   6.1. $\|v + v''\| = \|w + w''\|$;
   6.2. $\sphericalangle(v, v', v'') = \sphericalangle(w, w', w'')$;

20.2.6 Espaço projetivo

O espaço projetivo é o espaço de retas pela origem. Ele pode ser descrito de diferentes formas. Quando $V = \mathbb{R}^d$ (ou um espaço normado), podemos descrevê-lo como o quociente da esfera unitária pela antípoda $-\mathrm{I}$. Quando $\boldsymbol{V}$ é um espaço com produto interno, podemos induzir em $\mathbb{P}V$ uma distância, o ângulo entre as retas de V. De qualquer forma, esse espaço é um espaço topológico, como será também descrito a seguir.

20.2.6.1 Estrutura topológica

A relação de paralelismo $\parallel$ é uma equivalência não só em V, mas em $V \smallsetminus \{0\}$. Além disso, ela não depende de produto interno, está definida para qualquer espaço linear $\boldsymbol{V}$ sobre um corpo $\boldsymbol{C}$ qualquer. Isso permite que se quociente $V \smallsetminus \{0\}$ por $\parallel$, e esse é o *espaço projetivo* de $\boldsymbol{V}$.

:⊢ **Definição 20.14.** Seja $\boldsymbol{V}$ um espaço linear sobre um corpo $\boldsymbol{C}$. O *espaço projetivo* de $\boldsymbol{V}$ é o conjunto

$$\mathbb{P}V := {(V \smallsetminus \{0\})}\big/{\parallel}.$$

Os elementos de $\mathbb{P}V$ são as *retas* de $\boldsymbol{V}$.

Esse quociente pode ser entendido também como o quociente pela ação do grupo multiplicativo $C \smallsetminus \{0\}$. Desse modo, temos que

$$\mathbb{P}V = {(V \smallsetminus \{0\})}\big/{(C \smallsetminus \{0\})}.$$

O caso em que $\boldsymbol{V}$ é um espaço normado nos permite dar mais uma definição equivalente desse espaço. Lembremos que $\mathbb{S}^0 = \{1, -1\}$ é o grupo discreto multiplicativo com 2 elementos e é subgrupo do grupo multiplicativo $C \smallsetminus \{0\}$. Se $\boldsymbol{V}$ é um espaço normado, a esfera $\mathbb{S}V$ de $\boldsymbol{V}$ está definida. Nesse caso, $\mathbb{S}^0$ age em $\mathbb{S}V$ e o espaço quociente $\mathbb{S}/\mathbb{S}^0$ pode ser identificado com $\mathbb{P}V$ de um jeito bem natural, de modo que as estruturas definidas em um possam sempre ser passadas para o outro. Temos

$$\mathbb{P}V \simeq {\mathbb{S}V}\big/{\mathbb{S}^0}.$$

O isomorfismo é dado por

$$\begin{aligned} h\colon \mathbb{P}V &\longrightarrow {\mathbb{S}V}\big/{\mathbb{S}^0} \\ [v] &\longmapsto \left\{ \frac{v}{\|v\|}, -\frac{v}{\|v\|} \right\} \end{aligned}$$

com inversa

$$\begin{aligned} h^{-1}\colon {\mathbb{S}V}\big/{\mathbb{S}^0} &\longrightarrow \mathbb{P}V \\ \{u, -u\} &\longmapsto \{cu \mid c \in C \smallsetminus \{0\}\}. \end{aligned}$$

20.2.6.2 Estrutura métrica

Quando o espaço linear tem um produto interno, a função ângulo pode ser usada para definir uma função distância no espaço projetivo.

:⊢ **Definição 20.15.** Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno. O *ângulo* entre retas r e r' de $\boldsymbol{V}$ é o menor ângulo entre vetores de r e r'

$$\sphericalangle(r, r') := \min_{v\in r, v'\in r'} \sphericalangle(v, v').$$

Mostremos como achar uma expressão para $\sphericalangle(r, r')$, o que facilitará a demonstração de que essa função é de fato uma distância. Queremos mostrar que

$$\sphericalangle(r, r') = \cos^{-1}\left(\frac{|\langle v, v'\rangle|}{\|v\|\,\|v'\|}\right),$$

em que $v \in r, v' \in r'$. Para todos $\bar{v} \in r$ e $\bar{v}' \in r'$, existem escalares $c, c' \in C \smallsetminus \{0\}$ tais que $\bar{v} = cv$ e $\bar{v}' = c'v'$. Portanto

$$\begin{aligned}\sphericalangle(r, r') &= \min_{\bar{v}\in r, \bar{v}'\in r'} \sphericalangle(\bar{v}, \bar{v}') \\ &= \min_{c,c'\in C\smallsetminus\{0\}} \sphericalangle(cv, cv') \\ &= \min\left\{\sphericalangle(v, v'), \frac{\tau}{2} - \sphericalangle(v, v')\right\}.\end{aligned}$$

Como $\sphericalangle(v, v') \in \left[0, \frac{\tau}{2}\right]$, segue que

$$\min\left\{\sphericalangle(v, v'), \frac{\tau}{2} - \sphericalangle(v, v')\right\} \in \left[0, \frac{\tau}{4}\right].$$

Assim, escolhendo $v \in r$ e $v' \in r'$ tais que $\langle v, v'\rangle \geq 0$, temos $\sphericalangle(v, v') \in \left[0, \frac{\tau}{4}\right]$, o que implica que

$$\sphericalangle(r, r') = \min\left\{\sphericalangle(v, v'), \frac{\tau}{2} - \sphericalangle(v, v')\right\} = \cos^{-1}\left(\frac{|\langle v, v'\rangle|}{\|v\|\,\|v'\|}\right).$$

Podemos restringir ainda mais a escolha dos $v \in r$ e $v' \in r'$ notando que

$$\cos^{-1}\left(\frac{|\langle v, v'\rangle|}{\|v\|\,\|v'\|}\right) = \cos^{-1}\left(\left|\left\langle \frac{v}{\|v\|}, \frac{v'}{\|v'\|}\right\rangle\right|\right)$$

e que $\frac{v}{\|v\|}, \frac{v'}{\|v'\|} \in \mathbb{S}$. Como existe $u \in V$ tal que $r \cap \mathbb{S} = \{u, -u\}$, podemos tomar $u \in r \cap \mathbb{S}$ e $u' \in r' \cap \mathbb{S}$ e temos que

$$\sphericalangle(r, r') = \cos^{-1}(|\langle u, u'\rangle|).$$

:⊢ **Definição 20.16.** Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno. A distância em $\mathbb{P}V$ é a função

$$\begin{aligned} \sphericalangle(\cdot,\cdot)\colon \mathbb{P}V \times \mathbb{P}V &\longrightarrow \left[0, \frac{\tau}{4}\right] \\ (r, r') &\longmapsto \sphericalangle(r, r')\,. \end{aligned}$$

⊢ **Proposição 20.24.** *Seja $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ um espaço com produto interno. A função $\sphericalangle(\cdot,\cdot)$ é uma distância em $\mathbb{P}V$.*

□ *Demonstração.* (Separação) Sejam $r \in \mathbb{P}V$ e $v \in r$. Então

$$\sphericalangle(r, r) = \cos^{-1}\left(\frac{|\langle v, v\rangle|}{\|v\|\,\|v\|}\right) = \cos^{-1}(1) = 0.$$

Reciprocamente, sejam $r, r' \in \mathbb{P}V$, $v \in r$ e $v' \in r'$. Se $\sphericalangle(r, r') = 0$, então $\frac{|\langle v,v\rangle|}{\|v\|\|v\|} = 1$, o que implica que $v' = v$ ou $v' = -v$, logo $r = r'$. (Simetria) Sejam $r, r' \in \mathbb{P}V$, $v \in r$ e $v' \in r'$. Então

$$\sphericalangle(r, r') = \cos^{-1}\left(\frac{|\langle v, v'\rangle|}{\|v\|\,\|v'\|}\right) = \cos^{-1}\left(\frac{|\langle v', v\rangle^*|}{\|v'\|\,\|v\|}\right) = \cos^{-1}\left(\frac{|\langle v', v\rangle|}{\|v'\|\,\|v\|}\right) = \sphericalangle(r', r)\,.$$

(Desigualdade Triangular) Sejam $r, r', r'' \in \mathbb{P}V$, $v \in r$, $v' \in r'$ e $v'' \in r''$. Queremos mostrar que

$$\sphericalangle(r, r'') \leq \sphericalangle(r, r') + \sphericalangle(r', r'')\,.$$

Consideremos a projeção $\bar{v}$ de v' no plano $\langle\{v, v''\}\rangle$.

Se mostrarmos que

$$|\langle u, u''\rangle| \geq |\langle u, u'\rangle| + |\langle u', u''\rangle|\,,$$

segue da monotonicidade decrescente de $\cos^{-1}$.

$$|\langle u, u'\rangle + \langle u', u''\rangle| \leq |\langle u, u'\rangle| + |\langle u', u''\rangle|$$

$$\langle u, u'\rangle + \langle u'', u'\rangle = \langle u + u'', u'\rangle$$

■

20.2.7 Funções ortogonais e conformes

:⊢ **Definição 20.17.** Sejam $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ e $(\boldsymbol{V}', \langle\cdot,\cdot\rangle')$ espaços com produto interno. Uma função *ortogonal* de $\boldsymbol{V}$ para $\boldsymbol{V}'$ é uma função linear $f\colon V \to V'$ tal que, para todos $v, v' \in V$,

$$\langle f(v), f(v')\rangle' = \langle v, v'\rangle\,.$$

O conjunto dessas funções é $\mathscr{L}_{\langle,\rangle}(\boldsymbol{V}, \boldsymbol{V}')$.

Uma função *conforme* de $\boldsymbol{V}$ para $\boldsymbol{V}'$ é uma função linear $f\colon V \to V'$ tal que, para todos $v, v' \in V$,

$$\sphericalangle'(f(v), f(v')) = \sphericalangle\,(v, v')\,.$$

O conjunto dessas funções é $\mathscr{L}_{\sphericalangle}(\boldsymbol{V}, \boldsymbol{V}')$.

Note que, na definição de uma função conforme, está implícito que

$$\sphericalangle'(f(v), f(v'))$$

existe, portanto $f(v) \neq 0$ para todo $v \in V \smallsetminus \{0\}$, o que significa que f é injetiva.

⊢ **Proposição 20.25.** *Sejam $(\boldsymbol{V}, \langle\cdot,\cdot\rangle)$ e $(\boldsymbol{V}', \langle\cdot,\cdot\rangle')$ espaços com produto interno sobre um corpo $\boldsymbol{C}$ de característica diferente de 2 e $f\colon V \to V'$ uma função linear.*

1. *f é ortogonal se, e somente se, é uma isometria local:*

$$\mathscr{L}_{\langle,\rangle}(\boldsymbol{V}, \boldsymbol{V}') = \mathscr{L}_{\|\|}(\boldsymbol{V}, \boldsymbol{V}');$$

2. *f é conforme se, e somente se, existe $c \in \,]0, \infty[$ tal que, para todos $v, v' \in V$,*

$$\langle f(v), f(v')\rangle' = c\,\langle v, v'\rangle\,.$$

□ *Demonstração.* 1. Se f é ortogonal, então para todo $v \in V$,

$$\|f(v)\|' = \langle f(v), f(v)\rangle'^{\frac{1}{2}} = \langle v, v\rangle^{\frac{1}{2}} = \|v\|\,.$$

Reciprocamente, como a característica de $\boldsymbol{C}$ é diferente de 2, vale que, para todos $v, v' \in V$,

$$\langle v, v'\rangle = \frac{1}{4}\left(\left(\|v+v'\|^2 - \|v-v'\|^2\right) + \mathrm{i}\left(\|v+\mathrm{i}v'\|^2 - \|v-\mathrm{i}v'\|^2\right)\right).$$

Se f é isometria local, então, para todos $v, v' \in V$,

$$\begin{aligned}
\langle f(v), f(v')\rangle &= \frac{1}{4}\left(\|f(v)+f(v'))\|^2 - \|f(v)-f(v'))\|^2\right) \\
&\qquad + \frac{\mathrm{i}}{4}\left(\|f(v)+\mathrm{i}f(v')\|^2 - \|f(v)-\mathrm{i}f(v')\|^2\right) \\
&= \frac{1}{4}\left(\|f(v+v')\|^2 - \|f(v-v')\|^2\right) \\
&\qquad + \frac{\mathrm{i}}{4}\left(\|f(v+\mathrm{i}v')\|^2 - \|f(v-\mathrm{i}v')\|^2\right) \\
&= \frac{1}{4}\left(\left(\|v+v'\|^2 - \|v-v'\|^2\right) + \mathrm{i}\left(\|v+\mathrm{i}v'\|^2 - \|v-\mathrm{i}v'\|^2\right)\right) \\
&= \langle v, v'\rangle .
\end{aligned}$$

2. Se f é conforme, para todos $v, v' \in V$ vale

$$\sphericalangle'(f(v), f(v')) = \sphericalangle(v, v') .$$

Como $\cos^{-1}$ é bijeção, então

$$\frac{\langle f(v), f(v')\rangle'}{\|f(v)\|'\,\|f(v')\|'} = \frac{\langle v, v'\rangle}{\|v\|\,\|v'\|},$$

o que implica

$$\langle f(v), f(v')\rangle' = \frac{\|f(v)\|'\,\|f(v')\|'}{\|v\|\,\|v'\|}\langle v, v'\rangle .$$

Resta mostrar que $\frac{\|f(v)\|'\|f(v')\|'}{\|v\|\|v'\|}$ é constante. Note que isso ocorre se, e somente se, $N(v) := \frac{\|f(v)\|'}{\|v\|}$ é constante. Para mostrar que essa função é constante, vamos mostrar que sua diferencial é nula. Sejam $v \in V \setminus \{0\}$ e $h \in V$. Como f é linear e $\mathrm{D}\,\|v\|\,(h) = \frac{\langle v,h\rangle}{\|v\|}$ é a diferencial[6] de $\|\cdot\|$, a diferencial de N é

$$\begin{aligned}
\mathrm{D}N|_v(h) &= \frac{\|v\|\,\mathrm{D}\,\|f(v)\|'\,(h) - \|f(v)\|'\,\mathrm{D}\,\|v\|\,(h)}{\|v\|^2} \\
&= \frac{\frac{\|v\|}{\|f(v)\|'}\langle f(v), \mathrm{D}f(v)(h)\rangle' - \frac{\|f(v)\|'}{\|v\|}\langle v, h\rangle}{\|v\|^2} \\
&= \frac{\|v\|^2\langle f(v), f(h)\rangle' - \|f(v)\|'^2\langle v, h\rangle}{\|f(v)\|'\,\|v\|^3}.
\end{aligned}$$

[6] Não tenho certeza, mas acredito que isso dependa de a característica de $\boldsymbol{C}$ ser diferente de 2, pois quando calculamos a diferencial pela regra da cadeia cancelamos fatores de 2.

Notemos que, como f preserva ângulo, então se $v \perp v'$ segue que $f(v) \perp f(v')$. Escrevendo $h = \mathrm{p}_{\|v}(h) + \mathrm{p}_{\perp v}(h) = h^{\|} + h^{\perp}$, temos que $h^{\perp} \perp v$, e $h^{\|} \parallel v$ ou $h^{\|} = 0$, logo

$$\mathrm{D}N|_v(h) = \mathrm{D}N|_v(h^{\|} + h^{\perp}) = \mathrm{D}N|_v(h^{\|}) + \mathrm{D}N|_v(h^{\perp}).$$

Calculemos $\mathrm{D}N|_v(h^{\|})$. Como $h^{\|} = cv$, com $c = \frac{\langle h,v\rangle}{\|v\|^2}$, segue que

$$\begin{aligned}
\mathrm{D}N|_v(h^{\|}) &= \frac{\|v\|^2 \langle f(v), f(cv)\rangle' - \|f(v)\|'^2 \langle v, cv\rangle}{\|f(v)\|' \|v\|^3} \\
&= \frac{c\,\|v\|^2 \langle f(v), f(v)\rangle' - c\|f(v)\|'^2 \langle v, v\rangle}{\|f(v)\|' \|v\|^3} \\
&= \frac{c\,\|v\|^2 \|f(v)\|'^2 - c\|f(v)\|'^2 \|v\|^2}{\|f(v)\|' \|v\|^3} \\
&= 0.
\end{aligned}$$

Calculemos agora $\mathrm{D}N|_v(h^{\perp})$. Como $\left\langle h^{\perp}, v\right\rangle = 0$, então $\left\langle f(v), f(h^{\perp})\right\rangle = 0$, logo

$$\mathrm{D}N|_v(h^{\perp}) = \frac{\|v\|^2 \left\langle f(v), f(h^{\perp})\right\rangle' - \|f(v)\|'^2 \left\langle v, h^{\perp}\right\rangle}{\|f(v)\|' \|v\|^3} = 0.$$

Assim, concluímos que $\mathrm{D}N|_v(h) = \mathrm{D}N|_v(h^{\|}) + \mathrm{D}N|_v(h^{\perp}) = 0$, ou seja, $\mathrm{D}N|_v = 0$ para todo $v \in V \smallsetminus \{0\}$, o que implica que existe $c' \in \mathbb{R}$ tal que, para todo $v \in V \smallsetminus \{0\}$, $N(v) = \frac{\|f(v)\|}{\|v\|} = c'$. Como f preserva ângulos, é injetiva, portanto $f(v) = 0$ se, e somente se, $v = 0$, o que significa que $c' \neq 0$. Definindo $c := (c')^2$, segue que $c \in \,]0, \infty[$, portanto

$$\langle f(v), f(v')\rangle' = \frac{\|f(v)\|' \|f(v')\|'}{\|v\| \|v'\|} \langle v, v'\rangle = (c')^2 \langle v, v'\rangle = c \langle v, v'\rangle .$$

Reciprocamente, se existe $c \in \,]0, \infty[$ tal que

$$\langle f(v), f(v')\rangle' = c \langle v, v'\rangle ,$$

então $\|f(v)\| = \langle f(v), f(v)\rangle^{\frac{1}{2}} = c^{\frac{1}{2}} \langle v, v'\rangle^{\frac{1}{2}} = c^{\frac{1}{2}} \|v\|$ e segue que

$$\begin{aligned}
\sphericalangle' (f(v), f(v')) &= \cos^{-1}\left(\frac{\langle f(v), f(v')\rangle'}{\|f(v)\|' \|f(v')\|'}\right) \\
&= \cos^{-1}\left(\frac{c \langle v, v'\rangle}{c \|v\| \|v'\|}\right) \\
&= \sphericalangle (v, v') .
\end{aligned}$$

■

20.3 Espaço de funções quadrado somáveis

Sejam X um conjunto e $\boldsymbol{C} \subseteq \mathbb{R}^2$. Para todo $p \in [1, \infty]$, o espaço $\mathscr{S}^p(\boldsymbol{X}, \boldsymbol{C})$ é um espaço normado completo, mas somente para $p = 2$ esse espaço admite um produto interno. As funções absolutamente 2-somáveis do espaço $Smvl^2(\boldsymbol{X}, \boldsymbol{C})$ são também chamadas de *funções quadrado somáveis*. Note que, diferente do caso mais geral de funções absolutamente p-somáveis, aqui fixamos um subcorpo dos números complexos para usar o conjugação complexa. A generalização do complexo conjugado é mais difícil e detalhada e não será feita aqui.

:⊢ **Definição 20.18.** Sejam X um conjunto e $\boldsymbol{C} \subseteq \mathbb{R}^2$ um corpo. O *produto interno* entre $f, f' \in \mathscr{S}^2(\boldsymbol{X}, \boldsymbol{C})$ é

$$\langle f, f' \rangle := \left(\mathop{+}_{x \in X} f(x) \overline{f'(x)} \right)^{2^{-1}} .$$

⊢ **Proposição 20.26.** *Sejam X um conjunto e $\boldsymbol{C} \subseteq \mathbb{R}^2$ um corpo. O espaço $(\mathscr{S}^2(\boldsymbol{X}, \boldsymbol{C}), \langle \cdot, \cdot \rangle)$ é um espaço com produto interno completo.*

Capítulo 21

Variedades

21.1 Variedades topológicas e diferenciais

21.1.1 Cartas e atlas

:⊢ **Definição 21.1.** Sejam V um conjunto e $d \in \mathbb{N}$. Uma *carta* d-dimensional de V é um par (A, x) em que $A \subseteq V$ é um conjunto e $x\colon A \to \mathbb{R}^d$ é uma função injetiva tal que $x(A)$ é um aberto de $\mathbb{R}^d$. O *domínio* de (A, x) é A, o domínio de x, e o *mapa de coordenadas* de (A, x) é a função x. A *i-ésima coordenada* da carta (A, x) é a função $x^i := \pi^i \circ x\colon A \to \mathbb{R}$.

:⊢ **Definição 21.2.** Sejam V um conjunto e $(A, x), (\bar{A}, \bar{x})$ cartas d-dimensionais de V. A *transição de coordenadas* de (A, x) para $(\bar{A}, \bar{x})$ é a função

$$\bar{x} \circ x^{-1} : x(A \cap \bar{A}) \to \bar{x}(A \cap \bar{A}).$$

A composição $\bar{x} \circ x^{-1}$ não está necessariamente definida em todo $x(A)$, mas somente na interseção do domínio de x^{-1} com a imagem inversa do domínio de $\bar{x}$ por x^{-1}. Para ver que esse é o conjunto acima, basta notar que da injetividade de x segue

$$x(A) \cap (x^{-1})^{-1}(\bar{A}) = x(A \cap \bar{A}).$$

Da mesma forma definimos o contradomínio como acima.

:⊢ **Definição 21.3.** Seja V um conjunto. Cartas *$\mathscr{C}^k$-compatíveis* de V são cartas (A, x) e $(\bar{A}, \bar{x})$ tais que $x(A \cap \bar{A})$ e $\bar{x}(A \cap \bar{A})$ são abertos e a transição de coordenadas

$$\bar{x} \circ x^{-1} : x(A \cap \bar{A}) \to \bar{x}(A \cap \bar{A})$$

é um difeomorfismo $\mathscr{C}^k$. Para $k = 0$, as cartas são *topologicamente compatíveis*, para $k > 0$, são *diferencialmente compatíveis*, sendo para $k = \infty$ *suavemente compatíveis*.

A relação de compatibilidade entre cartas não é uma relação de equivalência, mas a compatibilidade de atlas, uma noção associada a ela, é. A seguir, definimos o que é um atlas. A primeira condição, além de aparentemente necessária para que todos os pontos de V tenham cartas, será uma hipótese essencial para que a compatibilidade de atlas seja uma relação de equivalência. Isso é um resultado muito simples, mas que indica que nossa axiomatização é boa, de certa forma, pois ao termos a estrutura de um atlas do conjunto, não temos simplesmente várias cartas, mas cartas que se comportam de acordo com certa relação de equivalência. Essa relação de equivalência nos permitirá definir um objeto abstrato chamado atlas maximal, que na prática nunca é calculado, mas que como objeto teórico é belo e conveniente.

:⊢ **Definição 21.4.** Seja V um conjunto. Um *atlas* $\mathscr{C}^k$ d-dimensional de V é um conjunto $\mathcal{A}$ de cartas d-dimensionais de V tal que

1. (Cobertura) O conjunto V é coberto pelos domínios das cartas de $\mathcal{A}$

$$V = \bigcup_{(A,x)\in\mathcal{A}} A;$$

2. (Compatibilidade $\mathscr{C}^k$) Todas cartas $(A, x), (\bar{A}, x) \in \mathcal{A}$ são $\mathscr{C}^k$-compatíveis.

Para $k = 0$, $\mathcal{A}$ é um atlas *topológico*, para $k > 0$, um atlas *diferencial*, sendo para $k = \infty$ um atlas *suave*.

Na definição de atlas, se não especificássemos que todas cartas devem ser d-dimensionais, poderíamos ter partes do conjunto V que fossem mapeadas em espaços reais de dimensão diferente. No entanto, nas cartas $(A, x), (\bar{A}, \bar{x}) \in \mathcal{A}$ em que $A \cap \bar{A} \neq \emptyset$, teríamos ao menos um homeomorfismo entre os conjuntos $x(A \cap \bar{A})$ e $\bar{x}(A \cap \bar{A})$, de modo que as cartas poderiam ser restritas a espaços reais de mesma dimensão; se $k \geq 1$, diferenciando a transição de coordenadas $\bar{x} \circ x^{-1}$ teríamos um homeomorfismo linear entre os espaços reais, garantindo a mesma dimensão. Isso mostra que, a menos de cartas que não tenham interseção no domínio, teríamos que ter espaços reais de mesma dimensão modelando o conjunto V. A partir de agora, sempre consideraremos que os atlas têm a mesma dimensão, e a dimensão de um atlas só será mencionada quando necessário.

:⊢ **Definição 21.5.** Seja V um conjunto. Atlas $\mathscr{C}^k$*-compatíveis* de V são atlas $\mathscr{C}^k$ $\mathcal{A}$ e $\bar{\mathcal{A}}$ de V tais que todas as cartas $(A, x) \in \mathcal{A}$ e $(\bar{A}, \bar{x}) \in \bar{\mathcal{A}}$ são $\mathscr{C}^k$-compatíveis.

Essa definição é equivalente a dizer que $\mathcal{A} \cup \bar{\mathcal{A}}$ é um atlas $\mathscr{C}^k$.

⊢ **Proposição 21.1.** *Seja V um conjunto. A relação de $\mathscr{C}^k$-compatibilidade de atlas de V é uma relação de equivalência.*

□ *Demonstração.* A reflexividade e a simetria são evidentes, pois valem entre cartas. Para conferir a transitividade, sejam $\mathcal{A}, \bar{\mathcal{A}}$ e $\{(A_i, x_i)\}_{i\in I}$ atlas de V e sejam $(A, x) \in \mathcal{A}$ e $(\bar{A}, \bar{x}) \in \bar{\mathcal{A}}$ cartas. Como $\mathcal{A}$ e $\bar{\mathcal{A}}$ são compatíveis com $\{(A_i, x_i)\}_{i\in I}$, segue que para todo $i \in I$, $x(A \cap A_i)$ e $x_i(A_i \cap \bar{A})$ são abertos de $\mathbb{R}^d$ e as transições de coordenadas

$$x_i \circ x^{-1} : x(A \cap A_i) \to x_i(A \cap A_i).$$

e

$$\bar{x} \circ x_i^{-1} : x_i(A_i \cap \bar{A}) \to \bar{x}(A_i \cap \bar{A}).$$

são difeomorfismos $\mathscr{C}^k$. Compondo essas transições de coordenada, e notando que

$$x(A \cap A_i) \cap (x_i \circ x^{-1})^{-1}(x_i(A_i \cap \bar{A})) = x(A \cap A_i \cap \bar{A})$$

(e analogamente para $\bar{x}(A \cap A_i \cap \bar{A})$), concluímos que

$$(\bar{x} \circ x_i^{-1}) \circ (x_i \circ x^{-1}) : x(A \cap A_i \cap \bar{A}) \to \bar{x}(A \cap A_i \cap \bar{A})$$

e é um difeomorfismo $\mathscr{C}^k$.

Agora, notemos que o conjunto $x(A \cap A_i \cap \bar{A})$ é um aberto de $\mathbb{R}^d$ já que $x(A \cap A_i)$ e $x_i(A_i \cap \bar{A})$ são abertos e $x_i \circ x^{-1}$ é contínua. Como

$$V = \bigcup_{i\in I} A_i,$$

segue que

$$\bigcup_{i\in I} x(A \cap A_i \cap \bar{A}) = x\left(A \cap \bigcup_{i\in I} A_i \cap \bar{A}\right) = x(A \cap \bar{A}),$$

então $x(A \cap \bar{A})$ é aberto de $\mathbb{R}^d$. Analogamente, o contradomínio de $\bar{x} \circ x^{-1}$ é $\bar{x}(A \cap \bar{A})$ e é aberto de $\mathbb{R}^d$. Portanto $\bar{x} \circ x^{-1}$ é uma função bem definida em $x(A \cap \bar{A})$ e segue que

$$\bar{x} \circ x^{-1} : x(A \cap \bar{A}) \to \bar{x}(A \cap \bar{A}).$$

é um difeomorfismo $\mathscr{C}^k$, o que mostra que (A, x) e $(\bar{A}, \bar{x})$ são $\mathscr{C}^k$-compatíveis, e finalmente $\mathcal{A}$ e $\bar{\mathcal{A}}$ são $\mathscr{C}^k$-compatíveis. ■

Isso implica, em particular, que os atlas $\mathscr{C}^k$ d-dimensionais de V podem ser particionados em classes de equivalência. Além disso, segue da proposição anterior que se dois atlas são compatíveis, sua união é um atlas, o que motiva a próxima definição.

:⊢ **Definição 21.6.** Seja V um conjunto. Um atlas *maximal* $\mathscr{C}^k$ (ou uma *estrutura diferencial* $\mathscr{C}^k$) d-dimensional de V é a união de todos os atlas de uma classe de equivalência de atlas $\mathscr{C}^k$ d-dimensionais de V.

Como comentado acima, o atlas maximal é um atlas, pois é uma união de atlas compatíveis. Com essa definição, podemos finalmente definir uma variedade.

21.1.2 Variedades e estrutura topológica e diferencial

:⊢ **Definição 21.7.** Uma *variedade* $\mathscr{C}^k$ d-dimensional é um par $\boldsymbol{V} = (V, \mathcal{A})$ em que V é um conjunto e $\mathcal{A}$ é um atlas maximal $\mathscr{C}^k$ d-dimensional de V. Para $k = 0$, a variedade $\boldsymbol{V}$ é uma variedade *topológica* e, para $k > 0$, é uma variedade *diferencial*, sendo para $k = \infty$ uma variedade *suave*.

Quando não for necessário explicitar detalhes, uma variedade $\mathscr{C}^k$ d-dimensional será simplesmente referida como uma variedade. Denotamos ainda a dimensão da variedade por um expoente: $\boldsymbol{V}^d$. É possível mostrar que toda variedade diferencial tem um subatlas maximal suave, de modo que não há perda de informação quando se separam as variedades apenas em topológicas e diferenciais, mas isso não será feito aqui.

:⊢ **Definição 21.8.** Seja $\boldsymbol{V}$ uma variedade. A *topologia* de $\boldsymbol{V}$ é o conjunto

$$\mathcal{T} := \left\langle \bigcup_{(A,x)\in\mathcal{A}} x^\star(\mathcal{T}_{\mathbb{R}^d}) \right\rangle,$$

em que $\mathcal{T}_{\mathbb{R}^d}$ é a topologia de $\mathbb{R}^d$.

Na notação acima, $\langle C \rangle$ denota a topologia gerada pelo conjunto $C \subseteq \wp(V)$ e $x^\star(\mathcal{T})$ denota a topologia puxada de $\mathcal{T}$ pela função x, ou seja, o conjunto de imagens inversas por x de abertos da topologia $\mathcal{T}$. Essa é a menor topologia tal que todos os mapas de coordenadas x das cartas do atlas maximal são contínuos, a topologia inicial com respeito aos mapas de coordenadas do atlas da variedade. Pode-se ver que essa é a topologia cuja base são os domínios das cartas do atlas maximal. Ainda, vale notar que essa é a mesma topologia da gerada pelas topologias puxadas de $\mathcal{T}|_{x(A)}$, a topologia induzida de $\mathcal{T}$ em $x(A)$. Isso ocorre porque os mapas de coordenadas x são bijeções no seu domínio, logo puxam qualquer subconjunto de $x(A)^\complement$ no vazio. De fato, há várias definições equivalentes de como induzir uma topologia em uma variedade, e a acima parece ser a mais bem definida em questão de teoria; no entanto, em geral só conseguimos conhecer um atlas de uma variedade, e não o atlas maximal todo, portanto saber gerar a topologia sem precisar do atlas maximal seria uma ferramenta muito boa, essencial às vezes. É isso que as proposições a seguir oferecem. A primeira afirma que, se gerarmos uma topologia a partir de uma atlas, qualquer carta compatível com esse atlas terá um mapa de coordenadas contínuo, o que implica, em particular, que as topologias geradas pelo atlas maximal e as geradas por qualquer subatlas são a mesma.

⊢ **Proposição 21.2.** *Sejam V um conjunto, $\mathcal{A} = \{(A_i, x_i)\}_{i\in I}$ um atlas de V e*

$$\mathcal{T} := \left\langle \bigcup_{i\in I} {x_i}^\star(\mathcal{T}_{\mathbb{R}^d}) \right\rangle.$$

Se (A, x) é uma carta compatível com $\mathcal{A}$, então x é uma função contínua.

□ *Demonstração.* Seja U um aberto de $\mathbb{R}^d$. Para todo $i \in I$, definamos $U_i := (x \circ x_i^{-1})(x_i(A_i))$ e $S := \mathbb{R}^d \smallsetminus \bigcup_{i\in I} U_i$. Então

$$U = U \cap \mathbb{R}^d = \bigcup_{i\in I}(U \cap U_i) \cup (U \cap S).$$

Como $V = \bigcup_{i\in I} A_i$ e, para todo $i \in I$, $x^{-1}(U_i) = A \cap A_i$, então

$$\begin{aligned} x^{-1}(S) &= x^{-1}(\mathbb{R}^d) \cap x^{-1}\left(\left(\bigcup_{i\in I} U_i\right)^{\complement}\right) \\ &= A \cap \left(\bigcup_{i\in I} x^{-1}(U_i)\right)^{\complement} \\ &= A \cap \left(\bigcup_{i\in I} A \cap A_i\right)^{\complement} \\ &= A \cap (A \cap V)^{\complement} = A \cap (A)^{\complement} = \emptyset. \end{aligned}$$

Disso, segue que $x^{-1}(U \cap S) = \emptyset$ e, portanto,

$$\begin{aligned} x^{-1}(U) &= \bigcup_{i\in I} x^{-1}(U \cap U_i) \cup x^{-1}(U \cap S) \\ &= \bigcup_{i\in I}(A \cap A_i) \cup (A \cap \emptyset) \\ &= \bigcup_{i\in I} A \cap A_i. \end{aligned}$$

Mostraremos, agora, que $A \cap A_i$ é aberto para todo $i \in I$, e com isso concluiremos que x^{-1} é aberto. Para isso, notemos que

$$x_i(A \cap A_i) = x_i(A_i) \cap (x_i \circ x^{-1})(x(A)).$$

Como $x(A)$ é aberto, pois (A, x) é carta, e $x \circ x_i^{-1}$ é difeomorfismo, pois (A, x) é compatível com $\mathcal{A}$, então $(x_i \circ x^{-1})(x(A))$ é aberto e, portanto, $x_i(A \cap A_i)$ é aberto, o que implica que $x_i^{-1}(x_i(A \cap A_i)) = A \cap A_i$ é aberto. Assim, concluímos que $x^{-1}(U)$ é aberto de $\mathcal{T}$, portanto x é contínua. ■

⊢ **Proposição 21.3.** *Sejam $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade e $\bar{\mathcal{A}}$ um atlas compatível com o atlas maximal $\mathcal{A}$. A topologia*

$$\bar{\mathcal{T}} := \left\langle \bigcup_{(A,x)\in\bar{\mathcal{A}}} x^{\star}(\mathcal{T}_{\mathbb{R}^d}) \right\rangle,$$

é igual à topologia $\mathcal{T}$ de $\boldsymbol{V}$.

□ *Demonstração.* Por definição da topologia de $\boldsymbol{V}$, temos imediatamente que $\bar{\mathcal{T}} \subseteq \mathcal{T}$. Para a inclusão contrária, a proposição anterior mostra que toda carta de $\mathcal{A}$ tem mapa de coordenadas contínuo em $\bar{\mathcal{T}}$, o que implica que $\mathcal{T} \subseteq \bar{\mathcal{T}}$. ■

Uma consequência óbvia das definições é que qualquer subconjunto aberto de $\mathbb{R}^d$ é uma variedade. Antes de continuar a discussão sobre as propriedades topológicas de variedades, um comentário importante é que alguns autores definem uma variedade a partir de um espaço topológico desde o início, exigindo que os mapas de coordenadas sejam homeomorfismos com sua imagem no espaço real. A partir dessa definição, a topologia da variedade já existe desde o começo e não é induzida pelas cartas.

▷ **Exercício 21.1.** *Sejam* $(\boldsymbol{V}_i)_{i\in[n]} = ((V_i, \mathcal{A}_i))_{i\in[n]}$ *variedades* $\mathscr{C}^k$ *de dimensões* $(d_i)_{i\in[n]}$*, respectivamente. O par*

$$\prod_{i\in[n]} V_i = \left(\prod_{i\in[n]} V_i, \prod_{i\in[n]} \mathcal{A}_i\right),$$

em que $\prod_{i\in[n]} V_i$ *é o produto de conjuntos e*

$$\prod_{i\in[n]} \mathcal{A}_i := \left\{ \left(\prod_{i\in[n]} A_i, \prod_{i\in[n]} x_i\right) \,\middle|\, \forall_{i\in[n]} (A_i, x_i) \in \mathcal{A}_i \right\}$$

é uma variedade $\mathscr{C}^k$ *de dimensão* $+_{i\in[n]} d_i$

21.1.3 Exemplos de variedades

21.1.3.1 Esfera

A esfera d-dimensional é o conjunto

$$\mathbb{S}^d = \left\{ x \in \mathbb{R}^{d+1} \,\middle|\, \|x\| = 1 \right\}.$$

A partir dessa notação de esfera unitária, podemos escrever facilmente qualquer n-esfera de raio $r \in \mathbb{R}$ e centro $c \in \mathbb{R}^{d+1}$ em $\mathbb{R}^{d+1}$ como $c + r\mathbb{S}^d$ usando a notação de adição e multiplicação de um elemento de um anel com um subconjunto do anel.

Vamos considerar, na esfera, um atlas com apenas duas cartas: as *projeções estereográficas* focadas no norte e no sul.

:⊢ **Definição 21.9.** O *polo norte* de $\mathbb{S}^d$ é o ponto $N := \mathrm{i}_d = (0, \ldots, 0, 1)$ e a *projeção estereográfica norte* de $\mathbb{S}^d$ é a função

$$\begin{aligned} \pi_N \colon \mathbb{S}^d \smallsetminus \{N\} &\longrightarrow \mathbb{R}^d \\ x &\longmapsto \frac{1}{1 - x_d}(x_0, \ldots, x_{d-1}). \end{aligned}$$

O *polo sul* de $\mathbb{S}^d$ é o ponto $S := -N == -\mathrm{i}_d = (0, \ldots, 0, -1)$ e a *projeção estereográfica sul* de $\mathbb{S}^d$ é a função

$$\begin{aligned} \pi_S \colon \mathbb{S}^d \smallsetminus \{S\} &\longrightarrow \mathbb{R}^d \\ x &\longmapsto \frac{1}{1 + x_d}(x_0, \ldots, x_{d-1}). \end{aligned}$$

⊢ **Proposição 21.4.** *Sejam π_N e π_S as projeções estereográficas norte e sul de $\mathbb{S}^d$. O conjunto*

$$\mathcal{A} := \{(\mathbb{S}^d \smallsetminus \{N\}, \pi_N), (\mathbb{S}^d \smallsetminus \{S\}, \pi_S)\}$$

é um atlas de $\mathbb{S}^d$.

□ *Demonstração.* As inversas de π_N e π_S são dadas por

$$\begin{aligned} \pi_N^{-1} \colon \mathbb{R}^d &\longrightarrow \mathbb{S}^d \\ x &\longmapsto \frac{1}{\|x\|^2 + 1}\left(2x_0, \ldots, 2x_{d-1}, \|x\|^2 - 1\right) \end{aligned}$$

e

$$\begin{aligned} \pi_S^{-1} \colon \mathbb{R}^d &\longrightarrow \mathbb{S}^d \\ x &\longmapsto \frac{-1}{\|x\|^2 + 1}\left(2x_0, \ldots, 2x_{d-1}, \|x\|^2 - 1\right). \end{aligned}$$

A transição de coordenadas entre as cartas é dada por

$$\begin{aligned} \pi_N \circ \pi_S^{-1} \colon \mathbb{R}^d \smallsetminus \{0\} &\longrightarrow \mathbb{R}^d \smallsetminus \{0\} \\ x &\longmapsto \frac{1}{\|x\|^2} x. \end{aligned}$$

Claramente, todas as funções são suaves. ■

FIGURA 21.1: Esfera 2-dimensional $\mathbb{S}^2$

21.1.3.2 Espaço projetivo

:⊢ **Definição 21.10.** Sejam $x, y \in \mathbb{R}^d \setminus \{0\}$. A relação de *equivalência homogênea* entre os pontos é definida por

$$x : y \quad \Longleftrightarrow \quad \exists t \in \mathbb{R}^* \qquad x = ty.$$

Essa relação é realmente uma relação de equivalência, pois tomando $t = 1$ temos a reflexividade, tomando $\bar{t} = t^{-1}$ temos a simetria e tomando $t_1 t_0$ temos a transitividade. A classe de equivalência de um ponto $x \in \mathbb{R}^d$ é a reta sem a origem de $\mathbb{R}^d$ que passa pela origem e por x, e é denotada por $[x_0 : \cdots : x_{d-1}] := [x]$.

:⊢ **Definição 21.11.** O *espaço projetivo real* d-dimensional é o conjunto

$$\mathbb{PR}^d := \left\{ [x_0 : \cdots : x_d] \;\middle|\; x \in \mathbb{R}^{d+1} \right\}.$$

Considerando os conjuntos

$$A_i := \left\{ [x_0 : \cdots : x_d] \in \mathbb{PR}^d \;\middle|\; x_i \neq 0 \right\}.$$

e as funções

$$\begin{aligned} \varphi_i \colon A_i &\longrightarrow \mathbb{R}^d \\ [x_0 : \cdots : x_d] &\longmapsto \frac{1}{x_i}(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d), \end{aligned}$$

cujas inversas são

$$\begin{aligned}\varphi_i^{-1}\colon \mathbb{R}^d &\longrightarrow A_i\\ x &\longmapsto [x_0 : \ldots, x_{i-1} : 1 : x_i :, \ldots : x_{d-1}].\end{aligned}$$

Uma descrição equivalente é

$$\mathbb{PR}^d := \mathbb{S}^d \big/ \mathbb{S}^0 = \left\{ \{x, -x\} \;\middle|\; x \in \mathbb{S}^d \right\}.$$

Essa construção considera a ação de $\mathbb{S}^0$ em $\mathbb{S}^d$, a saber,

$$\begin{aligned}\times\colon \mathbb{S}^0 \times \mathbb{S}^d &\longrightarrow \mathbb{S}^d\\ (u, x) &\longmapsto ux.\end{aligned}$$

O espaço $\mathbb{S}^0 = \{1, -1\} \subseteq \mathbb{R}$ tem estrutura de grupo com a multiplicação de $\mathbb{R}$ e pode ser identificado com o grupo $\mathbb{Z}_2$ pelo isomorfismo de grupos

$$\begin{aligned}h\colon (\mathbb{Z}_2, +) &\longrightarrow (\mathbb{S}^0, \times)\\ n &\longmapsto (-1)^n,\end{aligned}$$

que é homomorfismo pois $(-1)^{n+n'} = (-1)^n(-1)^{n'}$ e é claramente bijeção.

21.1.3.3 Toro

:⊢ **Definição 21.12.** O *toro* d-dimensional é o conjunto

$$\mathbb{T}^d := \mathbb{R}^d \diagup \mathbb{Z}^d.$$

Figura 21.2: Toro

21.1.3.4 Hiperboloide

Relembremos que $\mathbb{R}^{1,d}$ é o espaço vetorial $\mathbb{R}^{1+d}$ munido da 2-forma simétrica não degenerada $\langle \cdot, \cdot \rangle_{1,d}$ de assinatura $(1, d)$, definida para todos $x, x' \in \mathbb{R}^{1,d}$ por

$$\langle x, x' \rangle_{1,d} = x_0 x'_0 - \mathop{+}_{k=1}^{d} x_k x'_k,$$

em que $x = (x_0, \ldots, x_d)$ e $x' = (x'_0, \ldots, x'_d)$. A forma quadrática associada a $\langle \cdot, \cdot \rangle_{1,d}$ é $[\![\cdot]\!]_{1,d}$, definida para todo $x \in \mathbb{R}^{1,d}$ por

$$[\![x]\!]_{1,d} = {x_0}^2 - \mathop{+}_{k=1}^{d} {x_k}^2.$$

:⊢ **Definição 21.13.** O *hiperboloide* d-dimensional é o conjunto

$$\mathbb{H}^d := \left\{ x \in \mathbb{R}^{1,d} \;\middle|\; [\![x]\!]_{1,d} = 1 \wedge x_0 > 0 \right\}.$$

21.1.4 Propriedades topológicas

Uma variedade definida como acima não precisa necessariamente ser um espaço separado (T_2), ou seja, espaço cujos pontos distintos são separados por vizinhanças. Em geral, na definição de variedade consideram-se espaços separados, mas ainda não faremos essa distinção. Ainda, outra hipótese comum é que a topologia admita base enumerável. Isso é equivalente, para uma variedade, a existir um subatlas enumerável.

⊢ **Proposição 21.5.** *Seja* $\boldsymbol{V} = (V, \mathcal{A})$ *uma variedade. Então*

1. $\boldsymbol{V}$ *é um espaço topológico acessível* (T_1)*;*
2. $\boldsymbol{V}$ *é um espaço topológico separado* (T_2) *se, e somente se, para todo* p, $p' \in V$ *distintos, existe* $(A, x) \in \mathcal{A}$ *tal que* p, $p' \in A$ *ou existem cartas* (A, x), $(A', x) \in \mathcal{A}$ *tais que* $A \cap A' = \emptyset$, $p \in A$ *e* $p' \in A'$*;*
3. $\boldsymbol{V}$ *tem base topológica enumerável se, e somente se, existe um subatlas enumerável de* $\mathcal{A}$*;*
4. *Cada componente conexa de* $\boldsymbol{V}$ *é conexa por caminhos.*

□ *Demonstração.* 1. Sejam $p, \bar{p}$ pontos distintos de V. Queremos mostrar que existe vizinhança de cada um que não contém o outro. Para isso, tomamos carta (A, x) em p. Se A não contém $\bar{p}$, então A é vizinhança de p que não contém $\bar{p}$; caso contrário, como $x(p)$ e $x(\bar{p})$ são distintos, pois x é injetiva, segue que existe vizinhança U de $x(p)$ que não contém $x(p)$, pois $\mathbb{R}^d$ é acessível, logo $x^{-1}(p)$ é uma vizinhança de p que não contém $\bar{p}$. Analogamente, mostramos o mesmo para $\bar{p}$, portanto V eles são separados e concluímos que $\boldsymbol{V}$ é um espaço topológico acessível.

2. Exercício.
3. ($\Rightarrow$) Como $\boldsymbol{V}$ tem base enumerável, então toda cobertura de V tem subcobertura enumerável. Sendo assim, dada a cobertura por abertos das cartas de $\mathcal{A}$, tomamos uma subcobertura enumerável e segue que esse é um subatlas enumerável de $\mathcal{A}$.
   ($\Leftarrow$) Seja $\bar{\mathcal{A}}$ um subatlas enumerável de $\mathcal{A}$. Para cada $(A, x) \in \bar{\mathcal{A}}$, consideramos as bolas $\mathrm{B}(x; r) \subseteq x(A) \subseteq \mathbb{R}^d$ tais que $r \in \mathbb{Q}$ e $x \in \mathbb{Q}^d$. A união das imagens inversas dessas bolas pelos mapas de coordenadas é o conjunto enumerável
   $$\bar{\mathcal{B}} = \bigcup_{(A,x)\in\bar{\mathcal{A}}} \left\{ x^{-1}(\mathrm{B}(x; r)) \;\middle|\; \mathrm{B}(x; r) \subseteq x(A), r \in \mathbb{Q}, x \in \mathbb{Q}^d \right\}.$$
   Esse conjunto é uma base de $\boldsymbol{V}$: (1) O conjunto $\bar{\mathcal{B}}$ cobre V, pois os domínios das cartas de $\bar{\mathcal{A}}$ cobrem V e as imagens inversas das bolas de centro e raio racionais cobrem o domínio de cada carta de $\bar{\mathcal{A}}$; (2) Sejam A_1 e $A_2 \in \bar{\mathcal{B}}$. Então $A_1 \cap A_2$ é aberto e, para todo $p \in A_1 \cap A_2$, existem $r \in \mathbb{Q}$ e $x \in \mathbb{Q}^d$ tais que $x_1(p) \in \mathrm{B}(x; r) \subseteq x_1(A_1 \cap A_2)$, logo $p \in {x_1}^{-1}(\mathrm{B}(x; r)) \subseteq A_1 \cap A_2$.
4. Sejam $C(p)$ uma componente conexa de $\boldsymbol{V}$ em p e U o conjunto de todos pontos de $C(p)$ ligados por um caminho a p. Mostraremos que U é aberto e fechado, portanto igual a $C(p)$. Mostremos primeiro que U é aberto. Seja $q \in C(p)$ e A uma vizinhança de q homeomorfa a uma bola de $\mathbb{R}^d$. Como a bola é conexa por caminhos e homeomorfa a A, segue que A é conexa por caminhos. Portanto todo ponto de A está em $C(p)$, já que existe caminho do um ponto de A a q e caminho de q a p. Isso mostra que U é aberto. Agora, seja $q \in U^{\complement}$. Novamente, tomamos uma vizinhança A de q homeomorfa a uma bola de $\mathbb{R}^d$ e segue que A é conexa por caminhos. Nesse caso, nenhum ponto de A está em $C(p)$, caso contrário existiria caminho ligando q a p, o que contradiz a hipótese. Isso mostra que $U^{\complement}$ é fechado, portanto U aberto. Assim, como U é aberto e fechado, $C(p)$ é conexo e $p \in U$, segue que $C(p) = U$.

■

21.2 Funções diferenciáveis

21.2.1 Funções diferenciáveis

:⊢ **Definição 21.14.** Sejam $\boldsymbol{V_0}$ e $\boldsymbol{V_1}$ $\mathscr{C}^k$-variedades de dimensões d_0 e d_1, respectivamente. Uma função *$\mathscr{C}^k$-diferenciável* de $\boldsymbol{V_0}$ para $\boldsymbol{V_1}$ é uma função $F\colon V_0 \to V_1$ que satisfaz: para todo $p \in V_0$, existem cartas $(A_0, x_0) \in \mathcal{A}_0$ em p e $(A_1, x_1) \in \mathcal{A}_1$

em $F(p)$ tais que $F(A_0) \subseteq A_1$ e

$$x_1 \circ F \circ x_0^{-1} : x_0(A_0) \to x_1(A_1)$$

é uma função $\mathscr{C}^k$-diferenciável de $x_0(A_0) \subseteq \mathbb{R}^{d_0}$ para $x_1(A_1) \subseteq \mathbb{R}^{d_1}$. Denota-se $F\colon \boldsymbol{V_0} \to \boldsymbol{V_1}$. O conjunto das funções $\mathscr{C}^k$-diferenciáveis de $\boldsymbol{V_0}$ para $\boldsymbol{V_1}$ é denotado $\mathscr{C}^k(\boldsymbol{V_0}, \boldsymbol{V_1})$.

A diferenciabilidade independe da carta escolhida no seguinte sentido.

□ *Demonstração.* Sejam (A_0, x_0), $(\bar{A}_0, \bar{x}_0)$ cartas em p e (A_1, x_1), $(\bar{A}_1, \bar{x}_1)$ cartas em $f(p)$ tais que $F(A_0) \subseteq A_1$ e $F(\bar{A}_0) \subseteq \bar{A}_1$. Então $F(A_0 \cap \bar{A}_0) \subseteq A_1 \cap \bar{A}_1$, $p \in A_0 \cap \bar{A}_0$ e $F(p) \in A_1 \cap \bar{A}_1$. Restringindo domínios adequadamente,

$$\bar{x}_1 \circ F \circ \bar{x}_0^{-1} = (\bar{x}_1 \circ x_1^{-1}) \circ (x_1 \circ F \circ x_0^{-1}) \circ (x_0 \circ \bar{x}_0^{-1})$$

Como as transições de coordenadas $(\bar{x}_1 \circ x_1^{-1})$ e $(x_0 \circ \bar{x}_0^{-1})$ são $\mathscr{C}^k$-difeomorfismos, então $(x_1 \circ F \circ x_0^{-1})$ é $\mathscr{C}^k$-diferenciável se, e somente se, $(\bar{x}_1 \circ F \circ \bar{x}_0^{-1})$ o é. ■

Como uma $\mathscr{C}^k$-variedade é também uma $\mathscr{C}^l$-variedade para todo $l \leq k$, sempre se pode definir entre $\mathscr{C}^k$ e $\mathscr{C}^l$-variedade funções $\mathscr{C}^m$-diferenciáveis para qualquer $m \leq \min\{k, l\}$.

⊢ **Proposição 21.6.** *Sejam $\boldsymbol{V_0}$ e $\boldsymbol{V_1}$ $\mathscr{C}^k$-variedades e $F_0 : \boldsymbol{V_0} \to \boldsymbol{V_1}$ uma função $\mathscr{C}^k$-diferenciável. Então F é uma função contínua.*

⊢ **Proposição 21.7.** *Sejam $\boldsymbol{V_0}$, $\boldsymbol{V_1}$ e $\boldsymbol{V_2}$ $\mathscr{C}^k$-variedades e $F_0 : \boldsymbol{V_0} \to \boldsymbol{V_1}$ e $F_1 : \boldsymbol{V_1} \to \boldsymbol{V_2}$ funções $\mathscr{C}^k$-diferenciáveis. Então $F_1 \circ F_0 : \boldsymbol{V_0} \to \boldsymbol{V_1}$ é uma função $\mathscr{C}^k$-diferenciável.*

Casos particulares relevantes são quando uma das duas variedades é um subconjunto de $\mathbb{R}^d$. Toma-se assim a carta trivial nesse subconjunto. Um desses caso é evidenciado a seguir.

:⊢ **Definição 21.15.** Sejam $\boldsymbol{V}$ uma $\mathscr{C}^k$-variedade e $I \subseteq \mathbb{R}$ um intervalo aberto. Uma $\mathscr{C}^k$-*curva* em $\boldsymbol{V}$ é uma função $\gamma : I \to V$ $\mathscr{C}^k$-diferenciável.

Isso é equivalente a dizer que, para todo $p \in \gamma(V)$ existe carta (A, x) em p tal que

$$x \circ \gamma : I \to x(A)$$

é uma função $\mathscr{C}^k$-diferenciável, ao escolher acima a função identidade.

21.2.2 Funções separadoras e partições da unidade

Consideraremos inicialmente a função

$$\begin{aligned} f\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 0, & x \leq 0 \\ e^{-\frac{1}{x}}, & x > 0. \end{cases} \end{aligned}$$

Pode ser mostrado que essa função é suave, mas não faremos isso aqui. Sejam $r_0, r_1 \in \mathbb{R}$ tais que $r_0 < r_1$. Construímos a função

$$\begin{aligned} h\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{f(r_1 - x)}{f(x - r_0) - f(r_1 - x)}. \end{aligned}$$

Essa função é suave e tem a seguinte propriedade: $h(x) = 1$ para todo $x \leq r_0$, $0 < h(x) < 1$ para todo $x \in \,]r_0, r_1[$ e $h(x) = 0$ para todo $x \geq r_1$. Isso quer dizer que ela é uma função suave que separa (precisamente) os conjuntos $]-\infty, r_0]$ e $[r_1, +\infty[$. Em $\mathbb{R}^d$, podemos separar (precisamente) por função suave a bola fechada $\mathrm{B}\,(0; r_0)$ e o complementar da bola fechada $\mathrm{B}\,(0; r_1)$. Basta usar a função h para definir a função

$$\begin{aligned} H\colon \mathbb{R}^d &\longrightarrow \mathbb{R} \\ x &\longmapsto h(\|x\|). \end{aligned}$$

Uma função suave que separa esses conjuntos é geralmente chamada de uma "bump function", mas aqui a chamaremos de função separadora, pois ela separa (precisamente) os conjuntos. Em francês, essa função também é conhecida como uma "fonction plateau", algo como "função planalto". Essas funções são também chamadas de "funções teste". Mais geralmente, uma função como essa é uma função para $[0, 1]$ com suporte em um aberto A que vale 1 em um compacto K, ou seja, uma função contínua que separa um fechado $A^{\complement}$ e um compacto K. Essa funções sempre existem em espaços topológicos separados por vizinhanças (T_2) e localmente compactos.

:⊢ **Definição 21.16.** Sejam $\boldsymbol{X}$ um espaço topológico e $f : X \to \mathbb{R}$ uma função contínua. O *suporte fechado* de f é o conjunto

$$\overline{\mathrm{supp}}(f) := \overline{\mathrm{supp}(f)}.$$

⊢ **Proposição 21.8.** *Sejam $\boldsymbol{X}$ um espaço topológico separado por vizinhanças (T_2) e localmente compacto, $A \subseteq X$ um aberto e $K \subseteq A$ um compacto. Existe função $f : X \to [0, 1]$ tal que $\overline{\mathrm{supp}}(f) \subseteq A$ e $f(K) = \{1\}$.*

Notemos que tal função separa continuamente $A^{\complement}$ e K, pois

$$\operatorname{supp}(f) \subseteq \overline{\operatorname{supp}}(f) \subseteq A$$

implica $A^{\complement} \subseteq \operatorname{supp}(f)^{\complement}$, portanto $f(A^{\complement}) \subseteq f(\operatorname{supp}(f)^{\complement}) = \{0\}$.

:⊢ **Definição 21.17.** Sejam $\boldsymbol{X}$ um espaço topológico (variedade suave) e $\mathcal{C} = (C_i)_{i\in I}$ uma cobertura aberta de $\boldsymbol{X}$. Uma *partição da unidade* (*suave*) subordinada a $\mathcal{C}$ é uma família de funções contínuas (suaves) $\psi_i \colon X \to \mathbb{R}$ tal que

1. (Unidade) Para todos $i \in I$ e $x \in X$, $0 \leq \psi_i(x) \leq 1$;
2. (Partição) A família $(\overline{\operatorname{supp}}(\psi_i))_{i\in I}$ é localmente finita (todo ponto tem uma vizinhança que interseciona finitos dos conjuntos da família) e, para todo $x \in X$,
$$\mathop{+}_{i\in I} \psi_i(x) = 1.$$
3. (Subordinação) Para todo $i \in I$, $\overline{\operatorname{supp}}(\psi_i) \subseteq C_i$.

⊢ **Proposição 21.9.** *Sejam $\boldsymbol{V}$ uma variedade (suave) e $\mathcal{C} = (C_i)_{i\in I}$ uma cobertura aberta de $\boldsymbol{V}$. Existe partição da unidade (suave) $\psi_i \colon X \to \mathbb{R}$ subordinada a $\mathcal{C}$.*

21.3 Espaço tangente

Nesta seção, consideraremos somente variedade diferenciais, ou seja, variedades $\mathscr{C}^r$ para $r \geq 1$. Denotaremos isso dizendo que a variedade é $\mathscr{C}^{1+r}$, sem referência a r, ficando implícito que $r \geq 0$, logo $1 + r \geq 1$.

21.3.1 Espaço tangente em um ponto

Para definir a noção de um espaço tangente a uma variedade em um ponto, pensamos localmente através de cartas. Um vetor tangente é puxado do espaço normado modelo para a variedade usando uma carta, mas para que essa definição seja independente da carta escolhida devemos estabelecer uma relação de equivalência entre vetores tangentes da variedade modelo. Para ter a informação de que um vetor veio de uma carta, consideramos primeiro a noção de *vetor coordenado*.

:⊢ **Definição 21.18** (Vetores coordenados)**.** Seja $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$. Um *vetor coordenado* de $\boldsymbol{V}$ é uma tripla (A, x, u), em que $(A, x) \in \mathcal{A}$ é uma carta de $\boldsymbol{V}$ e $u \in E$ é um vetor.

Seja $p \in V$. Vetores coordenados *equivalentes em p* são vetores coordenados (A, x, u) e (A', x', u') tais que $p \in A \cap A'$ e

$$\mathrm{D}(x' \circ x^{-1})|_{x(p)}(u) = u'.$$

Denota-se $(A, x, u) \asymp_p (A', x', u')$.

Para mostrar que $\asymp_p$ é uma relação de equivalência, devemos considerar o conjunto $\mathcal{A}|_p$ das cartas em torno de p:

$$\mathcal{A}|_p = \{(A, x) \in \mathcal{A} \mid p \in A\}.$$

$\vdash$ **Proposição 21.10.** *Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$ e $p \in V$. A equivalência de vetores coordenados é uma relação de equivalência em $\mathcal{A}|_p \times E$.*

$\square$ *Demonstração.* Devemos demonstrar reflexividade, simetria e transitividade.

1. (Reflexividade) Para todo $(A, x, u) \in \mathcal{A}|_p \times E$, vale $p \in A \cap A$ e
$$\mathrm{D}(x \circ x^{-1})|_{x(p)}(u) = \mathrm{I}(u) = u,$$
portanto $(A, x, u) \asymp_p (A, x, u)$.
2. (Simetria) Para todos (A, x, u) e $(A', x', u') \in \mathcal{A}|_p \times E$ tais que $(A, x, u) \asymp_p (A', x', u')$, vale $p \in A \cap A' = A' \cap A$ e
$$\mathrm{D}(x' \circ x^{-1})|_{x(p)}(u) = u',$$
logo, como $x' \circ x^{-1}$ é difeomorfismo,
$$\mathrm{D}(x \circ x'^{-1})|_{x'(p)}(u') = u,$$
portanto $(A', x', u') \asymp_p (A, x, u)$.
3. (Transitividade) Para todos (A, x, u), (A', x', u') e $(A'', x'', u'') \in \mathcal{A}|_p \times E$ tais que $(A, x, u) \asymp_p (A', x', u')$ e $(A', x', u') \asymp_p (A'', x'', u'')$, vale $p \in A \cap A''$ e $p \in A' \cap A''$, logo $p \in A \cap A''$, e vale também $\mathrm{D}(x' \circ x^{-1})|_{x(p)}(u) = u'$ e $\mathrm{D}(x'' \circ x'^{-1})|_{x'(p)}(u') = u''$, logo segue da regra da cadeia que
$$\begin{aligned}\mathrm{D}(x'' \circ x^{-1})|_{x(p)}(u) &= \mathrm{D}(x'' \circ x'^{-1})|_{x'(p)} \circ \mathrm{D}(x' \circ x^{-1})|_{x(p)}(u) \\ &= \mathrm{D}(x'' \circ x'^{-1})|_{x'(p)}(u') \\ &= u'',\end{aligned}$$
portanto $(A, x, u) \asymp_p (A'', x'', u'')$. $\blacksquare$

Como $\asymp_p$ é uma relação de equivalência, podemos definir o conjunto quociente. Este será nosso espaço tangente à variedade no ponto p.

$:\vdash$ **Definição 21.19** (Espaço tangente)**.** Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$ e $p \in V$. O *espaço tangente a $\boldsymbol{V}$ em p* é o conjunto de classes de equivalência
$$\mathrm{T}V|_p := (\mathcal{A}|_p \times E)\big/_{\asymp_p}.$$

Um *vetor tangente a $\boldsymbol{V}$ em p* é um elemento de $v \in \mathrm{T}V|_p$. Denota-se $[u]_x|_p := [(A, x, u)]_{\asymp_p}$.

Podemos puxar a estrutura linear de $\boldsymbol{E}$ para $\mathrm{T}V|_p$ usando a carta (A, x) para relacionar $[(A, x, u)]_{\asymp_p}$ com u. Essa identificação não é independente da carta, mas a estrutura linear induzida no espaço tangente é. Primeiro, provemos que um vetor tangente à variedade em um ponto sempre tem um único representante em qualquer carta em torno daquele ponto.

$\vdash$ **Proposição 21.11.** *Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$, $p \in V$ e $v \in \mathrm{T}V|_p$. Para toda carta $(A, x) \in \mathcal{A}|_p$, existe um único $u \in E$ tal que $v = [u]_x|_p$. A função*

$$\begin{aligned} \mathrm{D}x|_p \colon \mathrm{T}V|_p &\longrightarrow E \\ [u]_x|_p &\longmapsto u \end{aligned}$$

está bem definida e é uma bijeção.

$\square$ *Demonstração.* Como $v \in \mathrm{T}V|_p$, existe vetor coordenado $(A', x', u') \in \mathcal{A}|_p \times E$ tal que $v = [(A', x', u')]_{\asymp_p}$. Como $p \in A \cap A'$, basta definir $u := \mathrm{D}(x \circ x'^{-1})|_{x'(p)}(u')$, e então segue que $(A', x', u') \asymp_p (A, x, u)$. Como $x \circ x'^{-1}$ é difeomorfismo, então $\mathrm{D}(x \circ x'^{-1})|_{x'(p)}$ é um isomorfismo linear, portanto tal u é único.

Isso mostra que a função $\mathrm{D}x|_p \colon \mathrm{T}V|_p \longrightarrow E$ está bem definida. (Injetividade) Sejam $v, v' \in \mathrm{T}V|_p$ tais que $\mathrm{D}x|_p(v) = \mathrm{D}x|_p(v') \in E$. Pelo que foi demonstrado, existem $u, u' \in E$ tais que $v = [u]_x|_p$ e $v' = [u']_x|_p$. Mas então

$$u = \mathrm{D}x|_p([u]_x|_p) = \mathrm{D}x|_p(v) = \mathrm{D}x|_p(v') = \mathrm{D}x|_p([u']_x|_p) = u',$$

portanto $v = [u]_x|_p = [u']_x|_p = v'$. (Sobrejetividade) Seja $u \in E$. Se definirmos $v := [u]_x|_p$, segue que $\mathrm{D}x|_p(v) = u$.

$\blacksquare$

$\vdash$ **Definição 21.20.** Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$ e $p \in V$. A *adição* em $\mathrm{T}V|_p$ é

$$\begin{aligned} + \colon \mathrm{T}V|_p \times \mathrm{T}V|_p &\longrightarrow \mathrm{T}V|_p \\ ([u]_x|_p, [u']_x|_p) &\longmapsto [u + u']_x|_p. \end{aligned}$$

A *inversão aditiva* em $\mathrm{T}V|_p$ é

$$\begin{aligned} - \colon \mathrm{T}V|_p &\longrightarrow \mathrm{T}V|_p \\ [u]_x|_p &\longmapsto [-u]_x|_p. \end{aligned}$$

O *zero* em $\mathrm{T}V|_p$ é

$$0 := [0]_x|_p.$$

A *multiplicação por escalar* em $\mathrm{T}V|_p$ é

$$\begin{aligned} \cdot\colon \mathbb{R} \times \mathrm{T}V|_p &\longrightarrow \mathrm{T}V|_p \\ (c, [u]_x|_p) &\longmapsto [cu]_x|_p. \end{aligned}$$

Denota-se $\mathbf{T}\boldsymbol{V}|_{\boldsymbol{p}} := (\mathrm{T}V|_p, +, -, 0, \cdot)$.

▷ **Exercício 21.2.** *Prove que as operações em* $\mathrm{T}V|_p$ *estão bem definidas (independem do representante na carta* (A, x) *escolhida) e que* $\mathbf{T}\boldsymbol{V}|_{\boldsymbol{p}}$ *é um espaço linear sobre* $\mathbb{R}$.

21.3.2 Diferencial de uma função

Para definir a diferencial de uma função diferenciável, como de costume usamos cartas para diferenciar a função localmente, e então puxamos essa estrutura de volta para a variedade.

:⊢ **Definição 21.21.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades $\mathscr{C}^{1+r}$ modeladas em espaços normados $\boldsymbol{E}$ e $\boldsymbol{E}'$, respectivamente, $p \in V$ e $f\colon V \longrightarrow V'$ uma função $\mathscr{C}^{1+r}$-diferenciável. A *diferencial de* f *em* p é a função

$$\begin{aligned} \mathrm{D}f|_p\colon \mathrm{T}V|_p &\longrightarrow \mathrm{T}V'|_{f(p)} \\ [u]_x|_p &\longmapsto [\mathrm{D}(x' \circ f \circ x^{-1})|_p(u)]_{x'}|_{f(p)} \end{aligned}$$

⊢ **Proposição 21.12.** *Sejam* $\boldsymbol{V}$ *e* $\boldsymbol{V}'$ *variedades* $\mathscr{C}^{1+r}$ *modeladas em espaços normados* $\boldsymbol{E}$ *e* $\boldsymbol{E}'$*, respectivamente,* $p \in V$ *e* $f\colon V \longrightarrow V'$ *uma função* $\mathscr{C}^{1+r}$*-diferenciável. A diferencial* $\mathrm{D}f|_p\colon \mathrm{T}V|_p \longrightarrow \mathrm{T}V'|_{f(p)}$ *está bem definida e é uma função linear.*

⊢ **Proposição 21.13** (Regra da Cadeia)**.** *Sejam* $\boldsymbol{V}$*,* $\boldsymbol{V}'$ *e* $\boldsymbol{V}''$ *variedades* $\mathscr{C}^{1+r}$ *modeladas em espaços normados* $\boldsymbol{E}$*,* $\boldsymbol{E}'$ *e* $\boldsymbol{E}''$*, respectivamente,* $p \in V$ *e* $f\colon V \longrightarrow V'$ *e* $f'\colon V' \longrightarrow V''$ *funções* $\mathscr{C}^{1+r}$*-diferenciáveis. Então*

$$\mathrm{D}(f' \circ f)|_p = \mathrm{D}f'|_{f(p)} \circ \mathrm{D}f|_p.$$

Note que uma carta $x\colon A \subseteq V \longrightarrow E$ é diferenciável e que

$$\begin{aligned} \mathrm{D}x|_p\colon \mathrm{T}V|_p &\longrightarrow E \\ [u]_x|_p &\longmapsto u \end{aligned}$$

estabelece um isomorfismo entre $\mathrm{T}V|_p$ e E que associa a cada vetor tangente $v \in \mathrm{T}V|_p$ o vetor $\mathrm{D}x|_p(v) \in E$, e a cada vetor $u \in E$, o vetor $[u]_x|_p = \mathrm{D}x^{-1}|_{x(p)}(u) \in \mathrm{T}V|_p$. Em particular isso mostra que os espaços tangentes de uma variedade têm a mesma dimensão que o espaço normado no qual a variedade é modelada.

21.3.3 Fibrado tangente

:⊢ **Definição 21.22.** Seja $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ modelada num um espaço normado $\boldsymbol{E}$. O *fibrado tangente* de $\boldsymbol{V}$ é

$$\mathrm{T}V := \bigsqcup_{p \in V} \mathrm{T}V|_p = \{(p, v) \mid p \in V, v \in \mathrm{T}V|_p\},$$

a união disjunta dos espaços tangentes a $\boldsymbol{V}$ em todos pontos de V. A *projeção canônica* do fibrado $\mathrm{T}V$ é a função

$$\begin{aligned} \mathrm{p}\colon \mathrm{T}A &\longrightarrow A \\ (p, v) &\longmapsto p. \end{aligned}$$

Seja $(A, x) \in \mathcal{A}$. A *carta de* $\mathrm{T}V$ *induzida por* (A, x) é o par $(\mathrm{T}A, \bar{x})$, em que

$$\mathrm{T}A := \{(p, v) \mid p \in A, v \in \mathrm{T}V|_p\} \subseteq \mathrm{T}V$$

e

$$\begin{aligned} \bar{x}\colon \mathrm{T}A &\longrightarrow E \times E \\ (p, v) &\longmapsto (x(p), \mathrm{D}x|_p(v)). \end{aligned}$$

O *atlas de* $\mathrm{T}V$ *induzido por* $\mathcal{A}$ é o conjunto

$$\bar{\mathcal{A}} := \{(\mathrm{T}A, \bar{x}) \mid (A, x) \in \mathcal{A}\}.$$

⊢ **Proposição 21.14.** *Seja* $\boldsymbol{V} = (V, \mathcal{A})$ *uma variedade* $\mathscr{C}^{1+r}$ *modelada num um espaço normado* $\boldsymbol{E}$. *O atlas induzido* $\bar{\mathcal{A}}$ *é atlas* $\mathscr{C}^r$ *de* $\mathrm{T}V$ *modelado em* $\boldsymbol{E} \times \boldsymbol{E}$.

□ *Demonstração.* Primeiro mostremos que, para toda carta $(A, x) \in \mathcal{A}$, a carta induzida $(\mathrm{T}A, \bar{x}) \in \bar{\mathcal{A}}$ é um carta de $\mathrm{T}V$ modelada em $\boldsymbol{E} \times \boldsymbol{E}$.

- ($\bar{x}$ é injetiva) Sejam $(p, v), (p', v') \in \mathrm{T}A$ tais que $\bar{x}(p, v) = \bar{x}(p', v')$. Então $x(p) = x(p')$ e $\mathrm{D}x|_p(v) = \mathrm{D}x|_{p'}(v')$. Como (A, x) é carta, x é injetiva, portanto $p = p'$, o que implica que $\mathrm{D}x|_p(v) = \mathrm{D}x|_p(v')$; como $\mathrm{D}x|_p$ é injetiva, então $v = v'$, portanto $(p, v) = (p', v')$.
- ($\bar{x}(\mathrm{T}A)$ é aberto em $E \times E$) Notemos que

  $$\bar{x}(\mathrm{T}A) = x(A) \times E.$$

  Seja $(y, u) \in E \times E$. Note que $(y, u) \in \bar{x}(\mathrm{T}A)$ se, e somente se, $(p, v) \in \mathrm{T}A$ tal que $\bar{x}(p, v) = (y, u)$. Como $\bar{x}(p, v) = (x(p), \mathrm{D}x|_p(v))$, isso é equivalente a existirem $p \in A$ e $v \in \mathrm{T}V|_p$ tais que $y = x(p)$ e $u = \mathrm{D}x|_p(v)$. Como $\mathrm{D}x|_p\colon \mathrm{T}V|_p \longrightarrow E$ é isomorfismo, isso é o mesmo que existir $p \in A$ tal que $y = x(p)$, o que equivale a $(y, u) \in x(A) \times E$. Sendo assim, como (A, x) é carta, $x(A)$ é aberto em E, portanto $\bar{x}(\mathrm{T}A) = x(A) \times E$ é aberto em $E \times E$.

Agora, devemos mostrar que $\bar{\mathcal{A}}$ é um atlas.

- (Cobertura) Por definição temos $\bigcup_{(\mathrm{T}A,\bar{x})\in\bar{\mathcal{A}}} \mathrm{T}A \subseteq \mathrm{T}V$. Para a inclusão contrária, seja $(p,v) \in \mathrm{T}V$. Como $p \in V$ e $\mathcal{A}$ é atlas, existe carta (A,x) tal que $p \in A$. Como e $v \in \mathrm{T}V|_p$, então $\mathrm{D}x|_p(v) \in E$, portanto $v = [\mathrm{D}x|_p(v)]|_x|_p$, o que mostra que $(p,v) \in \mathrm{T}A$.
- (Compatibilidade $\mathscr{C}^{1+r}$) Sejam $(\mathrm{T}A,\bar{x})$ e $(\mathrm{T}A',\bar{x}') \in \bar{\mathcal{A}}$ cartas induzidas por (A,x) e $(A',x') \in \mathcal{A}$, respectivamente. Como $\mathcal{A}$ é atlas $\mathscr{C}^{1+r}$, as cartas (A,x) e (A',x') são $\mathscr{C}^{1+r}$-compatíveis.
  Primeiro mostremos que os conjuntos $\bar{x}(\mathrm{T}A\cap\mathrm{T}A')$ e $\bar{x}'(\mathrm{T}A\cap\mathrm{T}A')$ são abertos em $E\times E$. Notemos que
  $$\bar{x}(\mathrm{T}A\cap\mathrm{T}A') = x(A\cap A')\times E.$$
  Seja $(y,u) \in E\times E$. Note que $(y,u) \in \bar{x}(\mathrm{T}A\cap\mathrm{T}A')$ se, e somente se, existe $(p,v) \in \mathrm{T}A\cap\mathrm{T}A'$ tal que $\bar{x}(p,v) = (y,u)$. Como $\bar{x}(p,v) = (x(p), \mathrm{D}x|_p(v))$, isso é equivalente a existirem $p \in V$ e $v \in \mathrm{T}V|_p$ tais que $p \in A$, $p \in A'$, $y = x(p)$ e $u = \mathrm{D}x|_p(v)$. Como $\mathrm{D}x|_p\colon \mathrm{T}V|_p \longrightarrow E$ é isomorfismo, isso é o mesmo que existir $p \in A\cap A'$ tal que $y = x(p)$, o que equivale a $(y,u) \in x(A\cap A')\times E$. Da mesma forma, vale $\bar{x}'(\mathrm{T}A\cap\mathrm{T}A') = x'(A\cap A')\times E$. Sendo assim, como (A,x) e (A',x') são $\mathscr{C}^{1+r}$-compatíveis, os conjuntos $x(A\cap A')$ e $x'(A\cap A')$ são abertos em E, portanto os conjuntos $\bar{x}(\mathrm{T}A\cap\mathrm{T}A')$ e $\bar{x}'(\mathrm{T}A\cap\mathrm{T}A')$ são abertos em $E\times E$.
  Agora mostremos que a transição de coordenadas $\bar{x}'\circ\bar{x}^{-1}\colon \bar{x}(\mathrm{T}A\cap\mathrm{T}A') \longrightarrow \bar{x}'(\mathrm{T}A\cap\mathrm{T}A')$ é um difeomorfismo $\mathscr{C}^{1+r}$. Como $\bar{x}(\mathrm{T}A\cap\mathrm{T}A') = x(A\cap A')\times E$, a inversa de $\bar{x}$ em $\mathrm{T}A\cap\mathrm{T}A'$ é dada por
  $$\begin{aligned}\bar{x}^{-1}\colon x(A\cap A')\times E &\longrightarrow \mathrm{T}A\cap\mathrm{T}A'\\ (y,v) &\longmapsto (x^{-1}(y), [v]|_x|_{x^{-1}(y)}).\end{aligned}$$
  Portanto, para todo $(y,v) \in x(A\cap A')\times E$,
  $$\begin{aligned}\bar{x}'\circ\bar{x}^{-1}(y,v) &= \bar{x}'((x^{-1}(y), [v]|_x|_{x^{-1}(y)})\\ &= (x'(x^{-1}(y)), \mathrm{D}x'|_{x^{-1}(p)}([v]|_x|_{x^{-1}(y)}))\\ &= (x'\circ x^{-1}(y), \mathrm{D}(x'\circ x^{-1})|_y(u))\end{aligned}$$
  Sendo assim, como (A,x) e (A',x') são $\mathscr{C}^{1+r}$-compatíveis, a transição de coordenadas $x'\circ x^{-1}\colon x(A\cap A') \longrightarrow x'(A\cap A')$ é um difeomorfismo $\mathscr{C}^{1+r}$, portanto sua diferencial $\mathrm{D}(x'\circ x^{-1})\colon x(A\cap A') \longrightarrow \mathscr{L}(E,E)$ é um difeomorfismo $\mathscr{C}^r$, o que implica que $\bar{x}'\circ\bar{x}^{-1}$ é um difeomorfismo $\mathscr{C}^r$, portanto as cartas $(\mathrm{T}A,\bar{x})$ e $(\mathrm{T}A',\bar{x}')$ são $\mathscr{C}^r$-compatíveis. ■

21.3.3.1 Referencial local coordenado

Dada uma carta (A, x) de V, as funções $\left(\frac{\partial}{\partial x^i}\Big|_p\right)_{i\in[d]}$, definidas por

$$\begin{aligned}\frac{\partial}{\partial x^i}\bigg|_p : \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathbb{R}\\ f &\longmapsto \partial_i(f\circ x^{-1})|_{x(p)}\end{aligned}$$

são uma base ordenada de $\mathrm{T}V|_p$. Essas funções são lineares porque, para todas $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$ e $c \in \mathbb{R}$,

$$\begin{aligned}\frac{\partial}{\partial x^i}\bigg|_p (cf + f') &= \partial_i((cf + f')\circ x^{-1})|_{x(p)}\\ &= \partial_i((cf\circ x^{-1}) + (f'\circ x^{-1}))|_{x(p)}\\ &= c\partial_i(f\circ x^{-1})|_{x(p)} + \partial_i(f'\circ x^{-1}))|_{x(p)}\\ &= c\,\frac{\partial}{\partial x^i}\bigg|_p f + \frac{\partial}{\partial x^i}\bigg|_p f,\end{aligned}$$

e são derivações porque, para todas $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$,

$$\begin{aligned}\frac{\partial}{\partial x^i}\bigg|_p (ff') &= \partial_i((ff')\circ x^{-1})|_{x(p)}\\ &= \partial_i((f\circ x^{-1})(f'\circ x^{-1}))|_{x(p)}\\ &= (f\circ x^{-1})|_{x(p)}(\partial_i(f'\circ x^{-1})|_{x(p)}\\ &\quad + (f'\circ x^{-1})|_{x(p)}(\partial_i(f\circ x^{-1})|_{x(p)}\\ &= f(p)\,\frac{\partial}{\partial x^i}\bigg|_p f + f'(p)\,\frac{\partial}{\partial x^i}\bigg|_p f'.\end{aligned}$$

Vale lembrar que $\partial_i f(p)$ é a derivada direcional da função f na direção de e_i no ponto p (também identificada com a derivada parcial $\mathrm{D}_i f(p)$) e

$$\partial_i f(p) = \mathrm{D}f(p)\cdot e_i.$$

A motivação da notação $\frac{\partial}{\partial x^i}$, além do uso histórico, vem da ideia de que vale a regra da cadeia

$$\partial_i(f\circ x^{-1})|_{x(p)}\text{‘=’}\partial_i f|_{x^{-1}(x(p))}\circ\partial_i(x^{-1})|_{x(p)}\text{‘=’}\partial_i f|_p\circ(\partial_i x|_p)^{-1}\text{‘=’}\,\frac{\partial_i f}{\partial_i x}\bigg|_p.$$

A função

$$\begin{aligned} \frac{\partial}{\partial x^i} \colon A &\longrightarrow \mathrm{T}V|_A \\ p &\longmapsto \left.\frac{\partial}{\partial x^i}\right|_p \end{aligned}$$

é um campo vetorial em A — uma seção de $\mathrm{T}V|_A$ — e $\frac{\partial}{\partial x} = \left(\frac{\partial}{\partial x^i}\right)_{i\in[d]}$ é um referencial local de A.

21.3.4 Álgebra dos campos escalares

Dada uma variedade diferencial $\boldsymbol{V}$, um *campo escalar* em $\boldsymbol{V}$ é uma função do espaço $\mathscr{C}^\infty(\boldsymbol{V}, \mathbb{R})$, as funções diferenciáveis de $\boldsymbol{V}$ para $\mathbb{R}$. Esse espaço será denotado $\mathfrak{T}^0(\boldsymbol{V})$[1]. O espaço $\mathfrak{T}^0(\boldsymbol{V})$ é um espaço vetorial sobre $\mathbb{R}$ com as operações de soma e produto por escalar induzidas pontualmente de $\mathbb{R}$: a *adição* é dada por

$$\begin{aligned} +\colon \mathfrak{T}^0(\boldsymbol{V}) \times \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^0(\boldsymbol{V}) \\ (f, f') &\longmapsto f + f' \colon V \longrightarrow \mathbb{R} \\ &\qquad\qquad p \longmapsto f(p) + f'(p), \end{aligned}$$

a *inversa* da adição é dada por

$$\begin{aligned} -\colon \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^0(\boldsymbol{V}) \\ f &\longmapsto -f \colon V \longrightarrow \mathbb{R} \\ &\qquad\qquad p \longmapsto -f(p), \end{aligned}$$

a *identidade* da adição é dada por

$$\begin{aligned} 0\colon V &\longrightarrow \mathbb{R} \\ p &\longmapsto 0 \end{aligned}$$

e o *produto por escalar* é dado por

$$\begin{aligned} \cdot\colon \mathbb{R} \times \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^0(\boldsymbol{V}) \\ (c, f) &\longmapsto cf \colon V \longrightarrow \mathbb{R} \\ &\qquad\qquad p \longmapsto cf(p). \end{aligned}$$

[1] Usualmente denota-se esse espaço por $\mathscr{C}^\infty(V)$, mas adotaremos essa notação por motivos que ficarão mais claros na seção de campos tensoriais mais à frente; em resumo, campos escalares são campos tensoriais de tipo $(0, 0)$

Ainda, pode-se definir um *produto* em $\mathfrak{T}^0(\boldsymbol{V})$, também induzido pontualmente de $\mathbb{R}$, dado por

$$\begin{aligned} \times\colon \mathfrak{T}^0(\boldsymbol{V}) \times \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^0(\boldsymbol{V}) \\ (f, f') &\longmapsto ff'\colon V \longrightarrow \mathbb{R} \\ &\qquad\qquad p \longmapsto f(p)f'(p). \end{aligned}$$

Esse produto é bilinear com respeito à estrutura linear de $\mathfrak{T}^0(\boldsymbol{V})$ e, portanto, faz desse espaço uma álgebra associativa e comutativa. Essa álgebra é a *álgebra de campos escalares* em $\boldsymbol{V}$. É simples ver que essas operações geram funções diferenciáveis de $\mathfrak{T}^0(\boldsymbol{V})$.

⊢ **Proposição 21.15.** *Seja $\boldsymbol{V}$ uma variedade diferencial. A lista*

$$\boldsymbol{\mathfrak{T}^0(V)} := (\mathfrak{T}^0(\boldsymbol{V}), +, -, 0, \times, 1, \cdot)$$

é uma álgebra associativa, comutativa e unitária sobre $\mathbb{R}$. Em particular,

$$(\mathfrak{T}^0(\boldsymbol{V}), +, -, 0, \times, 1)$$

é um anel.

21.3.5 Definições alternativas do espaço tangente

21.3.5.1 Derivações em pontos

:⊢ **Definição 21.23.** Sejam $\boldsymbol{V}$ uma variedade diferencial e $p \in V$. Uma *derivação* em p é um funcional linear $D\colon \mathfrak{T}^0(\boldsymbol{V}) \to \mathbb{R}$ que satisfaz, para todas $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$,

$$D(ff') = D(f)f'(p) + f(p)D(f').$$

O *espaço tangente* a $\boldsymbol{V}$ em p é o conjunto $\mathrm{T}V|_p$ de todas derivações em p e os *vetores tangentes* a V em p são os elementos de $\mathrm{T}V|_p$.

O espaço tangente $\mathrm{T}V|_p$ é um espaço vetorial. Para mostrar isso, devemos mostrar que ele é um subespaço vetorial do espaço dos funcionais lineares $\mathcal{L}(\mathfrak{T}^0(\boldsymbol{V}), \mathbb{R})$.

⊢ **Proposição 21.16.** *Sejam $\boldsymbol{V}$ uma variedade diferencial e $p \in V$. O espaço tangente a V em p é um subespaço vetorial de $\mathcal{L}(\mathfrak{T}^0(\boldsymbol{V}), \mathbb{R})$.*

□ *Demonstração.* Para isso, primeiro notamos que claramente $\mathrm{T}V|_p$ não é vazio, pois o funcional nulo que leva toda função diferenciável em $0 \in \mathbb{R}$ é uma derivação em p. Agora, para todos $v, v' \in \mathrm{T}V|_p$ e $c \in \mathbb{R}$, e todos $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$,

$$\begin{aligned} (v + v')(ff') &= v(ff') + v'(ff') \\ &= v(f)f'(p) + f(p)v(f') + v'(f)f'(p) + f(p)v'(f') \\ &= (v(f) + v'(f))f'(p) + f(p)(v(f') + v'(f')) \\ &= (v + v')(f)f'(p) + f(p)(v + v')(f') \end{aligned}$$

e

$$\begin{aligned}(cv)(ff') &= c(v(ff')) \\ &= c(v(f)f'(p) + f(p)v(f')) \\ &= cv(f)f'(p) + cf(p)v(f') \\ &= (cv)(f)f'(p) + f(p)(cv)(f'),\end{aligned}$$

o que mostra que $v + v'$ e cv são derivações em p. ■

Assim temos um espaço vetorial $\mathrm{T}V|_p$ associado a cada ponto p da variedade. As seguintes propriedades serão utilizadas na demonstração de algumas proposições mais à frente.

⊢ **Proposição 21.17.** *Sejam $\boldsymbol{V}$ uma variedade diferencial, $p \in V$ e $v \in \mathrm{T}V|_p$.*

1. *Para toda função constante $f \in \mathfrak{T}^0(\boldsymbol{V})$,*

$$v(f) = 0.$$

2. *Para todas funções $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$ tais que $f(p) = f'(p) = 0$,*

$$v(ff') = 0.$$

□ *Demonstração.* 1. Se f é constante, existe $c \in \mathbb{R}$ $f(p) = c$ para todo $p \in V$, portanto $f = c1_V$, em que 1_V é a função constante igual a 1. Notemos que

$$v(1_V) = v((1_V)^2) = v(1_V)1_V(p) + 1_V(p)v(1_V) = 2v(1_V),$$

o que implica $v(1_V) = 0$, portanto $v(f) = cv(1_V) = 0$.

2. Basta ver que

$$v(ff') = v(f)f'(p) + f(p)v(f') = 0.$$ ■

21.3.5.2 Diferencial de uma função diferenciável

:⊢ **Definição 21.24.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais, $F\colon V \to V'$ uma função diferenciável e $p \in V$. A *diferencial* de F em p é a função

$$\begin{aligned}\mathrm{D}F|_p\colon \mathrm{T}V|_p &\longrightarrow \mathrm{T}V'|_{F(p)} \\ v &\longmapsto \mathrm{D}F|_p v\colon \mathfrak{T}^0(\boldsymbol{V}') \longrightarrow \mathbb{R} \\ &\qquad\qquad\qquad f \longmapsto v(f \circ F).\end{aligned}$$

Pode-se denotar $\mathrm{D}F_p$ por simplicidade.

A diferencial $\mathrm{D}F|_p$ aplicada em um vetor tangente v é uma derivação em $F(p)$ de $v\mathfrak{T}^0(\boldsymbol{V}')$. Para mostrar isso, sejam $f, f' \in v\mathfrak{T}^0(\boldsymbol{V}')$ e $c \in \mathbb{R}$. Como v é linear,

$$\begin{aligned}(\mathrm{D}F|_p v)(f+f') &= v((f+f')\circ F)\\ &= v((f\circ F)+(f'\circ F))\\ &= v(f\circ F)+v(f'\circ F)\\ &= (\mathrm{D}F|_p v)(f)+(\mathrm{D}F|_p v)(f')\end{aligned}$$

e

$$\begin{aligned}(\mathrm{D}F|_p v)(cf) &= v((cf)\circ F)\\ &= v(c(f\circ F))\\ &= cv(f\circ F)\\ &= c(\mathrm{D}F|_p v)(f),\end{aligned}$$

o que mostra que $\mathrm{D}F|_p v$ é linear; como v é derivação em p,

$$\begin{aligned}(\mathrm{D}F|_p v)(ff') &= v((ff')\circ F)\\ &= v((f\circ F)(f'\circ F))\\ &= v(f\circ F)(f'\circ F)(p)+(f\circ F)(p)v(f'\circ F)\\ &= v((f\circ F))f'(F(p))+f(F(p))v(f'\circ F)\\ &= (\mathrm{D}F|_p v)(f)f'(F(p))+f(F(p))(\mathrm{D}F|_p v)(f')\end{aligned}$$

o que mostra que $\mathrm{D}F|_p v$ é derivação em $F(p)$.

⊢ **Proposição 21.18.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $F\colon V \to V'$ uma função diferenciável e $p \in V$. A diferencial $\mathrm{D}F|_p\colon \mathrm{T}V|_p \to \mathrm{T}V'|_{F(p)}$ é uma função linear.*

□ *Demonstração.* Sejam $v, v' \in \mathrm{T}V|_p$ e $c \in \mathbb{R}$. Então, para todo $f \in \mathfrak{T}^0(\boldsymbol{V})$,

$$\begin{aligned}(\mathrm{D}F|_p(v+v'))(f) &= (v+v')(f\circ F)\\ &= v(f\circ F)+v'(f\circ F)\\ &= (\mathrm{D}F|_p v)(f)+(\mathrm{D}F|_p v')(f)\end{aligned}$$

e

$$\begin{aligned}(\mathrm{D}F|_p(cv))(f) &= (cv)(f\circ F)\\ &= c(v(f\circ F))\\ &= (c\mathrm{D}F|_p v)(f),\end{aligned}$$

o que mostra que $\mathrm{D}F|_p(v+v') = \mathrm{D}F|_p v + \mathrm{D}F|_p v'$ e $\mathrm{D}F|_p(cv) = c\mathrm{D}F|_p v$. ■

⊢ **Proposição 21.19** (Regra da Cadeia)**.** *Sejam* $\boldsymbol{V}$*,* $\boldsymbol{V}'$ *e* $\boldsymbol{V}''$ *variedades diferenciais,* $F\colon V \to V'$ *e* $F'\colon V' \to V''$ *funções diferenciáveis, e* $p \in V$*. Então*

$$\mathrm{D}(F' \circ F)|_p = \mathrm{D}F'|_{F(p)} \circ \mathrm{D}F|_p.$$

□ *Demonstração.* Para todos $v \in \mathrm{T}V|_p$ e $f \in \mathfrak{T}^0(V'')$,

$$\begin{aligned}(\mathrm{D}(F' \circ F)|_p v)(f) &= v(f \circ (F' \circ F)) \\ &= v((f \circ F') \circ F)) \\ &= (\mathrm{D}F|_p v)(f \circ F') \\ &= ((\mathrm{D}F'|_{F(p)})(\mathrm{D}F|_p v))(f) \\ &= ((\mathrm{D}F'|_{F(p)} \circ \mathrm{D}F|_p)v)(f),\end{aligned}$$

portanto $\mathrm{D}(F' \circ F)|_p v = (\mathrm{D}F'|_{F(p)} \circ \mathrm{D}F|_p)v$ para todo $v \in \mathrm{T}V|_p$, o que implica

$$\mathrm{D}(F' \circ F)|_p = \mathrm{D}F'|_{F(p)} \circ \mathrm{D}F|_p. \qquad \blacksquare$$

21.3.5.3 Vetores tangentes e espaços tangentes ao $\mathbb{R}^d$

Os vetores dos espaços tangentes a $\mathbb{R}^d$ podem ser descritos como derivações. Seja $f \in \mathfrak{T}^0(\mathbb{R}^d)$ uma função escalar e consideremos sua diferencial $\mathrm{D}f|_p\colon \mathbb{R}^d \to \mathbb{R}$. Essa diferencial pode ser aplicada no vetor canônico $e_i \in \mathbb{R}^d$, definindo

$$\partial_i|_p f := \mathrm{D}f|_p e_i.$$

A função $\partial_i|_p\colon \mathfrak{T}^0(\mathbb{R}^d) \to \mathbb{R}$ assim definida é uma derivação em p. Ela é linear porque, para todas $f, f' \in \mathfrak{T}^0(\mathbb{R}^d)$ e $c \in \mathbb{R}$, segue da linearidade de D que

$$\begin{aligned}\partial_i|_p(cf + f') &= \mathrm{D}(cf + f')|_p e_i = (c\mathrm{D}f|_p + \mathrm{D}f'|_p)e_i \\ &= c\mathrm{D}f|_p e_i + \mathrm{D}f'|_p e_i \\ &= c\partial_i|_p(f) + \partial_i|_p(f'),\end{aligned}$$

e é uma derivação em p porque, para todas $f, f' \in \mathfrak{T}^0(\mathbb{R}^d)$, segue da regra do produto de D que

$$\begin{aligned}\partial_i|_p(ff') &= \mathrm{D}(ff')|_p e_i \\ &= (\mathrm{D}f|_p f'(p) + f(p)\mathrm{D}f'|_p)e_i \\ &= \mathrm{D}f|_p f'(p)e_i + f(p)\mathrm{D}f'|_p e_i \\ &= \partial_i|_p(f)f'(p) + f(p)\partial_i|_p(f').\end{aligned}$$

A derivação direcional de f an direção de $v = +_{i\in[d]} c^i e_i$ é

$$\mathrm{D}_v|_p f := \mathrm{D}f|_p v.$$

Note que, para todo $f \in \mathfrak{T}^0(\mathbb{R}^d)$,

$$\mathrm{D}_v|_p(f) = \mathrm{D}f|_p v = \mathrm{D}f|_p \left(\mathop{+}_{i\in[d]} v^i e_i \right) = \mathop{+}_{i\in[d]} v^i \mathrm{D}f|_p e_i = \mathop{+}_{i\in[d]} v^i \partial_i|_p(f),$$

logo $\mathrm{D}_v|_p = +_{i\in[d]} v^i \partial_i|_p$. Mostraremos que essa função é um isomorfismo de espaço lineares.

⊢ **Proposição 21.20.** *Seja $d \in \mathbb{N}$. Para todo $p \in \mathbb{R}^d$, a função*

$$\begin{aligned} \mathrm{D}.|_p \colon \mathbb{R}^d &\longrightarrow \mathrm{T}\mathbb{R}^d|_p \\ v &\longmapsto \mathrm{D}_v|_p \end{aligned}$$

é um isomorfismo de espaços lineares.

□ *Demonstração.* Seja $v = +_{i\in[d]} v^i e_i$ e, portanto, $\mathrm{D}_v|_p = +_{i\in[d]} v^i \partial_i|_p$. Primeiro mostramos que essa função é linear. Sejam $v, v' \in \mathbb{R}^d$ e $c \in \mathbb{R}$. Então, para toda $f \in \mathfrak{T}^0(\mathbb{R}^d)$, segue da linearidade de $\mathrm{D}f|_p$ que

$$\begin{aligned} \mathrm{D}_{cv+v'}|_p f &= \mathrm{D}f|_p(cv + v') \\ &= c\mathrm{D}f|_p v + \mathrm{D}f|_p v' \\ &= c\mathrm{D}_v|_p f + \mathrm{D}_{v'}|_p f \\ &= (c\mathrm{D}_v + \mathrm{D}_{v'})f. \end{aligned}$$

Agora, mostremos que ela é injetiva. Seja $v \in \mathbb{R}^d$ tal que $\mathrm{D}_v|_p = 0$. Então, para cada $i \in [d]$, como $\mathrm{D}\pi^i|_p = \pi^i$,

$$0 = \mathrm{D}_v|_p \pi^i = \mathrm{D}\pi^i|_p v = \pi^i(v) = v^i,$$

logo $v = 0$. Agora, mostremos que ela é sobrejetiva. Seja $D \in \mathrm{T}\mathbb{R}^d|_p$. Para cada $i \in [d]$, definimos $v^i := D(\pi^i)$ e $v := +_{i\in[d]} v^i e_i$. Mostraremos que $D = \mathrm{D}_v|_p$. Seja $f \in \mathfrak{T}^0(\mathbb{R}^d)$. Pela fórmula de Taylor, existem constantes $C_{i,j} \in \mathbb{R}$ tais que

$$f = f(p) + \mathop{+}_{i\in[d]} \partial_i f(p)(\pi^i - p^i) + \mathop{+}_{(i,j)\in[d]^2} C_{i,j}(\pi^i - p^i)(\pi^j - p^j).$$

Como $f(p)$ é constante, $D(f(p)) = 0$ (21.17), e, como as funções $(\pi^i - p^i)$ e $(\pi^j - p^j)$ se anulam em p, $D\left((\pi^i - p^i)(\pi^j - p^j)\right) = 0$ para todos $i, j \in [d]$ (21.17), logo

$$D\left(\mathop{+}_{(i,j)\in[d]^2} C_{i,j}(\pi^i - p^i)(\pi^j - p^j) \right) = \mathop{+}_{(i,j)\in[d]^2} C_{i,j} D\left((\pi^i - p^i)(\pi^j - p^j)\right) = 0,$$

o que implica que

$$\begin{aligned}
D(f) &= D\left(\mathop{+}_{i\in[d]} \partial_i f(p)(\pi^i - p^i)\right) \\
&= \mathop{+}_{i\in[d]} D\left(\partial_i f(p)(\pi^i - p^i)\right) \\
&= \mathop{+}_{i\in[d]} \left(\partial_i f(p) D(\pi^i - p^i)\right) \\
&= \mathop{+}_{i\in[d]} \partial_i f(p)(D(\pi^i) - D(p^i)) \\
&= \mathop{+}_{i\in[d]} \partial_i f(p) v^i \\
&= \mathrm{D}_v|_p f.
\end{aligned}$$

■

Essa proposição mostra que $(\partial_i|_p)_{i\in[d]}$ é uma base ordenada do espaço tangente $\mathrm{T}\mathbb{R}^d|_p$, o qual é isomorfo a $\mathbb{R}^d$ e deve ser visto como o espaço dos vetores tangentes a $\mathbb{R}^d$ em p.

21.3.5.4 Curvas equivelozes

O espaço tangente pode ser alternativamente descrito com curvas.

:⊢ **Definição 21.25.** Sejam $\boldsymbol{V}$ uma variedade diferencial e $p \in V$. Duas curvas $\gamma_0 \colon I_0 \to V$ e $\gamma_1 \colon I_1 \to V$ iniciadas em p ($\gamma(0) = p$) são *equivelozes* se, e somente se, existe uma carta (A, x) em p tal que

$$(x \circ \gamma_0)'(0) = (x \circ \gamma_1)'(0).$$

Denota-se $\gamma_0 \asymp \gamma_1$.

A relação está bem definida, no sentido de que não depende da escolha de carta, e é uma relação de equivalência.

□ *Demonstração.* Sejam (A, x), $(\bar{A}, \bar{x})$ cartas em p. Então, para $i = 0$ e $i = 1$,

$$\bar{x} \circ \gamma_i = (\bar{x} \circ x^{-1}) \circ (x \circ \gamma_i)$$

e, diferenciando essas funções em 0, obtemos pela regra da cadeia que

$$\begin{aligned}
(\bar{x} \circ \gamma_i)'(0) &= \mathrm{D}\Big((\bar{x} \circ x^{-1}) \circ (x \circ \gamma_i)\Big)(0) \\
&= \mathrm{D}(\bar{x} \circ x^{-1})(x \circ \gamma_i(0)) \circ (x \circ \gamma_i)'(0) \\
&= \mathrm{D}(\bar{x} \circ x^{-1})(x(p)) \circ (x \circ \gamma_i)'(0).
\end{aligned}$$

Como $\bar{x} \circ x^{-1}$ é difeomorfismo, $\mathrm{D}(\bar{x} \circ x^{-1})(x(p))$ é invertível, portanto

$$(x \circ \gamma_0)'(0) = (x \circ \gamma_1)'(0)$$

se, e somente se,

$$(\bar{x} \circ \gamma_0)'(0) = (\bar{x} \circ \gamma_1)'(0).$$

■

:⊢ **Definição 21.26.** Sejam $\boldsymbol{V}$ uma variedade diferencial, $p \in V$ e $\gamma : I \to V$ uma curva iniciada em p. O *vetor tangente a* $\boldsymbol{V}$ *em* p, denotado $\gamma'(0)$, é a classe de equivalência de γ sob a relação $\asymp$ de equivelocidade de curvas.

O *espaço tangente* a $\boldsymbol{V}$ em p é o conjunto de vetores tangentes a $\boldsymbol{V}$ em p, denotado

$$\mathrm{T}V|_p := \{\gamma'(0) \mid \gamma \colon I \to V, \gamma(0) = p\},$$

em que os I são intervalos abertos de $\mathbb{R}$ contendo o 0.

O *fibrado tangente* de $\boldsymbol{V}$ é a união disjunta dos espaços tangentes a $\boldsymbol{V}$ em todos pontos de V

$$\mathrm{T}V := \bigsqcup_{p \in V} \mathrm{T}V|_p = \{(p, v) \mid p \in V, v \in \mathrm{T}V|_p\}.$$

Os elementos do espaço tangente a $\boldsymbol{V}$ em um ponto $p \in V$ são chamados de vetores porque $\mathrm{T}_p\boldsymbol{V}$ é um espaço vetorial se definidas operações apropriadas. Isso que fazemos a seguir.

Para isso, vamos mostrar que $\mathrm{D}f(p) : \mathrm{T}_pV \to \mathbb{R}^n$ é uma bijeção. Assim poderemos puxar as operações de espaço vetorial de $\mathbb{R}^n$ para T_pV.

:⊢ **Definição 21.27.** Sejam $\boldsymbol{V}$ e $\bar{\boldsymbol{V}}$ variedades, $p \in V$ e $f : V \to \bar{V}$ uma função diferenciável em p. A *diferencial* de f em p é a função

$$\begin{aligned} \mathrm{D}f(p) \colon \mathrm{T}_pV &\longrightarrow \mathrm{T}_{f(p)}\bar{V} \\ \gamma'(0) &\longmapsto (f \circ \gamma)'(0). \end{aligned}$$

Devemos mostrar que essa função está bem definida.

Tomando uma carta (A, x) em p, definimos a função

$$\begin{aligned} \mathrm{D}x(p) \colon \mathrm{T}_p\boldsymbol{V} &\longrightarrow \mathbb{R}^d \\ \gamma'(0) &\longmapsto (x \circ \gamma)'(0) \end{aligned}$$

de modo que a regra da composição é simulada. Essa função é uma bijeção e é usada para puxar as operações de $\mathbb{R}^d$ para $\mathrm{T}_p\boldsymbol{V}$, de modo que $\mathrm{T}_p\boldsymbol{V}$ é um espaço vetorial. Essas operações puxadas não dependem da carta (A, x) escolhida.

21.4 Subvariedades

21.4.1 Imersão

21.4.2 Submersão

21.4.3 Mergulho

21.4.4 Subvariedades imersas e mergulhadas

21.5 Transversalidade

21.5.1 Conjuntos nulos

Comentaremos nesta seção sobre o conceito de conjuntos nulos. Esses conceitos precedem o conceito mais amplo de uma medida de conjuntos, e podem ser definidos em $\mathbb{R}^d$ independentemente das medidas mais usuais, como a de Borel e a de Lebesgue.

:⊢ **Definição 21.28.** Seja $d \in \mathbb{N}$. Um *cubo d-dimensional* é um produto de d intervalos de $\mathbb{R}$. O *volume d-dimensional* de um cubo d-dimensional C cujo interior é $C^\circ =]a_0, b_0[\times \cdots \times]a_{d-1}, b_{d-1}[$ é o número real

$$\mathrm{vol}^d(C) := |b_0 - a_0| \cdots |b_{d-1} - a_{d-1}| \,.$$

Admitimos nessa definição que $\emptyset$ é um intervalo aberto, portanto que é um cubo aberto, e temos que seu volume é 0.

:⊢ **Definição 21.29.** Seja $d \in \mathbb{N}$. Um conjunto *nulo* em $\mathbb{R}^d$ é um conjunto $N \subseteq \mathbb{R}^d$ tal que, para todo $\varepsilon > 0$, existe cobertura $(C_n)_{n\in\mathbb{N}}$ de N por cubos d-dimensionais tal que

$$\mathop{+}_{n\in\mathbb{N}} \mathrm{vol}^d(C_n) < \varepsilon.$$

Denota-se $N \mathrel{\overset{\circ}{=}} \emptyset$.

O motivo da notação $N \mathrel{\overset{\circ}{=}} \emptyset$ ficará mais claro quando uma medida for definida.

⊢ **Proposição 21.21.** *Seja $d \in \mathbb{N}$. O conjunto de conjuntos nulos é um σ-ideal:*

1. $\emptyset \mathrel{\overset{\circ}{=}} \emptyset$;
2. *Se $S \subseteq N \subseteq \mathbb{R}^d$ e $N \mathrel{\overset{\circ}{=}} \emptyset$, então $S \mathrel{\overset{\circ}{=}} \emptyset$;*
3. *Para toda sequência $(N_n)_{n\in\mathbb{N}}$ de subconjuntos nulos em $\mathbb{R}^d$,*

$$\bigcup_{n\in\mathbb{N}} N_n \mathrel{\overset{\circ}{=}} \emptyset.$$

□ *Demonstração.* 1. Basta considerar a sequência de conjuntos vazios $(\emptyset)_{n\in\mathbb{N}}$. Essa sequência cobre $\emptyset$ e $\mathrm{vol}^d(\emptyset) = 0$. Portanto, para todo $\varepsilon > 0$,

$$\mathop{+}_{n\in\mathbb{N}} \mathrm{vol}^d(\emptyset) = 0 < \varepsilon,$$

o que mostra que $\emptyset \mathrel{\mathring{=}} \emptyset$.

2. Seja $\varepsilon > 0$. Como $N \mathrel{\mathring{=}} \emptyset$, existe cobertura $(C_n)_{n\in\mathbb{N}}$ de N por cubos tal que $+_{n\in\mathbb{N}} \mathrm{vol}^d(C_n) < \varepsilon$. Como $S \subseteq N$, essa cobertura é também uma cobertura de S, logo $S \mathrel{\mathring{=}} \emptyset$.
3. Seja $\varepsilon > 0$. Para cada $n \in \mathbb{N}$, existe cobertura $(C_i^n)_{i\in\mathbb{N}}$ de N_n por cubos tal que $+_{i\in\mathbb{N}} \mathrm{vol}^d(C_i^n) < \varepsilon 2^{-(n+1)}$, logo $(C_i^n)_{(n,i)\in\mathbb{N}^2}$ é uma cobertura de $(N_n)_{n\in\mathbb{N}}$ e

$$\mathop{+}_{(n,i)\in\mathbb{N}^2} \mathrm{vol}^d(C_i^n) < \mathop{+}_{n\in\mathbb{N}} \varepsilon 2^{-(n+1)} = \varepsilon. \qquad \blacksquare$$

⊢ **Proposição 21.22.** *Sejam $d \in \mathbb{N}$ e $C \subseteq \mathbb{R}^d$ um conjunto contável. Então $C \mathrel{\mathring{=}} \emptyset$.*

□ *Demonstração.* Basta mostrar que um conjunto unitário é nulo, pois C é uma união contável de conjunto unitários, portanto nulo pela proposição anterior. Para isso, seja $p \in \mathbb{R}^d$ e $\epsilon > 0$. Tomando $\alpha \in \,]0,1[$ e definindo os cubos

$$C_0 := \bigtimes_{i\in[d]} \left]p_i - \frac{(\alpha\varepsilon)^{\frac{1}{d}}}{2}, p_i + \frac{(\alpha\varepsilon)^{\frac{1}{d}}}{2}\right[$$

e, para todo $n \in \mathbb{N}^*$, $C_n := \emptyset$, segue que $p \in \bigcup_{n\in\mathbb{N}} C_n$, pois $p \in C_0$. Como

$$\mathrm{vol}^d(C_0) = \bigtimes_{i\in[d]} \left| p_i + \frac{(\alpha\varepsilon)^{\frac{1}{d}}}{2} - \left(p_i - \frac{(\alpha\varepsilon)^{\frac{1}{d}}}{2}\right)\right| = \left((\alpha\varepsilon)^{\frac{1}{d}}\right)^d = \alpha\varepsilon$$

e $\mathrm{vol}^d(\emptyset) = 0$, segue que

$$\mathop{+}_{n\in\mathbb{N}} \mathrm{vol}^d(C_n) = \alpha\varepsilon < \varepsilon,$$

logo $\{p\} \mathrel{\mathring{=}} \emptyset$. $\blacksquare$

Agora, definiremos conjuntos nulos em uma variedade $\boldsymbol{V}$.

:⊢ **Definição 21.30.** Seja $\boldsymbol{V}$ uma variedade diferencial. Um conjunto *nulo* em $\boldsymbol{V}$ é um conjunto $N \subseteq V$ tal que, para toda carta (A, x) de $\boldsymbol{V}$, o conjunto $x(Q \cap A)$ é nulo em $\mathbb{R}^d$. Denota-se $N \mathrel{\mathring{=}} \emptyset$.

Note que se (A, x) e (A', x') são cartas, a função $x' \circ x^{-1}$ preserva conjuntos nulos, pois é difeomorfismo. Isso significa, em particular, que se a propriedade vale para um atlas de $\boldsymbol{V}$, vale para seu atlas maximal. Na próxima seção, enunciaremos proposições que envolvem conjuntos nulos.

21.5.2 Valor regular

:⊢ **Definição 21.31.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $F\colon V \to V'$ uma função diferenciável. Um *valor regular* de F é um ponto $q \in V'$ tal que, para todo $p \in F^{-1}(q)$, $\mathrm{D}F|_p\colon \mathrm{T}V|_p \to \mathrm{T}V'|_q$ é sobrejetiva.

⊢ **Proposição 21.23.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais, $F\colon V \to V'$ uma função diferenciável e $q \in V'$ um valor regular de F. Então $F^{-1}(q) \subseteq V$ é uma subvariedade diferencial de dimensão* $\dim(V) - \dim(V')$.

21.5.3 Ponto crítico

:⊢ **Definição 21.32.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $F\colon V \to V'$ uma função diferenciável. Um *ponto crítico* de F é um ponto $p \in V$ tal que $\mathrm{D}F|_p\colon \mathrm{T}V|_p \to \mathrm{T}V'|_{F(p)}$ não é sobrejetiva.

Ou seja, um valor regular é um ponto cujos pontos de sua imagem inversa não são críticos.

⊢ **Proposição 21.24** (Teorema de Sard)**.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $F\colon V \to V'$ diferenciável. Se o conjunto $C \subseteq V$ de pontos críticos de F é nulo em $\boldsymbol{V}$, então $F(C) \subseteq V'$ é nulo em $\boldsymbol{V}'$.*

21.5.4 Transversalidade

:⊢ **Definição 21.33.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $\boldsymbol{S} \subseteq \boldsymbol{V}'$ uma subvariedade. Uma função *transversal* a $\boldsymbol{S}$ é uma função diferenciável $F\colon V \to V'$ tal que, para todo $p \in F^{-1}(S) \subseteq V$,

$$\mathrm{D}F|_p(\mathrm{T}V|_p) + \mathrm{T}S|_{F(p)} = \mathrm{T}V'|_{F(p)}.$$

Denota-se $F \pitchfork S$.

Um caso particular dessa definição é quando S é um conjunto unitário $\{q\}$. Nesse caso, para todo $p \in V$ tal que $F(p) = q$, $\mathrm{T}S|_{F(p)} = \{0\}$ e a condição acima se torna $\mathrm{D}F|_p\mathrm{T}V|_p = \mathrm{T}V'|_{F(p)}$, o que é equivalente a $\mathrm{D}F|_p$ ser sobrejetiva, portanto $F \pitchfork \{p\}$ é equivalente a p ser valor regular de F. Segue da proposição análoga para valores regulares a seguinte proposição.

⊢ **Proposição 21.25.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais, $\boldsymbol{S} \subseteq \boldsymbol{V}'$ uma subvariedade e $F\colon V \to V'$ uma função diferenciável tal que $F \pitchfork S$. Então $F^{-1}(S) \subseteq V$ é uma subvariedade diferencial e* $\operatorname{codim}_V(F^{-1}(S)) = \operatorname{codim}_{V'}(S)$.

□ *Demonstração.* Consideremos $p \in U \subseteq V$, $q = F(p)$, $U' = F(U) \subseteq V'$, uma carta adaptada $\varphi\colon U' \subseteq V' \to I^{\dim S} \times I^{\operatorname{codim}_{V'} S}$ tal que $\varphi(q) = 0$ e $\varphi(U \cap S) = I^{\dim S} \times \{0\}$, e a projeção $\pi\colon I^{\dim S} \times I^{\operatorname{codim}_{V'} S} \to I^{\operatorname{codim}_{V'} S}$. Então

$$\pi \circ \varphi \circ F\colon V \to I^{\operatorname{codim}_{V'} S}$$

é uma função diferenciável. Notemos agora que $0 \in I^{\operatorname{codim}_{V'} S}$ é um valor regular de $\pi \circ \varphi \circ F$, pois para todo $a \in F^{-1}(S)$, como $F \pitchfork S$,

$$\mathrm{T}V'|_{F(a)} = \mathrm{D}F|_p \mathrm{T}V|_a + \mathrm{T}S|_{F(a)},$$

portanto $\mathrm{D}\varphi|_{F(a)} \mathrm{T}S|_{F(a)} = \mathbb{R}^{\dim S} \times \{0\}$. Sendo assim, segue que $F^{-1}(S) = (\pi \circ \varphi \circ F)^{-1}(0)$, logo $F^{-1}(S)$ é uma subvariedade de V. ■

Um caso importante é quando tem-se uma variedade diferencial $\boldsymbol{V}$ e duas subvariedades diferenciais $\boldsymbol{S}$ e $\boldsymbol{S}'$ de $\boldsymbol{V}$. Se consideramos a inclusão $\iota\colon S \to V$, a condição $p \in \iota^{-1}(S')$ é equivalente a $p \in S \cap S'$, pois $\iota(p) = p$. Nesse caso, $\mathrm{D}\iota|_p = \mathrm{I}$, portanto $\iota \pitchfork S'$ é equivalente a, para todo $p \in S \cap S'$,

$$\mathrm{T}S|_p + \mathrm{T}S'|_p = \mathrm{T}V|_p.$$

Claramente, essa condição é simétrica com relação a S e S'. Definimos assim a transversalidade entre subvariedades.

:⊢ **Definição 21.34.** Seja $\boldsymbol{V}$ uma variedade diferencial. Duas subvariedades diferenciais $\boldsymbol{S}$ e $\boldsymbol{S}'$ de $\boldsymbol{V}$ são *transversais* se, e somente se, para todo $p \in S \cap S'$,

$$\mathrm{T}S|_p + \mathrm{T}S'|_p = \mathrm{T}V|_p.$$

Denota-se $S \pitchfork S'$.

⊢ **Proposição 21.26.** *Sejam $\boldsymbol{V}$ uma variedade diferencial e $\boldsymbol{S}$ e $\boldsymbol{S}'$ subvariedades diferenciais de $\boldsymbol{V}$ tais que $S \pitchfork S'$. Então $S \cap S'$ é uma subvariedade diferencial de V e*

$$\operatorname{codim}(S \cap S') = \operatorname{codim}(S) + \operatorname{codim}(S').$$

□ *Demonstração.* Corolário da proposição anterior. ■

⊢ **Proposição 21.27.** *Sejam $\boldsymbol{V}, \boldsymbol{V}'$ e $\boldsymbol{W}$ variedades diferenciais, $\boldsymbol{S} \subseteq \boldsymbol{V}'$ uma subvariedade diferencial e $F\colon V \times W \to V''$, tal que $F \pitchfork S$. Então existe $\tilde{W} \subseteq W$ de medida total tal que $F_w \pitchfork S$ para todo $w \in \tilde{W}$ e*

$$\begin{aligned} F_w\colon V &\longrightarrow V' \\ p &\longmapsto F(p, w). \end{aligned}$$

► **Exemplo 21.1.** *Sejam $f\colon \mathbb{R}^m \to \mathbb{R}^n$ suave e $Z \subseteq \mathbb{R}^n$. Defina*

$$\begin{aligned} f\colon \mathbb{R}^m \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (x, v) &\longmapsto f(x) + v. \end{aligned}$$

$\mathrm{D}F_{(x_0,v_0)}(0, w) = \mathrm{D}(F_{x_0})_{v_0}(w) = w$. *Então $F \pitchfork Z$, e portanto $\tilde{S} \subseteq \mathbb{R}^n$ (como no teorema acima) tal que $F_{v_0}(x) = f(x) + v_0$ é $\pitchfork Z$.*

⊢ **Proposição 21.28.** *Sejam $S \subseteq N$ subvariedade fechada. O conjunto*

$$\{f\colon M \to N \mid f \pitchfork S\}$$

é aberto e denso.

21.6 Campos tensoriais

21.6.1 Campos tensoriais, vetoriais, e derivações

:⊢ **Definição 21.35.** Seja $\boldsymbol{V}$ uma variedade diferencial. Um *campo tensorial* de tipo (k, l) em $\boldsymbol{V}$ é uma função

$$\begin{aligned} T\colon V &\longrightarrow \mathrm{T}V^{\otimes(k,l)} \\ p &\longmapsto T|_p \end{aligned}$$

tal que, para todo $p \in V$, $T|_p \in \mathrm{T}V^{\otimes(k,l)}|_p$. O conjunto dos campos tensoriais diferenciáveis[2] de tipo (k, l) em $\boldsymbol{V}$ é denotado $\mathfrak{T}^{(k,l)}(\boldsymbol{V})$.

Um campo tensorial diferenciável de tipo $(0, k)$ é também denotado $\mathfrak{T}^k(\boldsymbol{V})$ e um de tipo $(k, 0)$ é também denotado $\mathfrak{T}^{*k}(\boldsymbol{V})$. Um *campo vetorial* é um campo tensorial de tipo $(0, 1)$ e um *campo covetorial* é um campo tensorial de tipo $(1, 0)$.

Como já definido anteriormente, um campo escalar é uma função $f \in \mathfrak{T}^0(\boldsymbol{V}) \simeq \mathscr{C}^\infty(\boldsymbol{V}, \mathbb{R})$. O conjunto $\mathfrak{T}^1(\boldsymbol{V})$ é um espaço linear sobre $\mathbb{R}$ com a soma e o produto por escalar induzidos pontualmente pela soma e produto por escalar de $\mathrm{T}V|_p$: a *adição* é dada por

$$\begin{aligned} +\colon \mathfrak{T}^1(\boldsymbol{V}) \times \mathfrak{T}^1(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^1(\boldsymbol{V}) \\ (v, v') &\longmapsto v + v'\colon V \longrightarrow \mathrm{T}V \\ &\qquad\qquad\qquad p \longmapsto v|_p + v'|_p, \end{aligned}$$

[2] Aqui diferenciável quer dizer $\mathscr{C}^\infty$.

a *inversa* da adição é dada por

$$\begin{aligned} -\colon \mathfrak{T}^1(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^1(\boldsymbol{V}) \\ v &\longmapsto -v\colon V \longrightarrow \mathrm{T}V \\ & \qquad\qquad p \longmapsto -v|_p, \end{aligned}$$

a *identidade* da adição é dada por

$$\begin{aligned} 0\colon V &\longrightarrow \mathrm{T}V \\ v &\longmapsto 0|_p \end{aligned}$$

e o produto por escalar é dado por

$$\begin{aligned} \cdot\colon \mathbb{R} \times \mathfrak{T}^1(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^1(\boldsymbol{V}) \\ (c, v) &\longmapsto cv\colon V \longrightarrow \mathrm{T}V \\ & \qquad\qquad p \longmapsto cv|_p. \end{aligned}$$

Mais adiante introduziremos um produto bilinear nesse espaço linear de modo a lhe dar uma estrutura de álgebra sobre $\mathbb{R}$. Além dessa operações, podemos multiplicar um campo vetorial por uma função escalar diferenciável pontualmente:

$$\begin{aligned} \bullet\colon \mathfrak{T}^0(\boldsymbol{V}) \times \mathfrak{T}^1(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^1(\boldsymbol{V}) \\ (f, v) &\longmapsto fv\colon V \longrightarrow \mathrm{T}V \\ & \qquad\qquad p \longmapsto f(p)v|_p. \end{aligned}$$

Isso é uma ação do anel $\mathfrak{T}^0(\boldsymbol{V})$ que dá ao espaço $\mathfrak{T}^1(\boldsymbol{V})$ uma estrutura de módulo.

$\vdash$ **Proposição 21.29.** *Seja $\boldsymbol{V}$ uma variedade diferencial.*

1. *A lista $(\mathfrak{T}^1(\boldsymbol{V}), +, -, 0, \cdot)$ é um espaço linear sobre $\mathbb{R}$.*
2. *A lista $(\mathfrak{T}^1(\boldsymbol{V}), +, -, 0, \bullet)$ é um módulo sobre o anel de $\mathfrak{T}^0(\boldsymbol{V})$.*

A proposição a seguir dá uma condição equivalente a um campo vetorial $v\colon V \longrightarrow \mathrm{T}V$ ser diferenciável, ou seja, pertencer a $\mathfrak{T}^1(\boldsymbol{V})$.

$\vdash$ **Proposição 21.30.** *Sejam $\boldsymbol{V}$ uma variedade diferencial e $v\colon V \longrightarrow \mathrm{T}V$ um campo vetorial. O campo vetorial v é diferenciável*[3] *em $\boldsymbol{V}$ se, e somente se, para toda $f \in \mathfrak{T}^0(\boldsymbol{V})$,*

$$\begin{aligned} \partial_v f\colon V &\longrightarrow \mathbb{R} \\ p &\longmapsto v|_p f \end{aligned}$$

é um campo escalar diferenciável.

[3]Aqui diferenciável quer dizer $\mathscr{C}^\infty$. Não consideraremos os detalhes relacionados ao grau de regularidade do campo, mas eles podem ser definidos e estudados de modo análogo, substituindo $\mathscr{C}^\infty$ por $\mathscr{C}^k$.

21.6.2 Derivações e colchete de campos vetoriais

A proposição da seção anterior é equivalente a dizer que o campo $v \in \mathfrak{T}^1(\boldsymbol{V})$ induz uma função

$$\begin{aligned} \partial_v \colon \mathfrak{T}^0(\boldsymbol{V}) &\longrightarrow \mathfrak{T}^0(\boldsymbol{V}) \\ f &\longmapsto \partial_v f. \end{aligned}$$

Essa função é linear na álgebra $\mathfrak{T}^0(\boldsymbol{V})$, pois

$$\partial_v(cf + f')(p) = v|_p(cf + f') = cv|_p f + v|_p f = (c\partial_v f + \partial_v f')(p).$$

Além disso, ela é uma derivação em $\mathfrak{T}^0(\boldsymbol{V})$, pois

$$\partial_v(ff')(p) = v|_p(ff') = v|_p(f)f'(p) + f(p)v|_p(f') = (\partial_v(f)f' + f\partial_v(f'))(p).$$

Pode-se mostrar que existe uma bijeção entre o conjunto dos campos vetoriais diferenciáveis $\mathfrak{T}^1(\boldsymbol{V})$ e o conjunto das derivações $\mathfrak{D}(\boldsymbol{\mathfrak{T}^0(V)})$. Por esse motivo, é comum se ignorar a notação $\partial_v f$ da derivação e denotar-se simplesmente vf. É importante notar, no entanto, que fv e vf são objetos distintos, o primeiro sendo um elemento de $\mathfrak{T}^1(\boldsymbol{V})$, resultado da multiplicação de uma função escalar por um campo vetorial, uma operação da estrutura de $\mathfrak{T}^0(\boldsymbol{V})$-módulo de $\mathfrak{T}^1(\boldsymbol{V})$, e a segunda é um elemento de $\mathfrak{T}^0(\boldsymbol{V})$, um campo escalar, resultado de uma derivação na álgebra $\boldsymbol{\mathfrak{T}^0(V)}$.

Definiremos agora o *colchete*[4] de campos vetoriais. Esse colchete é um produto bilinear que dará, como comentado anteriormente, uma estrutura de álgebra sobre $\mathbb{R}$ a $\mathfrak{T}^1(\boldsymbol{V})$. O conjunto $\mathfrak{D}(\mathfrak{T}^0(\boldsymbol{V}))$ tem um produto $\circ$ dado pela composição de funções, induzido do produto de composição de $\mathscr{L}(\boldsymbol{\mathfrak{T}^0(V)}, \mathbb{R})$. Esse produto é associativo, mas não faz de $\mathfrak{D}(\mathfrak{T}^0(\boldsymbol{V}))$ uma álgebra de derivação adjunta[5], pois não é fechado; ou seja, a composição de quaisquer duas derivações não é uma derivação. Podemos formar uma álgebra de derivação adjunta a partir de uma álgebra associativa de um modo bem conhecido usando o colchete. Definamos, portanto, para duas derivações $\partial_v, \partial_{v'} \in \mathfrak{D}(\mathfrak{T}^0(\boldsymbol{V}))$, o produto

$$[\partial_v, \partial_{v'}] := \partial_v \circ \partial_{v'} - \partial_{v'} \circ \partial_v.$$

Esse produto é uma derivação pela proposição 10.8. Portanto existe um campo vetorial associado à derivação $[\partial_v, \partial_{v'}]$, que denotaremos por $[v, v'] \in \mathfrak{T}^1(\boldsymbol{V})$, de modo que

$$\partial_{[v,v']} = [\partial_v, \partial_{v'}]\,.$$

[4] Comumente chamado de 'colchete de Lie' na literatura.

[5] Álgebra de Lie

O campo vetorial $[v, v']$ é chamado de *colchete* dos campos $v, v' \in \mathfrak{T}^1(\boldsymbol{V})$. Em um ponto $p \in V$, e função $f \in \mathfrak{T}^0(\boldsymbol{V})$, o colchete é

$$[v, v']\,|_p(f) = v|_p(v'f) - v'|_p(vf).$$

⊢ **Proposição 21.31** (Colchete em Coordenadas Locais)**.** *Sejam $\boldsymbol{V}$ uma variedade diferencial e $v, w \in \mathfrak{T}^1(\boldsymbol{V})$ campos vetoriais. Para toda carta (A, x) de $\boldsymbol{V}$, se $v = v^i \frac{\partial}{\partial x^i}$ e $w = w^i \frac{\partial}{\partial x^i}$ em A, então o colchete entre v e w em A é*

$$[v, w] = \mathop{+}_{(i,j)\in[d]^2} \left(v^i \frac{\partial w^j}{\partial x^i} - w^i \frac{\partial v^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}.$$

□ *Demonstração.* Basta notar que, para toda $f \in \mathfrak{T}^0(\boldsymbol{V})$

$$\begin{aligned}
[v, w]\, f &= \mathop{+}_{i\in[d]} v^i \frac{\partial}{\partial x^i} \left(\mathop{+}_{j\in[d]} w^j \frac{\partial f}{\partial x^j} \right) - \mathop{+}_{j\in[d]} w^j \frac{\partial}{\partial x^j} \left(\mathop{+}_{i\in[d]} v^i \frac{\partial f}{\partial x^i} \right) \\
&= \mathop{+}_{(i,j)\in[d]^2} \left(v^i \frac{\partial w^j}{\partial x^i} \frac{\partial f}{\partial x^j} + v^i w^j \frac{\partial^2 f}{\partial x^i \partial x^j} - w^j \frac{\partial v^i}{\partial x^j} \frac{\partial f}{\partial x^i} - w^j v^i \frac{\partial^2 f}{\partial x^j \partial x^i} \right) \\
&= \mathop{+}_{(i,j)\in[d]^2} \left(v^i \frac{\partial w^j}{\partial x^i} - w^i \frac{\partial v^j}{\partial x^i} \right) \frac{\partial f}{\partial x^j},
\end{aligned}$$

pois $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$. ■

Usando a notação de Einstein e simplificando a fórmula da proposição, temos

$$[v, w] = (vw^i - wv^i) \frac{\partial}{\partial x^i}.$$

Em particular, temos que $\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0$ para todos $i, j \in [d]$, fato essencial na demonstração da proposição anterior.

O colchete satisfaz as seguintes propriedades.

⊢ **Proposição 21.32.** *Seja $\boldsymbol{V}$ uma variedade diferencial.*

1. *A lista*
$$\boldsymbol{\mathfrak{T}^1(V)} := (\mathfrak{T}^1(\boldsymbol{V}), +, -, 0, [\cdot, \cdot], \cdot)$$
é uma álgebra de derivação adjunta sobre $\mathbb{R}$;
2. *Para todos $v, v' \in \mathfrak{T}^1(\boldsymbol{V})$ e $f, f' \in \mathfrak{T}^0(\boldsymbol{V})$,*
$$[fv, f'v'] = ff'\,[v, v'] + f(vf')v' - f'(v'f)v.$$

□ *Demonstração.* 1. Corolário de 10.8.

2. Para toda $g \in \mathfrak{T}^0(\boldsymbol{V})$,

$$\begin{aligned}
[fv, f'v']\, g &= (fv)(f'v')g - (f'v')(fv)g \\
&= (fv)(f'v'g) - (f'v')(fvg) \\
&= (fv)(f')(v'g) + f'((fv)(v'g)) \\
&\qquad - (f'v')(f)(vg) - f((f'v')(vg)) \\
&= ff'\,[v, v']\,(g) + f(vf')v'(g) - f'(v'f)v(g) \\
&= (ff'\,[v, v'] + f(vf')v' - f'(v'f)v)\,(g).
\end{aligned}$$

■

21.6.3 Álgebra de campos tensoriais

Assim como definimos produto tensorial de espaços lineares, podemos definir um produto tensorial de módulos sobre anéis (comutativos). Pode-se mostrar que, nesse caso, o produto tensorial de tipo (k, l) do módulo $\mathfrak{T}^1(\boldsymbol{V})$ sobre o anel comutativo $\mathfrak{T}^0(\boldsymbol{V})$ é o espaço $\mathfrak{T}^{(k,l)}(\boldsymbol{V})$ de campos tensoriais de tipo (k, l) em $\boldsymbol{V}$:

$$\left(\mathfrak{T}^1(\boldsymbol{V})\right)^{\otimes(k,l)} \simeq \mathfrak{T}^{(k,l)}(\boldsymbol{V}).$$

Desse modo, os campos tensoriais $T \in \mathfrak{T}^{(k,l)}(\boldsymbol{V})$ são tensores em si, e não somente tensores em cada ponto $p \in V$. Assim, podemos formar a *álgebra de campos tensoriais* em $\boldsymbol{V}$

$$\mathfrak{T}(\boldsymbol{V}) := \bigoplus_{(k,l)\in\mathbb{N}^2} \mathfrak{T}^{(k,l)}(\boldsymbol{V}).$$

O espaço $\mathfrak{T}(\boldsymbol{V})$ é uma álgebra sobre $\mathfrak{T}^0(\boldsymbol{V})$. A soma e o produto por escalar são definidos pontualmente e o produto tensorial também.

21.6.4 Fluxo de campos vetoriais

As demonstrações dessa serão omitidas, mas podem ser encontradas em *Introduction to Smooth Manifolds* [Lee12].

Consideremos uma curva diferenciável $\gamma\colon I \to V$ definida em um intervalo aberto I. Como, para todo $t \in I$, $\mathrm{T}I|_t \simeq \mathbb{R}$, segue que sua diferencial em $t \in I$ é uma função linear $\mathrm{D}\gamma|_t\colon \mathrm{T}\mathbb{R}|_{t_0} \longrightarrow \mathrm{T}V|_{\gamma(t)}$, o que significa que $\mathrm{D}\gamma|_t$ pode ser identificada com um vetor $\dot{\gamma}(t) \in \mathrm{T}V|_{\gamma(t)}$. O vetor velocidade de tal curva no ponto t é dado por

$$\dot{\gamma}(t_0) := \mathrm{D}\gamma\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0}\right) \in \mathrm{T}|_{\gamma(t_0)}V,$$

em que $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0}$ é a base padrão em $\mathrm{T}\mathbb{R}|_{t_0} \simeq \mathbb{R}$.

:⊢ **Definição 21.36.** Sejam $\boldsymbol{V}$ um variedade diferencial e $v \in \mathfrak{T}^1(\boldsymbol{V})$. Uma *curva integral de v* é uma curva diferenciável $\gamma\colon I \to V$ tal que, para todo $t \in I$,

$$\dot{\gamma}(t) = v|_{\gamma(t)}.$$

Se $0 \in I$, $\gamma(0) \in M$ é o *ponto inicial* de γ e γ é *iniciada* em $\gamma(0)$.

Em coordenadas locais (A, x), a condição $\dot{\gamma}(t) = v|_{\gamma(t)}$ se torna

$$\dot{\gamma}^i(t) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)} = v^i(\gamma(t)) \left.\frac{\partial}{\partial x^i}\right|_{\gamma(t)},$$

que se reduz a um sistema autônomo de equações diferenciais parciais: para todo $i \in [d]$,

$$\dot{\gamma}^i(t) = v^i(\gamma^0(t), \ldots, \gamma^{d-1}(t)).$$

O teorema de existência, unicidade e diferenciabilidade garante que o sistema anterior tem solução única e diferenciável. Isso implica a seguinte proposição.

⊢ **Proposição 21.33.** *Sejam $\boldsymbol{V}$ um variedade diferencial e $v \in \mathfrak{T}^1(\boldsymbol{V})$. Para todo $p \in V$, existem $\varepsilon \in \,]0, \infty[$ e curva diferenciável $\gamma\colon\,]-\varepsilon, \varepsilon[\to V$ que é curva integral de v iniciada em p.*

⊢ **Proposição 21.34.** *Sejam $\boldsymbol{V}$ um variedade diferencial, $v \in \mathfrak{T}^1(\boldsymbol{V})$ e $\gamma\colon I \to V$ uma curva integral de v.*

1. *(Translação) Para todo $c \in \mathbb{R}$, a curva*

$$\begin{aligned} \gamma \circ (c+)\colon -c + I &\longrightarrow V \\ t &\longmapsto \gamma(c+t) \end{aligned}$$

   é uma curva integral de v;
2. *(Reescala) Para todo $c \in \mathbb{R} \smallsetminus \{0\}$, a curva*

$$\begin{aligned} \gamma \circ (c\times)\colon c^{-1}I &\longrightarrow M \\ t &\longmapsto \gamma(ct) \end{aligned}$$

   é uma curva integral de cv.

Lembremos que um fluxo em um conjunto V é uma ação de $\mathbb{R}$ em v, uma função

$$\begin{aligned} \Phi\colon \mathbb{R} \times V &\longrightarrow V \\ (t, p) &\longmapsto \Phi^t(p) \end{aligned}$$

tal que

1. Para todo $p \in V$, $\Phi^0(p) = p$;
2. Para todos $p \in V$ e $t, t' \in \mathbb{R}$, $\Phi^{t+t'}(p) = \Phi^t(\Phi^{t'}(p))$.

Quando V tem topologia, podemos considerar um fluxo contínuo, e quando V é uma variedade diferencial V, podemos considerar fluxos diferenciáveis. Para todo $t \in \mathbb{R}$, a translação por t ao longo de Φ é

$$\begin{aligned} \Phi^t \colon V &\longrightarrow V \\ p &\longmapsto \Phi^t(p) \end{aligned}$$

e, para todo $p \in V$, a curva de fluxo de Φ iniciada em p é

$$\begin{aligned} \Phi_{(p)} \colon \mathbb{R} &\longrightarrow V \\ t &\longmapsto \Phi^t(p). \end{aligned}$$

Note que $\Phi^t \colon V \to V$ é um difeomorfismo e $\Phi_{(p)} \colon \mathbb{R} \to V$ é uma curva diferenciável.

:⊢ **Definição 21.37.** Sejam $\boldsymbol{V}$ um variedade diferencial e $\Phi \colon \mathbb{R} \times V \to V$ um fluxo diferenciável em $\boldsymbol{V}$. O *gerador infinitesimal de* Φ é o campo vetorial

$$\begin{aligned} v \colon V &\longrightarrow \mathrm{T}V \\ p &\longmapsto \dot{\Phi}_{(p)}(0). \end{aligned}$$

⊢ **Proposição 21.35.** *Sejam $\boldsymbol{V}$ um variedade diferencial e $\Phi \colon \mathbb{R} \times V \to V$ um fluxo diferenciável em $\boldsymbol{V}$. O gerador infinitesimal v de Φ é um campo vetorial diferenciável e cada curva $\Phi_{(p)} \colon \mathbb{R} \to V$ é uma curva integral de v.*

No entanto, nem todo campo vetorial diferenciável gera um fluxo porque nem toda curva integral está definida para todo $\mathbb{R}$. Para permitir que exista uma forma de recíproca para essa proposição, devemos expandir o conceito de fluxo com o conceito de um fluxo local.

:⊢ **Definição 21.38.** Seja $\boldsymbol{V}$ um variedade diferencial. Um *domínio de fluxo local* em $\boldsymbol{V}$ é um conjunto aberto $\mathscr{D} \subseteq \mathbb{R} \times V$ que satisfaz: para todo $p \in V$, existem $e_p \in [-\infty, 0[$ e $\bar{e}_p \in]0, \infty]$ tais que

$$]e_p, \bar{e}_p[= \{t \in \mathbb{R} \mid (t, p) \in \mathscr{D}\}.$$

Em particular, $0 \in]e_p, \bar{e}_p[$. Podemos escrever

$$\mathscr{D} = \bigcup_{p \in V}]e_p, \bar{e}_p[\times \{p\}.$$

:⊢ **Definição 21.39.** Seja $\boldsymbol{V}$ um variedade diferencial. Um *fluxo local* em $\boldsymbol{V}$ é uma função $\Phi\colon \mathscr{D} \to V$, em que $\mathscr{D}$ um domínio de fluxo local em $\boldsymbol{V}$, e valem

1. Para todo $p \in V$,
$$\Phi^0(p) = p;$$
2. Para todos $p \in V$, $t \in \,]e_p, \bar{e}_p[$ e $t' \in \,]e_{\Phi^t(p)}, \bar{e}_{\Phi^t(p)}[$ tal que $t + t' \in \,]e_p, \bar{e}_p[$,
$$\Phi^{t+t'}(p) = \Phi^t(\Phi^{t'}(p)).$$

Para todo $p \in V$, a *curva de fluxo* de Φ iniciada em p é a curva

$$\begin{aligned} \Phi_{(p)}\colon]e_p, \bar{e}_p[&\longrightarrow V \\ t &\longmapsto \Phi^t(p). \end{aligned}$$

Para todo $t \in \mathbb{R}$, a *translação* por t ao longo de Φ é a função

$$\begin{aligned} \Phi^t\colon V_t &\longrightarrow V \\ p &\longmapsto \Phi^t(p), \end{aligned}$$

em que $V_t := \{p \in V \mid (t, p) \in \mathscr{D}\}$.

Temos as relações

$$p \in V_t \quad \Longleftrightarrow \quad t \in \,]e_p, \bar{e}_p[\quad \Longleftrightarrow \quad (t, p) \in \mathscr{D}.$$

Para ressaltar a diferença conceitual, às vezes chama-se um fluxo de *fluxo global*. Claro que, se Φ é um fluxo é um fluxo local em que, para todos $p \in V$ e $t \in \mathbb{R}$, $]e_p, \bar{e}_p[= \mathbb{R}$ e $V_t = V$. Assim como no caso do fluxo, um fluxo local pode ser contínuo ou diferenciável.

:⊢ **Definição 21.40.** Sejam $\boldsymbol{V}$ um variedade diferencial e $\Phi\colon \mathscr{D} \to V$ um fluxo local diferenciável em $\boldsymbol{V}$. O *gerador infinitesimal de* Φ é o campo vetorial

$$\begin{aligned} v\colon V &\longrightarrow \mathrm{T}V \\ p &\longmapsto \dot{\Phi}_{(p)}(0). \end{aligned}$$

⊢ **Proposição 21.36.** *Sejam $\boldsymbol{V}$ um variedade diferencial e $\Phi\colon \mathscr{D} \to V$ um fluxo local diferenciável em V. O gerador infinitesimal v de Φ é um campo vetorial diferenciável e cada curva $\Phi_{(p)}\colon \mathbb{R} \to V$ é uma curva integral de v.*

:⊢ **Definição 21.41.** Seja $\boldsymbol{V}$ uma variedade diferencial.

1. Um *fluxo local máximo* em $\boldsymbol{V}$ é um fluxo local em $\boldsymbol{V}$ que não admite extensão para um domínio de fluxo local maior.

2. Uma *curva integral máxima* é uma curva integral de um campo vetorial que não admite extensão para um intervalo maior.

$\vdash$ **Proposição 21.37.** *Sejam $\boldsymbol{V}$ um variedade diferencial e $v \in \mathfrak{T}^1(\boldsymbol{V})$. Existe único fluxo diferencial máximo $\Phi\colon \mathscr{D} \to V$ gerado por v e o fluxo Φ satisfaz*

1. *Para todo $p \in V$, a curva de fluxo $\Phi_{(p)}\colon\ ]e_p, \bar{e}_p[\to V$ é uma curva integral máxima de v iniciada em p;*
2. *Para todo $p \in V$ e todo $t \in\ ]e_p, \bar{e}_p[$,*

$$]e_{\Phi^t(p)}, \bar{e}_{\Phi^t(p)}[\ = \ ]e_p, \bar{e}_p[\ - t = \{t' - t \mid t' \in\ ]e_p, \bar{e}_p[\}\,;$$

3. *Para todo $t \in \mathbb{R}$, V_t é aberto em V e $\Phi^t\colon V_t \to V_{-t}$ é um difeomorfismo com inversa Φ^{-t}.*

Esse fluxo gerado por $v \in \mathfrak{T}^1(\boldsymbol{V})$ será denotado

$$\begin{aligned} \Phi^{(\cdot)v}(\bullet)\colon \mathscr{D} &\longrightarrow V \\ (t,p) &\longmapsto \Phi^{tv}(p), \end{aligned}$$

de modo que valem as seguintes propriedades. Pela propriedade de curvas integrais maximais, para todo $p \in V$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^{tv}(p) = v \circ \Phi^{tv}(p) = v|_{\Phi^{tv}(p)}.$$

Pela propriedades de fluxo local,

$$\Phi^{0v}(p) = \mathrm{I}(p) = p$$

e

$$\Phi^{(t+t')v}(p) = \Phi^{tv} \circ \Phi^{t'v}(p)$$

Além disso, pelo lema de reescala, temos que, para todo $t' \in \mathbb{R}$, o fluxo de do campo vetorial $t'v$ é dado por

$$\Phi^{t(t'v)}(p) = \Phi^{(tt')v}(p),$$

o que significa que a notação $\Phi^{tt'v}$ não é ambígua. Essas similaridades de propriedades com a exponencial justificam a notação. Além da similaridade, a função exponencial de grupos diferenciais é de fato definida a partir do fluxo de um campo vetorial invariante, e como um caso particular, a função exponencial em álgebras normadas está também relacionada ao fluxo acima definido.

21.6.5 Espaço cotangente

Dada uma variedade diferencial $\boldsymbol{V}$ e seu espaço

:⊢ **Definição 21.42.** Sejam $\boldsymbol{V}$ uma variedade diferencial e $p \in V$. O *espaço cotangente* a $\boldsymbol{V}$ em p é

$$\mathrm{T}^*V|_p := (\mathrm{T}V|_p)^*,$$

o espaço linear dual do espaço tangente a $\boldsymbol{V}$ em p. O *fibrado cotangente* de $\boldsymbol{V}$ é

$$\mathrm{T}^*V := \bigsqcup_{p\in V} \mathrm{T}^*V|_p,$$

a união disjunta dos espaço tangentes a $\boldsymbol{V}$ em todos pontos de V.

21.6.5.1 Referencial local coordenado

Dada uma carta (A, x) de V, as funções $(\mathrm{d}x^i|_p)_{i\in[d]}$, definidas por

$$\begin{aligned} \mathrm{d}x^i|_p \colon \mathrm{T}V|_p &\longrightarrow \mathbb{R} \\ v &\longmapsto v(x^i) \end{aligned}$$

são uma base ordenada de $\mathrm{T}^*V|_p$. Essas funções são lineares porque, para todos $v, v' \in \mathrm{T}V|_p$ e $c \in \mathbb{R}$,

$$\begin{aligned} \mathrm{d}x^i|_p(cv + v') &= (cv + v')(x^i) \\ &= cv(x^i) + v'(x^i) \\ &= c\mathrm{d}x^i|_p(v) + \mathrm{d}x^i|_p(v'). \end{aligned}$$

A base $(\mathrm{d}x^i|_p)_{i\in[d]}$ de $\mathrm{T}^*V|_p$ é a base dual de $\left(\frac{\partial}{\partial x^i}\Big|_p\right)_{i\in[d]}$, pois

$$\mathrm{d}x^i|_p\left(\frac{\partial}{\partial x^j}\bigg|_p\right) = \frac{\partial}{\partial x^j}\bigg|_p x^i = \partial_j(x^i \circ x^{-1})|_{x(p)} = \partial_j(\pi^i)|_{x(p)} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

21.6.6 Formas diferenciáveis

O fibrado de k-covetores (k-tensores covariantes) de uma variedade diferencial $\boldsymbol{V}$ é $\mathrm{T}^*V^{\otimes k}$, e o subfibrado de tensores alternados desse fibrado é denotado

$$\bigwedge^k \mathrm{T}^*V := \bigsqcup_{p\in V} \bigwedge^k \mathrm{T}^*V|_p.$$

Uma seção desse fibrado é um campo de funcionais k-lineares alternados sobre $\boldsymbol{V}$. A próxima definição consiste desses objetos.

:⊢ **Definição 21.43.** Seja $\boldsymbol{V}$ uma variedade diferencial. Uma *k-forma* em $\boldsymbol{V}$ é uma função

$$\omega\colon V \to \bigwedge^k \mathrm{T}^*V$$

tal que, para todo $p \in V$, $\omega|_p \in \bigwedge^k \mathrm{T}^*V|_p$ é um funcional k-linear alternado de $\mathrm{T}V|_p$. O conjunto de k-formas diferenciáveis[6] em $\boldsymbol{V}$ é denotado $\mathfrak{A}^k(\boldsymbol{V})$.

O produto exterior de duas formas $\omega \in \mathfrak{A}^k(\boldsymbol{V})$ e $\omega' \in \mathfrak{A}^{k'}(\boldsymbol{V})$ é definido pontualmente,

$$(\omega \wedge \omega')|_p := \omega|_p \wedge \omega'|_p,$$

e é uma $(k+k')$-forma em $\boldsymbol{V}$. Para $k = 0$, $\omega \wedge \omega'$ é simplesmente $\omega\omega'$, uma k'-forma.

A soma direta desses espaços é o espaço das formas em $\boldsymbol{V}$, denotado

$$\mathfrak{A}(\boldsymbol{V}) := \bigoplus_{i\in[d]} \mathfrak{A}^k(\boldsymbol{V}).$$

Esse espaço linear com o produto exterior

$$\begin{aligned} \wedge\colon \mathfrak{A}(\boldsymbol{V}) \times \mathfrak{A}(\boldsymbol{V}) &\longrightarrow \mathfrak{A}(\boldsymbol{V}) \\ (\omega, \omega') &\longmapsto \omega \wedge \omega' \end{aligned}$$

é uma álgebra associativa, anticomutativa e graduada. Ressaltamos isso na proposição seguinte, que não demonstraremos.

⊢ **Proposição 21.38.** *Seja $\boldsymbol{V}$ um variedade diferencial. A lista*

$$\boldsymbol{\mathfrak{A}(V)} := (\mathfrak{A}(\boldsymbol{V}), +, -, 0, \wedge, \cdot)$$

é uma álgebra associativa, anticomutativa e graduada sobre $\mathbb{R}$.

21.6.6.1 Representação coordenada

Seja $\omega \in \mathfrak{A}^k(V)$. Como $\omega|_p$ é um k-covetor de $\mathrm{T}V|_p$ para cada $p \in V$, podemos escrevê-lo com respeito a uma base de $\mathrm{T}^*V|_p$. Uma carta (A, x) de $\boldsymbol{V}$ induz uma base $(\mathrm{d}x^i|_p)_{i\in[d]}$ de $\mathrm{T}^*V|_p$ para cada $p \in A$. Em coordenadas locais, portanto, o funcional k-linear alternado $\omega|_p$ pode ser escrito

$$\omega|_p = \mathop{+}_{i\in[d]^{\uparrow k}} \omega_i(p)\mathrm{d}x^i|_p$$

em que $[d]^{\uparrow k} := \left\{(i_0, \ldots, i_{k-1})_{j\in[k]} \in [d]^k \;\middle|\; i_0 < \cdots < i_{k-1}\right\}$ é um conjunto de multi-índices crescentes e, para cada multi-índice $i = (i_0, \ldots, i_{k-1}) \in [d]^{\uparrow k}$,

$$\mathrm{d}x^i|_p = \mathrm{d}x^{i_0}|_p \wedge \cdots \wedge \mathrm{d}x^{i_{k-1}}|_p.$$

[6] Aqui diferenciável quer dizer $\mathscr{C}^\infty$.

Isso ocorre porque $(\mathrm{d}x^i|_p)_{i\in[d]^{\uparrow k}}$ é uma base de $\bigwedge^k \mathrm{T}^*V|_p$. A k-forma ω restrita a A pode ser escrita

$$\omega|_A = \mathop{+}_{i\in[d]^{\uparrow k}} \omega_i \mathrm{d}x^i.$$

⊢ **Proposição 21.39.** *Seja $\boldsymbol{V}$ uma variedade diferencial. Uma k-forma $\omega \in \mathfrak{A}^k(V)$ é diferenciável se, e somente se, para toda carta (A, x) de $\boldsymbol{V}$, as funções $\omega_i \colon A \to \mathbb{R}$ tais que*

$$\omega|_A = \mathop{+}_{i\in[d]^{\uparrow k}} \omega_i \mathrm{d}x^i$$

são diferenciáveis.

Diferenciabilidade independe da carta no sentido de que, se algum sub-atlas da variedade faz com que as funções a_i sejam todas diferenciáveis, então todas cartas compatíveis com esse atlas também farão.

21.6.6.2 Formas volume

:⊢ **Definição 21.44.** Seja $\boldsymbol{V}$ uma variedade diferencial d-dimensional. Uma *forma volume* em $\boldsymbol{V}$ é uma forma $\omega \in \mathfrak{A}^d(V)$ tal que, para todo $p \in V$, $\omega|_p \neq 0$.

⊢ **Proposição 21.40.** *Seja $\boldsymbol{V}$ uma variedade diferenciável. Então $\boldsymbol{V}$ é orientável se, e somente se, existe uma forma volume $\omega \in \mathfrak{A}^d(V)$.*

21.6.6.3 Formas puxadas

:⊢ **Definição 21.45.** Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais, $F \colon V \to V'$ uma função diferencial e $\omega \in \mathfrak{A}^k(V')$ uma k-forma em $\boldsymbol{V}'$. A k-forma *puxada* de ω por F é a a k-forma

$$\begin{aligned} F^\star\omega \colon V &\longrightarrow \bigwedge^k \mathrm{T}^*V \\ p &\longmapsto F^\star\omega|_p \colon \mathrm{T}V|_p \times \cdots \mathrm{T}V|_p \longrightarrow \mathbb{R} \\ &\qquad\qquad (v_0, \ldots, v_{k-1}) \longmapsto \omega|_{F(p)}(\mathrm{D}F|_p v_0, \ldots, \mathrm{D}F|_p v_{k-1}). \end{aligned}$$

em $\boldsymbol{V}$.

A definição é equivalente a, para todo $p \in V$,

$$(F^\star\omega)|_p := (\mathrm{D}F|_p)^\star \omega|_{F(p)}.$$

⊢ **Proposição 21.41.** *Sejam $\boldsymbol{V}$ e $\boldsymbol{V}'$ variedades diferenciais e $F \colon V \to V'$ uma função diferencial.*

1. *$F^\star \colon \mathfrak{A}^k(V') \to \mathfrak{A}^k(V')$ é uma função linear;*

2. *Para todas formas* $\omega \in \mathfrak{A}^k(V')$ *e* $\omega' \in \mathfrak{A}^{k'}(V')$,

$$F^\star(\omega \wedge \omega') = (F^\star\omega) \wedge (F^\star\omega');$$

3. *Para toda carta* (A, x) *de* $\boldsymbol{V}'$ *e toda forma* $\omega \in \mathfrak{A}^k(V')$,

$$F^\star\left(\mathop{+}_{I\in[d]^{\uparrow k}} \omega_I \mathrm{d}x^I\right) = \mathop{+}_{I\in[d]^{\uparrow k}} (\omega_I \circ F)\mathrm{d}(x \circ F)^I = \mathop{+}_{I\in[d]^{\uparrow k}} (F^\star\omega_I)\mathrm{d}(F^\star x)^I.$$

21.6.7 Derivada exterior

21.6.7.1 Formas no espaço real

Definiremos nesta seção a derivada exterior de formas. Primeiro definiremos a derivada exterior de formas em $\mathbb{R}^d$. Consideremos um campo escalar $\omega\colon A \subseteq \mathbb{R}^d \to \mathbb{R}$, que é uma 0-forma em $A \subseteq \mathbb{R}^d$. A diferencial de ω em um ponto $p \in A$ é uma função linear $\mathrm{D}\omega|_p\colon \mathbb{R}^d \to \mathbb{R}$ (que é um funcional 1-linear alternado em $\mathbb{R}^d$) dada por

$$\mathrm{D}\omega|_p = \mathop{+}_{i\in[d]} \partial_i\omega(p)\mathrm{d}\pi^i|_p,$$

em que $\mathrm{d}\pi^i|_p := \mathrm{D}\pi^i|_p$ é a diferencial de $\pi^i\colon A \subseteq \mathbb{R}^d \to \mathbb{R}$ em p, que é a funcional linear $e^i\colon \mathbb{R}^d \to \mathbb{R}$. (O funcional e^i é o mesmo funcional que π^i — as notações diferem somente por causa do contexto ser diferente — e $\mathrm{d}\pi^i|_p = \pi^i$, pois π^i é linear.) Isso significa que

$$\mathrm{D}\omega = \mathop{+}_{i\in[d]} \partial_i\omega\mathrm{d}\pi^i$$

é uma 1-forma em $A \subseteq \mathbb{R}^d$, ou seja, um campo de funcionais lineares alterados. Além disso, lembremos que, para todo multi-índice $I = (i_0, \ldots, i_{k-1}) \in [d]^{\uparrow k}$, define-se

$$\mathrm{d}\pi^I = \mathrm{d}\pi^{i_0} \wedge \cdots \wedge \mathrm{d}\pi^{i_{k-1}}.$$

Com isso em mente, definimos a derivada exterior de formas em $\mathbb{R}^d$.

:⊢ **Definição 21.46.** Sejam $d \in \mathbb{N}$, $A \subseteq \mathbb{R}^d$ um aberto e $\omega \in \mathfrak{A}^k(A)$ uma k-forma em A tal que

$$\omega := \mathop{+}_{I\in[d]^{\uparrow k}} \omega_I\mathrm{d}\pi^I,$$

sendo, para todo $I \in [d]^{\uparrow k}$, $\omega_I \in \mathscr{C}^\infty(A)$. A *derivada exterior* de ω é a $(k+1)$-forma

$$\mathrm{d}\omega := \mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}\omega_I \wedge \mathrm{d}\pi^I,$$

em que, para todo $I \in [d]^{\uparrow k}$, $\mathrm{d}\omega_I := \mathrm{D}\omega_I$ é a diferencial de ω_I.

► **Exemplo 21.2.** *Como já comentado, a derivada exterior de uma* 0*-forma* $\omega\colon A \subseteq \mathbb{R}^d \to \mathbb{R}$ *é dada por*

$$\mathrm{d}\omega = \mathop{\bigplus}_{i\in[d]} \partial_i \omega \mathrm{d}\pi^i.$$

Note também que a derivada exterior de uma 1*-forma* $\omega = \bigplus_{i\in[d]} \omega_i \mathrm{d}\pi^i$ *é dada por*

$$\begin{aligned}
\mathrm{d}\omega &= \mathop{\bigplus}_{i\in[d]} \mathrm{d}\omega_i \wedge \mathrm{d}\pi^i \\
&= \mathop{\bigplus}_{i\in[d]} \left(\mathop{\bigplus}_{j\in[d]} \partial_j \omega_i(p) \mathrm{d}\pi^j \right) \wedge \mathrm{d}\pi^i \\
&= \mathop{\bigplus}_{(i,j)\in[d]^2} \partial_j \omega_i(p) \mathrm{d}\pi^j \wedge \mathrm{d}\pi^i \\
&= \mathop{\bigplus}_{i<j} \partial_j \omega_i(p) \mathrm{d}\pi^j \wedge \mathrm{d}\pi^i + \mathop{\bigplus}_{i>j} \partial_j \omega_i(p) \mathrm{d}\pi^j \wedge \mathrm{d}\pi^i \\
&= \mathop{\bigplus}_{i<j} (\partial_i \omega_j - \partial_j \omega_i) \mathrm{d}\pi^i \wedge \mathrm{d}\pi^j.
\end{aligned}$$

Notemos que a derivada exterior acima definida para cada $k \in [d]$ é uma função

$$\begin{aligned}
\mathrm{d}\colon \mathfrak{A}^k(A) &\longrightarrow \mathfrak{A}^{k+1}(A) \\
\omega &\longmapsto \mathrm{d}\omega
\end{aligned}$$

e, portanto, é uma função

$$\begin{aligned}
\mathrm{d}\colon \mathfrak{A}(A) &\longrightarrow \mathfrak{A}(A) \\
\omega &\longmapsto \mathrm{d}\omega.
\end{aligned}$$

ao definirmos que, para todas $\omega \in \mathfrak{A}^k(A)$ e $\omega' \in \mathfrak{A}^{k'}(A)$,

$$\mathrm{d}(\omega + \omega') := \mathrm{d}\omega + \mathrm{d}\omega'.$$

A derivada exterior satisfaz as seguintes propriedades.

⊢ **Proposição 21.42.** *Sejam* $d \in \mathbb{N}$ *e* $A \subseteq \mathbb{R}^d$ *um aberto. A derivada exterior* $\mathrm{d}\colon \mathfrak{A}(A) \to \mathfrak{A}(A)$ *satisfaz*

1. d *é linear sobre* $\mathbb{R}$*;*
2. *Para todas formas* $\omega \in \mathfrak{A}^k(A)$ *e* $\omega' \in \mathfrak{A}^{k'}(A)$*,*

$$\mathrm{d}(\omega \wedge \omega') = \mathrm{d}\omega \wedge \omega' + (-1)^k \omega \wedge \mathrm{d}\omega';$$

3. $\mathrm{d} \circ \mathrm{d} = 0$*;*

4. *Para toda função diferenciável* $F\colon A \subseteq \mathbb{R}^d \to A' \subseteq \mathbb{R}^{d'}$ *e toda forma* $\omega \in \mathfrak{A}^k(A')$,
$$F^\star(\mathrm{d}\omega) = \mathrm{d}(F^\star\omega).$$

□ *Demonstração.* 1. Sejam $\omega, \omega' \in \mathfrak{A}^k(A)$ e $c \in \mathbb{R}$. Então
$$c\omega + \omega' = \mathop{+}_{I\in[d]^{\uparrow k}} (c\omega_I + \omega'_I) \wedge \mathrm{d}\pi^I,$$
portanto
$$\begin{aligned}
\mathrm{d}(c\omega + \omega') &= \mathrm{d}\left(\mathop{+}_{I\in[d]^{\uparrow k}} (c\omega_I + \omega'_I) \wedge \mathrm{d}\pi^I\right) \\
&= \mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}(c\omega_I + \omega'_I) \wedge \mathrm{d}\pi^I \\
&= \mathop{+}_{I\in[d]^{\uparrow k}} (c\mathrm{d}\omega_I + \mathrm{d}\omega'_I) \wedge \mathrm{d}\pi^I \\
&= c \mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}\omega_I \wedge \mathrm{d}\pi^I + \mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}\omega'_I \wedge \mathrm{d}\pi^I \\
&= c\mathrm{d}\omega + \mathrm{d}\omega'.
\end{aligned}$$
Para $\omega \in \mathfrak{A}^k(A)$ e $\omega' \in \mathfrak{A}^{k'}(A)$, segue da definição de d em $\mathfrak{A}(A)$.

2. Basta considerar as formas $f\mathrm{d}\pi^i \in \mathfrak{A}^k(A)$ e $f'\mathrm{d}\pi^{I'} \in \mathfrak{A}^{k'}(A)$, em que $f, f' \in \mathscr{C}^\infty(A)$ são campos escalares, pois por linearidade a propriedade valerá para todas formas. Temos que mostrar que $\mathrm{d}(f\mathrm{d}\pi^I) = \mathrm{d}f \wedge \mathrm{d}\pi^I$ para todo multi-índice $I \in [d]^k$, não somente os crescentes $I \in [d]^{\uparrow k}$. Se I tem alguma entrada repetida, $\mathrm{d}\pi^I = 0$, portanto $\mathrm{d}(f\mathrm{d}\pi^I) = 0 = \mathrm{d}f \wedge \mathrm{d}\pi^I$. Caso contrário, existe permutação $\sigma \in [k]$ tal que $\sigma(I) = (i_{\sigma(0)}, \ldots, i_{\sigma(k-1)})$ é um multi-índice crescente, portanto
$$\mathrm{d}(f\mathrm{d}\pi^I) = \mathrm{d}(\pm(\sigma)f\mathrm{d}\pi^{\sigma(I)}) = \pm(\sigma)\mathrm{d}f \wedge \mathrm{d}\pi^{\sigma(I)} = \mathrm{d}f \wedge \mathrm{d}\pi^I.$$
Assim, segue que
$$\begin{aligned}
\mathrm{d}((f\mathrm{d}\pi^I) \wedge (f'\mathrm{d}\pi^{I'})) &= \mathrm{d}(ff'\mathrm{d}\pi^I \wedge \mathrm{d}\pi^{I'}) \\
&= \mathrm{d}(ff'\mathrm{d}\pi^I \wedge \mathrm{d}\pi^{I'}) \\
&= ((\mathrm{d}f)f' + f\mathrm{d}f') \wedge \mathrm{d}\pi^I \wedge \mathrm{d}\pi^{I'} \\
&= (\mathrm{d}f)f' \wedge \mathrm{d}\pi^I \wedge \mathrm{d}\pi^{I'} + f\mathrm{d}f' \wedge \mathrm{d}\pi^I \wedge \mathrm{d}\pi^{I'} \\
&= (\mathrm{d}f \wedge \mathrm{d}\pi^I) \wedge (f'\mathrm{d}\pi^{I'}) + (-1)^k(f\mathrm{d}\pi^I) \wedge (\mathrm{d}f' \wedge \mathrm{d}\pi^{I'}) \\
&= \mathrm{d}(f\mathrm{d}\pi^I) \wedge (f'\mathrm{d}\pi^{I'}) + (-1)^k(f\mathrm{d}\pi^I) \wedge \mathrm{d}(f'\mathrm{d}\pi^{I'}),
\end{aligned}$$
pois $\mathrm{d}f' \wedge \mathrm{d}\pi^I = (-1)^k\mathrm{d}\pi^I \wedge \mathrm{d}f'$.

3. Provaremos primeiros para 0-formas. Para $f \in \mathfrak{A}^0(A)$,

$$\begin{aligned}
\mathrm{d}(\mathrm{d}f) &= \mathrm{d}\left(\mathop{+}_{i\in[d]} \mathrm{D}_i f \mathrm{d}\pi^i\right) \\
&= \mathop{+}_{i<j} (\mathrm{D}_{i,j} f - \mathrm{D}_{j,i} f)\mathrm{d}\pi^i \wedge \mathrm{d}\pi^j \\
&= 0,
\end{aligned}$$

pois $\mathrm{D}_{i,j} f = \mathrm{D}_{j,i} f$. Agora, para uma k-forma $\omega \in \mathfrak{A}^k(A)$, segue do resultado anterior e do item anterior que

$$\begin{aligned}
\mathrm{d}(\mathrm{d}\omega) &= \mathrm{d}\left(\mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}\omega_I \wedge \mathrm{d}\pi^I\right) \\
&= \mathop{+}_{I\in[d]^{\uparrow k}} \mathrm{d}(\mathrm{d}\omega_I) \wedge \mathrm{d}\pi^I \\
&\quad + \mathop{+}_{I\in[d]^{\uparrow k}} \mathop{+}_{j\in[k]} (-1)^j \mathrm{d}\omega_I \wedge \mathrm{d}\pi^{i_0} \wedge \cdots \wedge \mathrm{d}(\mathrm{d}\pi^{i_j}) \wedge \cdots \wedge \mathrm{d}\pi^{i_{k-1}} \\
&= 0.
\end{aligned}$$

4. Novamente, basta considerar $f\mathrm{d}\pi^I \in \mathfrak{A}^k(A')$, pois $F^\star$ é linear. Da definição de forma puxada, segue que

$$\begin{aligned}
F^\star(\mathrm{d}(f\mathrm{d}\pi^I)) &= F^\star(\mathrm{d}f \wedge \mathrm{d}\pi^I) \\
&= \mathrm{d}(f \circ F) \wedge \mathrm{d}(\pi^{i_0} \circ F) \wedge \cdots \mathrm{d}(\pi^{i_{k-1}} \circ F) \\
&= \mathrm{d}((f \circ F)\mathrm{d}(\pi^{i_0} \circ F) \wedge \cdots \mathrm{d}(\pi^{i_{k-1}} \circ F)) \\
&= \mathrm{d}(F^\star(f\mathrm{d}\pi^I)).
\end{aligned}$$

■

21.6.7.2 Formas em variedades diferenciais

Estendemos agora a definição da derivada exterior de formas para variedades diferenciais quaisquer.

⊢ **Proposição 21.43.** *Seja* $\boldsymbol{V}$ *uma variedade diferencial. Existe única função*

$$\mathrm{d}\colon \mathfrak{A}(V) \to \mathfrak{A}(V)$$

tal que

1. *(Linearidade) Para todo* $k \in [\dim V]$, $\mathrm{d}\colon \mathfrak{A}^k(V) \to \mathfrak{A}^{k+1}(V)$ *é linear sobre* $\mathbb{R}$*;*

2. *(Derivação graduada) Para todas formas* $\omega \in \mathfrak{A}^k(V)$ *e* $\omega' \in \mathfrak{A}^{k'}(V)$,

$$\mathrm{d}(\omega \wedge \omega') = \mathrm{d}\omega \wedge \omega' + (-1)^k \omega \wedge \mathrm{d}\omega';$$

3. *(Cobordismo)* $\mathrm{d} \circ \mathrm{d} = 0$;
4. *(Diferencial) Para todo campo escalar* $f \in \mathscr{C}^\infty(V) = \mathfrak{A}^0(V)$, *a diferencial* $\mathrm{D}f$ *de* f *é a derivada exterior* $\mathrm{d}f$ *de* f, *dada por* $\mathrm{d}f(v) = \partial_v f$.

21.6.8 Derivada por fluxo

A derivada por fluxo é um modo de generalizar o conceito de derivada direcional para variedades diferenciais. No caso de derivadas direcionais, o ambiente de estudo são espaços vetoriais normados e derivamos uma função f ao longo de um vetor v, avaliando o limite para o tempo t em 0 da diferença entre o valor da função em um ponto p deslocado por tv e o valor da função em p, pela fórmula

$$\nabla_v f(p) = \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

Como a função é avaliada em um espaço vetorial, o ponto p e o vetor de deslocamento v podem ser somados, de modo que $p + tv$ é um elemento do espaço. No caso que estamos estudando, no entanto, não podemos somar um ponto p a um vetor tangente v, portanto o termo $p + tv$ na fórmula de derivada direcional não tem sentido em variedades. Existem dois principais maneiras de se resolver esse problema, e ambos consistem na seguinte observação: o que o termo $p + tv$ representa na definição de derivada direcional em espaço vetoriais normados é o deslocamento do ponto p na direção v por um tempo t, e nesse caso esse deslocamento é dado pela soma do ponto com o vetor. No contexto de variedades diferenciais, pode-se tentar definir a derivada direcional ao longo de um vetor tangente a um ponto por coordenadas, mas é simples ver que isso não será invariante por escolha de coordenadas para derivação de qualquer campo tensorial. Assim, devemos estabelecer uma maneira de deslocar um ponto ao longo de um vetor tangente que seja invariante por coordenadas. A primeira maneira envolve definir o que se chama de 'transporte paralelo' na variedade, que dará origem ao conceito de 'derivada covariante'. Isso envolve, no entanto, dotar a variedade de uma estrutura geométrica que permite dizer como se transportar um vetor tangente ao longo de uma curva de modo que esses vetores se mantenham 'paralelos' ao longo dessa curva. Por isso, não envolve somente a estrutura da variedade. A segunda maneira, a que estudaremos nesta seção, consiste em se considerar, em vez de um vetor tangente, um campo vetorial na variedade. Dado esse campo vetorial, podemos deslocar um cada vetor tangente do campo por seu fluxo local, e assim podemos definir a 'derivada por fluxo'[7] ao longo

[7] A derivada por fluxo recebe o nome na literatura de 'derivada de Lie'

do um campo vetorial. A mesma ideia pode ser aplicada a derivada de qualquer campo tensorial, não somente de campos escalares.

21.6.8.1 Derivada de campos escalares

Denotamos o fluxo local do campo vetorial $v \in \mathfrak{T}^1(\boldsymbol{V})$ no tempo t por Φ^{tv}.

:⊢ **Definição 21.47.** Sejam $\boldsymbol{V}$ uma variedade diferencial, $v \in \mathfrak{T}^1(\boldsymbol{V})$ um campo vetorial e $f \in \mathfrak{T}^0(\boldsymbol{V})$ um campo escalar. A *derivada por fluxo* de f ao longo de v (ou na direção de v) é

$$\nabla_v f(p) := \frac{f(\Phi^{tv}(p)) - f(p)}{t}.$$

21.6.8.2 Definição algébrica

⊢ **Proposição 21.44.** *Sejam $\boldsymbol{V}$ uma variedade diferencial e $v \in \mathfrak{T}^1(\boldsymbol{V})$ um campo vetorial. Existe única função*

$$\nabla_v \colon \mathfrak{T}(\boldsymbol{V}) \longrightarrow \mathfrak{T}(\boldsymbol{V})$$

tal que

1. *(Derivada direcional) Para todo $f \in \mathfrak{T}^0(\boldsymbol{V})$,*

$$\nabla_v(f) = \partial_v f.$$

2. *(Derivação) Para todos $T \in \mathfrak{T}^{(k,l)}(\boldsymbol{V})$ e $T' \in \mathfrak{T}^{(k',l')}(\boldsymbol{V})$,*

$$\nabla_v(T \otimes T') = (\nabla_v T) \otimes T' + T \otimes \nabla_v(T');$$

3. *(Contração) Para todos $T \in \mathfrak{T}^{(k,l)}(\boldsymbol{V})$, $v_0, \ldots, v_{k-1} \in \mathfrak{T}^1(\boldsymbol{V})$ e $\nu_0, \ldots, \nu_{l-1} \in \mathfrak{T}^{*1}(\boldsymbol{V})$,*

$$\begin{aligned}\nabla_v(T(v_0, \ldots, v_{k-1}, \nu_0, \ldots, \nu_{l-1})) &= (\nabla_v T)(v_0, \ldots, v_{k-1}, \nu_0, \ldots, \nu_{l-1}) \\ &+ \bigplus_{i \in [k]} T(v_0, \ldots, (\nabla_v v_i), \ldots, v_{k-1}, \nu_0, \ldots, \nu_{l-1}) \\ &+ \bigplus_{i \in [l]} T(v_0, \ldots, v_{k-1}, \nu_0, \ldots, (\nabla_v \nu_i), \ldots, \nu_{l-1}).\end{aligned}$$

4. *(Derivada exterior) $[\nabla_v, \mathrm{d}] = 0$.*

21.7 Orientação

21.7.1 Orientação de espaços lineares

Denotaremos como $\mathcal{B}(L)$ o conjunto de bases ordenadas de $\boldsymbol{L}$, ou seja,

$$\mathcal{B}(L) := \left\{ b \in L^d \;\middle|\; L = \langle b \rangle \right\}.$$

:⊢ **Definição 21.48.** Seja $\boldsymbol{L}$ um espaço linear finito sobre um corpo $\boldsymbol{C}$. Duas bases ordenadas $b, b' \in \mathcal{B}(L)$ são *coorientadas* se, e somente se,

$$\det [\mathrm{I}]_{b'}^{b} > 0,$$

em que $[\mathrm{I}]_{b'}^{b}$ é a mudança de base de b para b'.

⊢ **Proposição 21.45.** *Seja $\boldsymbol{L}$ um espaço linear finito sobre um corpo $\boldsymbol{C}$. A relação de coorientação de bases ordenadas em $\mathcal{B}(L)$ é uma relação de equivalência.*

□ *Demonstração.* (Reflexividade) Seja b uma base ordenada de $\boldsymbol{L}$. Como $[\mathrm{I}]_{b}^{b} = \mathrm{I}$, segue que $\det [\mathrm{I}]_{b}^{b} = \det \mathrm{I} = 1 > 0$, portanto b é coorientada consigo mesma. (Simetria) Sejam b e b' bases ordenadas de $\boldsymbol{L}$ tais que b e b' são coorientadas. Isso significa que $\det [\mathrm{I}]_{b'}^{b} > 0$. Mas como $[\mathrm{I}]_{b}^{b'} = ([\mathrm{I}]_{b'}^{b})^{-1}$, segue que

$$\det [\mathrm{I}]_{b}^{b'} = \det ([\mathrm{I}]_{b'}^{b})^{-1} = \det {[\mathrm{I}]_{b'}^{b}}^{-1} > 0,$$

portanto b' e b são coorientadas. (Transitividade) Sejam b, b' e b'' bases ordenadas de $\boldsymbol{L}$ tais que b e b' são coorientadas e b' e b'' são coorientadas. Isso significa que $\det [\mathrm{I}]_{b'}^{b} > 0$ e $\det [\mathrm{I}]_{b''}^{b'} > 0$. Como $[\mathrm{I}]_{b''}^{b} = [\mathrm{I}]_{b'}^{b} \circ [\mathrm{I}]_{b''}^{b'}$, segue que

$$\det [\mathrm{I}]_{b''}^{b} = \det [\mathrm{I}]_{b'}^{b} \circ [\mathrm{I}]_{b''}^{b'} = \det [\mathrm{I}]_{b'}^{b} \det [\mathrm{I}]_{b''}^{b'} > 0,$$

logo b e b'' são coorientadas. ■

Isso permite que se quociente o espaço de bases ordenadas de $\boldsymbol{L}$ em classes de equivalências, e esse quociente é o conjunto de classes de bases coorientadas de $\boldsymbol{L}$, definido a seguir. Lembremos que $\mathbb{S}^0 = \{+1, -1\}$.

:⊢ **Definição 21.49.** Sejam $\boldsymbol{L}$ um espaço linear finito sobre um corpo $\boldsymbol{C}$ e b uma base ordenada de $\boldsymbol{L}$. A *classe de orientação* gerada por b é o conjunto

$$[b] = \left\{ b' \in \mathcal{B}(L) \;\middle|\; \det [\mathrm{I}]_{b'}^{b} > 0 \right\}.$$

O conjunto dessas classes de equivalência é denotado $[\mathcal{B}(L)]$. Uma *orientação* de $\boldsymbol{L}$ é uma função injetiva

$$O \colon [\mathcal{B}(L)] \to \mathbb{S}^0$$

O conjunto de orientações de $\boldsymbol{L}$ é denotado $\mathcal{O}(L)$.

Em geral, pode-se considerar que O está definida em $\mathcal{B}(L)$ por simplicidade de notação, pois teremos $O(b) \in \mathbb{S}^0$ em vez de $O([b]) \in \mathbb{S}^0$.

A classe de orientações gerada por uma base ordenada b é o conjunto de todas as bases ordenadas de $\boldsymbol{L}$ que são coorientadas com essa base, a classe de equivalência dessa base com respeito à relação de coorientação.

A pergunta a ser feita então é: quantas orientações tem um espaço linear? A resposta intuitiva para $\mathbb{R}^1$, $\mathbb{R}^2$ e $\mathbb{R}^3$ é 2, e essa é de fato a resposta para qualquer espaço linear finito. Uma orientação determina qual é a orientação positiva e qual é a orientação negativa. Mostraremos a seguir que, no caso de $d \geq 1$, essa função é uma bijeção e, no caso de $d = 0$, é uma escolha de $+1$ ou -1 para a orientação do ponto, pois $\emptyset$ é a única base de $\{0\}$.

⊢ **Proposição 21.46.** *Seja $\boldsymbol{L}$ um espaço linear finito sobre um corpo $\boldsymbol{C}$. Então*

$$\#\left(\mathcal{O}(L)\right) = 2.$$

□ *Demonstração.* Separamos a demonstração em dois casos: o primeiro em que $\dim(L) = 0$ e o segundo em que $\dim(L) > 0$. No primeiro caso, a única base de $\boldsymbol{L}$ é $\emptyset$, o que implica que existem duas funções injetivas do conjunto de classes de orientação de bases para $\{+1, -1\}$: a função $O(\emptyset) = +1$ e a função $O(\emptyset) = -1$. Portanto $\#\left(\mathcal{O}(L)\right) = 2$.

No segundo caso, se $\dim(L) > 0$, denotemos $d = \dim(L)$ e seja $b = (b_0, \ldots, b_{d-1})$ uma base ordenada de $\boldsymbol{L}$. Definimos a base $\bar{b} := (-b_0, \ldots, b_{d-1})$. Mostraremos que as únicas classes de orientação de $\boldsymbol{L}$ são $[b]$ e $[\bar{b}]$. Seja b' uma base de $\boldsymbol{L}$. Se $\det [\mathrm{I}]^b_{b'} > 0$, então $[b'] = [b]$; caso contrário, $\det [\mathrm{I}]^b_{b'} < 0$. Mas

$$\begin{aligned}
\det [\mathrm{I}]^b_{b'} &= \det \left[[b_0]_{b'} \cdots [b_{d-1}]_{b'}\right] \\
&= -\det \left[[-b_0]_{b'} \cdots [b_{d-1}]_{b'}\right] \\
&= -\det [\mathrm{I}]^{\bar{b}}_{b'},
\end{aligned}$$

portanto $\det [\mathrm{I}]^{\bar{b}}_{b'} > 0$, o que mostra que $b' \in [\bar{b}]$. Isso implica que só existem duas orientações de $\boldsymbol{L}$, pois da injetividade segue que $O([b]) = +1$ e $O([\bar{b}]) = -1$, ou $O([b]) = -1$ e $O([\bar{b}]) = +1$. Logo $\#\left(\mathcal{O}(L)\right) = 2$. ■

A orientação canônica em $\mathbb{R}^d$ é a função O definida por $O([e_0, \ldots, e_{d-1}]) = +1$ e $O([-e_0, \ldots, e_{d-1}]) = -1$.

21.7.1.1 Aritmética de orientações

Definimos a orientação de um isomorfismo linear $f \colon L \to L$ por

$$O(f) := \frac{\det f}{|\det f|}.$$

Essa é uma função $O\colon \mathcal{L}(L) \to \{+1, -1\}$. Assim, segue que, para todos isomorfismos $f, f' \in \mathcal{L}(L)$,

$$O(f' \circ f) = O(f')O(f),$$

pois $\det f' \circ f = \det f' \det f$.

Definimos, para toda classe de equivalência $[b]$ de bases ordenadas de $\boldsymbol{L}$ coorientadas e todo isomorfismo $f \in \mathcal{L}(L)$,

$$f[b_0, \ldots, b_{d-1}] := [f(b_0), \ldots, f(b_{d-1})].$$

Isso está bem definido. Assim, segue que, para toda base ordenada b de $\boldsymbol{L}$,

$$O(f[b]) = O(f)O([b]).$$

Vale notar que ambas as funções orientação, tanto a de isomorfismo como a de bases, define uma função com valores em $\{0, 1\}$, através do isomorfismo de grupos de $\{+1, -1\}$ para $\{0, 1\}$ definido por $(-1)^0 = +1$ e $(-1)^1 = -1$.

21.7.2 Orientação de variedades

:⊢ **Definição 21.50.** Seja $\boldsymbol{V}$ uma variedade diferencial. Uma *orientação* de $\boldsymbol{V}$ é uma função

$$O\colon V \to \bigcup_{p \in V} \mathcal{O}(\mathrm{T}V|_p)$$

tal que

1. Para todo $p \in V$, $O|_p\colon [\mathcal{B}(\mathrm{T}V|_p)] \to \{+1, -1\}$ é uma orientação de $\mathrm{T}V|_p$;
2. Para todo $\bar{p} \in V$, existe um referencial local diferenciável $(B_i)_{i \in [d]}$ de $\boldsymbol{V}$ em uma vizinhança $A \subseteq V$ de $\bar{p}$ tal que, para todos $p, p' \in A$,

$$O|_p([B_0|_p, \ldots, B_{d-1}|_p]) = O|_{p'}([B_0|_{p'}, \ldots, B_{d-1}|_{p'}]).$$

Uma variedade diferencial *orientável* é uma variedade diferencial que admite uma orientação $\mathcal{O}$.

21.7.2.1 Atlas orientados

:⊢ **Definição 21.51.** Seja $\boldsymbol{V}$ uma variedade diferencial. Duas cartas (A, x), (A', x') de $\boldsymbol{V}$ são *coorientadas* se, e somente se, para todo $p \in A \cap A'$,

$$\det \mathrm{D}(x' \circ x^{-1})|_p > 0.$$

Caso contrário, elas são *contraorientadas*, e existe $p \in A \cap A'$ tal que

$$\det \mathrm{D}(x' \circ x^{-1})|_p < 0.$$

O caso em que $\det \mathrm{D}(x' \circ x^{-1})|_p = 0$ não ocorre, pois as cartas são diferencialmente compatíveis, o que implica que a transição de coordenadas é um difeomorfismo, logo $\mathrm{D}(x' \circ x^{-1})(p) \colon \mathbb{R}^d \to \mathbb{R}^d$ é invertível.

:⊢ **Definição 21.52.** Um atlas diferencial *orientado* é um atlas cujas cartas são coorientadas duas a duas. Atlas diferenciais orientados *consistentes* são atlas diferenciais orientados cuja união é um atlas orientado.

Pode-se verificar que a relação de coorientabilidade de atlas é uma relação de equivalência.

⊢ **Proposição 21.47.** *Seja $\boldsymbol{V}$ uma variedade diferencial orientável com n componentes conexas. Então*

$$\#\left(\mathcal{O}(V)\right) = 2^n.$$

21.8 Folheações

As definições e os principais resultados desta seção são baseados no livro *Foliations I*, Candel e Conlon.

21.8.1 Cartas e atlas folheados

:⊢ **Definição 21.53.** Seja $d \in \mathbb{N}$. Um *retângulo aberto* d-dimensional é um conjunto $R \subseteq \mathbb{R}^d$ tal que, para intervalos abertos $I_0, \ldots, I_{d-1} \subseteq \mathbb{R}$,

$$R = I_0 \times \cdots \times I_{d-1}.$$

:⊢ **Definição 21.54.** Sejam $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional e $k \in [d]$. Uma *carta k-folheada* (ou *k-dimensionalmente folheada*) de $\boldsymbol{V}$ é uma carta $(A, x) \in \mathcal{A}$ para a qual existem retângulos abertos $B_{\pitchfork} \subseteq \mathbb{R}^k$ e $B_{\pitchfork} \subseteq \mathbb{R}^{d-k}$ tais que

$$x(A) = B_{\pitchfork} \times B_{\pitchfork}.$$

Para todo $y \in B_{\pitchfork}$, a *placa* (*tangencial*) de (A, x) relativa a y é o conjunto $P_y = x^{-1}(B_{\pitchfork} \times \{y\}) \subseteq A$; para todo $x \in B_{\pitchfork}$, a (*seção*) *transversal* de (A, x) relativa a x é o conjunto $P_x = x^{-1}(\{x\} \times B_{\pitchfork}) \subseteq A$.

Os símbolos $\pitchfork$ e $\pitchfork$ são lidos *tangente* e *transverso*, respectivamente.

⊢ **Proposição 21.48.** *Seja $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d- dimensional e $(A, x) \in \mathcal{A}$ uma carta k-folheada de $\boldsymbol{V}$. As placas $\{P_y\}_{y \in B_{\pitchfork}}$ são uma partição do domínio A de (A, x).*

□ *Demonstração.* Para mostrar isso, primeiro notamos que $P_y = x^{-1}(B_{\sqcap} \times \{y\}) \neq \emptyset$, pois x é sobrejetiva. Segundo, como $x(A) = B_{\sqcap} \times B_{\pitchfork}$, então

$$\begin{aligned} A &= x^{-1}(B_{\sqcap} \times B_{\pitchfork}) \\ &= x^{-1}\left(B_{\sqcap} \times \bigcup_{y \in B_{\pitchfork}} \{y\}\right) \\ &= \bigcup_{y \in B_{\pitchfork}} x^{-1}(B_{\sqcap} \times \{y\}) \\ &= \bigcup_{y \in B_{\pitchfork}} P_y. \end{aligned}$$

Por fim, notemos que, para todos $y, y' \in B_{\pitchfork}$, temos

$$\begin{aligned} P_y \cap P_{y'} &= x^{-1}(B_{\sqcap} \times \{y\}) \cap x^{-1}(B_{\sqcap} \times \{y'\}) \\ &= x^{-1}(B_{\sqcap} \times \{y\} \cap B_{\sqcap} \times \{y'\}) \\ &= x^{-1}(B_{\sqcap} \times (\{y\} \cap \{y'\})). \end{aligned}$$

portanto, se $y = y'$, $P_y = P_{y'}$ e, se $y \neq y'$, $P_y \cap P_{y'} = \emptyset$. ■

:⊢ **Definição 21.55.** Seja $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. Cartas folheadas de $\boldsymbol{V}$ *coerentemente folheadas* são cartas folheadas $(A, x), (A', x') \in \mathcal{A}$ tais que, para todas placas P de (A, x) e P' de (A', x'), a interseção $P \cap P'$ é aberta em P e em P'. Denota-se $(A, x) \approx (A', x')$.

Por definição, as cartas têm que ser ambas folheadas com placas de mesma dimensão caso tenham interseção não vazia, caso contrário a interseção de duas placas não seria aberta na placa de dimensão maior. Note que a relação de coerência de folheação é simétrica por definição e é reflexiva pois, para todos $y, y' \in B_{\pitchfork}$ com placas relativas $P_y, P_{y'} \subseteq A$ de uma carta folheada (A, x), temos $P_y = P_{y'}$, se, e somente se, $y = y'$, logo $P_y \cap P_{y'}$ é P_y se $y = y'$, ou $\emptyset$ se $y \neq y'$. Como P_y é aberto em si mesmo e $\emptyset$ é aberto, segue que (A, x) é coerentemente folheada consigo mesma. Para mostrar que a relação é transitiva, no entanto, precisamos de um atlas de cartas coerentemente folheadas.

:⊢ **Definição 21.56.** Seja $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. Um *atlas k-folheado*[8] de $\boldsymbol{V}$ é um subatlas $\mathcal{A}' \subseteq \mathcal{A}$ cujas cartas são cartas k-folheadas e coerentemente folheadas duas a duas.

:⊢ **Definição 21.57.** Seja $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. Atlas *coerentemente folheados* de $\boldsymbol{V}$ são atlas folheados $\mathcal{A}$ e $\mathcal{A}'$ de $\boldsymbol{V}$ tais que todas cartas $(A, x) \in \mathcal{A}$ e $(A', x') \in \mathcal{A}'$ são coerentemente folheadas. Denota-se $\mathcal{A} \approx \mathcal{A}'$.

[8] Novamente, ou *k-dimensionalmente folheado.*

Essa definição é equivalente a dizer que $\mathcal{A} \cup \mathcal{A}$ é um atlas folheado.

$\vdash$ **Proposição 21.49.** *Seja $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. A relação de coerência de folheação de atlas é uma relação de equivalência.*

$\square$ *Demonstração.* A reflexividade e a simetria são evidentes, pois valem entre cartas folheadas. Para conferir a transitividade, basta mostrar que, se $\mathcal{A} = \{(A_i, x_i)\}_{i \in I}$ é um atlas k-folheado de $\boldsymbol{V}$ e $(A, x), (A', x')$ são cartas k-folheadas de $\boldsymbol{V}$ que são coerentemente folheadas com cada carta de $\mathcal{A}$, então as cartas (A, x) e (A', x') são coerentemente folheadas. Sejam P e P' placas de (A, x) e (A', x'), respectivamente. Se $P \cap P' = \emptyset$, $P \cap P'$ é aberto em P e em P'. Caso contrário, seja $p \in P \cap P'$. Como $\mathcal{A}$ é atlas, existe $i \in I$ tal que $p \in A_i$. Seja Q_p a placa de (A_i, x) tal que $p \in Q_p$. Como $(A, x) \approx (A_i, x_i)$, segue que $P \cap Q_p$ é aberto em P e em Q_p e, como $(A_i, x_i) \approx (A', x')$ que $Q_p \cap P'$ é aberto em Q_p e em P'. Assim, segue que $P \cap Q_p \cap P'$ é aberto em Q_p e, por uma mudança de parametrização, segue que é aberto em P e em P' também. Como

$$P \cap P' = \bigcup_{p \in P \cap P'} P \cap Q_p \cap P',$$

concluímos que $P \cap P'$ é aberto em P e em P', pois é união de abertos. Portanto $(A, x) \approx (A', x')$, o que mostra que $\mathcal{A} \approx \mathcal{A}'$. ■

21.8.2 Componentes tangencial e transversal da carta

$:\vdash$ **Definição 21.58.** Sejam $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional e $(A, x) \in \mathcal{A}$ uma carta k-folheada de $\boldsymbol{V}$. A *projeção tangencial* de (A, x) é a projeção canônica $\mathrm{p}_{\pitchfork}\colon B_{\pitchfork} \times B_{\pitchfork} \to B_{\pitchfork}$ e a *componente tangencial* de x é a função $x_{\pitchfork} := \mathrm{p}_{\pitchfork} \circ x\colon A \to B_{\pitchfork}$. A *projeção transversal* de (A, x) é a projeção canônica $\mathrm{p}_{\pitchfork}\colon B_{\pitchfork} \times B_{\pitchfork} \to B_{\pitchfork}$ e a *componente transversal* de x é a função $x_{\pitchfork} := \mathrm{p}_{\pitchfork} \circ x\colon A \to B_{\pitchfork}$.

Mostraremos uma equivalência da definição de coerência de folheação que será útil.

$\vdash$ **Proposição 21.50.** *Sejam $\boldsymbol{V} = (V, \mathcal{A})$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional e (A, x), $(A', x') \in \mathcal{A}$ cartas k-folheadas de $\boldsymbol{V}$. As cartas são coerentemente folheadas se, e somente se, todo ponto $p \in A \cap A'$ tem vizinhança em que a transição de coordenadas*

$$x' \circ x^{-1}\colon x(A \cap A') \to x'(A \cap A')$$

satisfaz, para todo $(p, q) \in x(A \cap A')$, com $p \in \mathbb{R}^k$ e $q \in \mathbb{R}^{d-k}$,

$$x' \circ x^{-1}(p, q) = (h_{\pitchfork}(p, q), h_{\pitchfork}(q)) \in \mathbb{R}^k \times \mathbb{R}^{d-k}.$$

□ *Demonstração.* (De Candel e Conlon) Suponhamos que as cartas são coerentemente folheadas. Consideremos em $x(A \cap A')$ a mudança de coordenadas

$$x' \circ x^{-1}(p,q) = (x'_{\pitchfork} \circ x^{-1}(p,q), x'_{\pitchfork} \circ x^{-1}(p,q)).$$

Sejam P placa de (A,x) e P' placa de (A',x'). Como as cartas são coerentemente folheadas, as componentes conexas de $P \cap A'$ estão contidas em placas de A', e o mesmo vale para P'. Equivalentemente, como as placas P e P' são conjuntos de nível de y e y', respectivamente, cada ponto em $A \cap A'$ tem vizinhança em que vale

$$x'_{\pitchfork} \circ x^{-1}(p,q) = x'_{\pitchfork} \circ x^{-1}(p',q),$$

ou seja, $x'_{\pitchfork} \circ x^{-1}$ independe de $B_{\pitchfork}$. Definindo $h_{\pitchfork}(p,q) := x'_{\pitchfork} \circ x^{-1}(p,q)$ e $h_{\pitchfork}(q) = x'_{\pitchfork} \circ x^{-1}(p,q)$ para qualquer p, segue o resultado. A recíproca fica como exercício. ■

21.8.3 Folheação

:⊢ **Definição 21.59.** Seja $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. Uma *k-folheação* de $\boldsymbol{V}$ é uma partição $\mathcal{F}$ de V em subvariedades k-dimensionais imersas[9] de V, as *folhas* de $\mathcal{F}$, para a qual existe atlas k-folheado $\mathcal{A}_{\mathcal{F}}$ de V satisfazendo que, para toda carta folheada $(A,x) \in \mathcal{A}_{\mathcal{F}}$ e toda $F \in \mathcal{F}$, $F \cap A$ é uma união de placas de $\mathcal{A}_{\mathcal{F}}$.

⊢ **Proposição 21.51.** *Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional e $\mathcal{F}$ uma k-folheação de $\boldsymbol{V}$. As folhas de $\mathcal{F}$ são conexas.*

21.8.3.1 Definições alternativas

Folheações (baseado em Camacho)

:⊢ **Definição 21.60.** Seja $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional. Uma *folheação* $\mathscr{C}^s$ k-dimensional de $\boldsymbol{V}$ (para $s \leq r$ e $k \leq d$) é um atlas $\mathscr{C}^s$ $\mathcal{F}$ de $\boldsymbol{V}$ que é maximal segundo as propriedades:

1. Para toda carta $(A,x) \in \mathcal{F}$, existem bolas abertas $B_{\pitchfork} \subseteq \mathbb{R}^k$ e $B_{\pitchfork} \subseteq \mathbb{R}^{d-k}$ tais que
$$x(A) = B_{\pitchfork} \times B_{\pitchfork};$$
2. Para todas cartas (A,x) e $(A',x') \in \mathcal{F}$, a transição de coordenadas
$$x' \circ x^{-1} \colon x(A \cap A') \to x'(A \cap A')$$
satisfaz
$$x' \circ x^{-1}(x,y) = (h_{\pitchfork}(x,y), h_{\pitchfork}(y)) \in \mathbb{R}^k \times \mathbb{R}^{d-k}$$
para todo $(x,y) \in x(A \cap A')$, com $x \in \mathbb{R}^k$ e $y \in \mathbb{R}^{d-k}$.

[9] Isso quer dizer que a subvariedade é imagem de uma imersão injetiva em V.

Nesse caso, a variedade $\boldsymbol{V}$ é *folheada* por $\mathcal{F}$ e as cartas de $\mathcal{F}$ são *cartas de folheação*.

:⊢ **Definição 21.61.** Sejam $\boldsymbol{V}$ uma variedade $\mathscr{C}^{1+r}$ d-dimensional, $\mathcal{F}$ uma folheação $\mathscr{C}^s$ k-dimensional de $\boldsymbol{V}$, $(A, x) \in \mathcal{F}$ uma carta de folheação e $x(A) = B_{\pitchfork} \times B_{\pitchfork}$. Uma *placa* de $\mathcal{F}$ baseada em (A, x) é um conjunto

$$P_c = x^{-1}(B_{\pitchfork} \times \{c\}),$$

em que $c \in B_{\pitchfork}$.

⊢ **Proposição 21.52.** *1. A função*

$$x^{-1}|_{B_{\pitchfork} \times \{c\}} \colon B_{\pitchfork} \times \{c\} \to A$$

é um mergulho $\mathscr{C}^{1+r}$. Portanto as placas são subvariedades $\mathscr{C}^{1+r}$ k-dimensionais.

2. As placas baseadas em (A, x) são disjuntas duas a duas.

Folheações (Baseado em R. Bowen e B. Marcus, Unique Ergodicity for Horocycle Foliations)

:⊢ **Definição 21.62.** Sejam X uma variedade topológica d-dimensional, $x \in X$ e $\mathcal{F}$ uma partição de X em subvariedades conexas de X. Um conjunto *k-transversal* a $\mathcal{F}$ em x é uma vizinhança compacta K de x (de $\operatorname{codim}_X(K) = n = d - k$) satisfazendo que

1. Existe uma função contínua injetiva $\phi \colon K \times \mathbb{B}^n \to X$ tal que $\phi(K \times \mathbb{B}^n)$ é uma vizinhança[10] de x;
2. Para todo $y \in K$, existe folha $F \in \mathcal{F}$ tal que $\phi(y, 0) = y$ e $\phi(\{y\} \times \mathbb{B}^n) \subseteq F$.

Denota-se $K \pitchfork \mathcal{F}$.

:⊢ **Definição 21.63.** Seja X uma variedade d-dimensional. Uma *folheação* k-dimensional de X é uma partição $\mathcal{F}$ de X em subvariedades conexas de X tal que, para todo $x \in X$, existe uma vizinhança compacta K de x $(d - k)$-transversal a $\mathcal{F}$ em x. As *folhas* de $\mathcal{F}$ são as partes de $\mathcal{F}$.

[10] Aqui aparece a dimensão $n = d - k$ das subvariedades.

Capítulo 22

Fibrados

22.1 Fibrados topológicos

:⊢ **Definição 22.1.** Sejam $\boldsymbol{X}$ e $\boldsymbol{F}$ espaços topológicos. Um *fibrado (topológico) de* $\boldsymbol{F}$ *sobre* $\boldsymbol{X}$ é um par $(\boldsymbol{E}, \mathrm{p})$, em que $\boldsymbol{E}$ é um espaço topológico e $\mathrm{p}\colon E \to X$ é uma função contínua sobrejetiva que satisfaz: para todo $e \in E$, existem vizinhança $A \subseteq X$ de $\mathrm{p}(e)$ e, definindo $E|_A := \mathrm{p}^{-1}(A) \subseteq E$, homeomorfismo $h\colon E|_A \to A \times F$ tais que $\mathrm{p}_A \circ h = \mathrm{p}$ (o diagrama comuta).

$$\begin{tikzcd}
E|_A \arrow[r, "h"] \arrow[d, "\mathrm{p}"'] & A \times F \arrow[dl, "\mathrm{p}_A"] \\
A &
\end{tikzcd}$$

O espaço $\boldsymbol{E}$ é o *espaço fibrado*, o espaço $\boldsymbol{X}$ é a *base*, o espaço $\boldsymbol{F}$ é a *fibra* e a função $\mathrm{p}\colon E \to X$ é a *projeção fibrada* de $(\boldsymbol{E}, \mathrm{p})$. Para cada $x \in X$, a *fibra de* E *em* x é $E|_x := \mathrm{p}^{-1}(\{x\})$.

A definição significa que o espaço fibrado $\boldsymbol{E}$ é localmente um produto de sua base $\boldsymbol{X}$ e sua fibra $\boldsymbol{F}$. Globalmente isso não precisa ocorrer, o espaço fibrado não precisa ser homeomorfo ao produto de sua base com sua fibra. Quanto isso ocorre, o fibrado é chamado trivial.

⊢ **Proposição 22.1.** *Sejam* $\boldsymbol{X}$ *e* $\boldsymbol{F}$ *espaços topológicos e* $(\boldsymbol{E}, \mathrm{p})$ *um fibrado de* $\boldsymbol{F}$ *sobre* $\boldsymbol{X}$*.*

1. *Para todo* $x \in X$*, a fibra* $E|_x$ *é homeomorfa a* F*;*

2. *A projeção fibrada* $\mathrm{p}\colon E \to X$ *é aberta e a topologia de* $\boldsymbol{X}$ *é a topologia quociente de* $\boldsymbol{E}$ *por* p.

⊢ **Proposição 22.2.** *Sejam* $\boldsymbol{X}$ *e* $\boldsymbol{F}$ *espaços topológicos. O par* $(\boldsymbol{X} \times \boldsymbol{F}, \mathrm{p}_X)$ *é um fibrado de* $\boldsymbol{F}$ *sobre* $\boldsymbol{X}$.

□ *Demonstração.* Basta tomar $A = X$ e $h = \mathrm{I}\colon X \times F \to X \times F$ e segue que h é homeomorfismo e $\mathrm{p}_X \circ \mathrm{I} = \mathrm{p}_X$. ■

► **Exemplo 22.1** (Faixa Torcida (Möbius))**.** *Consideremos em* $\mathbb{R}^2$ *a equivalência*

$$(x, y) \sim (x', y') \quad \Longleftrightarrow \quad \exists_{n\in\mathbb{Z}}(x', y') = (x + n, (-1)^n y).$$

Defina $E := \mathbb{R}^2/\!\sim$ *e denote por* $q\colon \mathbb{R}^2 \to E$ *o mapa quociente. Consideremos o mapa* $\mathrm{p}_0\colon \mathbb{R}^2 \to \mathbb{R}$ *a projeção na coordenada* 0 *e* $r\colon \mathbb{R} \to \mathbb{S}^1$ *o recobrimento*

$$\begin{aligned} r\colon \mathbb{R} &\longrightarrow \mathbb{S}^1 \\ x &\longmapsto (\cos(\tau x), \operatorname{sen}(\tau x)). \end{aligned}$$

Como $r \circ \mathrm{p}_0\colon \mathbb{R}^2 \to \mathbb{S}^1$ *é constante em cada classe de equivalência, existe única função contínua* $\mathrm{p}\colon E \to \mathbb{S}^1$ *tal que* $q \circ \mathrm{p} = r \circ \mathrm{p}_0$ *(o diagrama comuta).*

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "q"] \arrow[d, "\mathrm{p}_0"'] & E \arrow[d, dashed, "\mathrm{p}"] \\ \mathbb{R} \arrow[r, "r"'] & \mathbb{S}^1 \end{tikzcd}$$

O par $(\boldsymbol{E}, \mathrm{p})$ *é um fibrado de* $\mathbb{R}$ *sobre* $\mathbb{S}^1$*, chamado de* faixa torcida.

A faixa torcida é um importante objeto topológico e geométrico. Ela está intrinsecamente ligada ao fenômeno de não-orientabilidade de variedades diferenciais, como será demonstrado em capítulos mais à frente.

22.2 Fibrados vetoriais

Nessa seção, consideraremos espaços lineares reais de dimensão finita, ou seja, $\mathbb{R}^n$, mas poderíamos considerar espaços lineares topológicos $\boldsymbol{L}$.

:⊢ **Definição 22.2.** Sejam $\boldsymbol{X}$ um espaço topológico e $n \in \mathbb{N}$. Um *fibrado vetorial de posto* n *sobre* $\boldsymbol{X}$ (ou *fibrado vetorial de* $\mathbb{R}^n$ *sobre* $\boldsymbol{X}$) é um par $(\boldsymbol{E}, \mathrm{p})$, em que $\boldsymbol{E}$ é um espaço topológico e $\mathrm{p}\colon E \to X$ é uma função contínua sobrejetiva que, definido, para todo $A \subseteq X$ e todo $x \in X$, $E|_A := \mathrm{p}^{-1}(A) \subseteq E$ e $E|_x := E|_{\{x\}}$, satisfaz:

1. Para todo $x \in X$, $E|_x$ tem estrutura linear;
2. Para todo $e \in E$, existem vizinhança $A \subseteq X$ de $\mathrm{p}(e)$ e homeomorfismo $h\colon E|_A \to A \times \mathbb{R}^n$ tais que

   2.1. $\mathrm{p}_A \circ h = \mathrm{p}$ (o diagrama comuta).

$$\begin{tikzcd} E|_A \arrow[r, "h"] \arrow[d, "\mathrm{p}"'] & A \times \mathbb{R}^n \arrow[dl, "\mathrm{p}_A"] \\ A & \end{tikzcd}$$

   2.2. Para todo $x \in A$, $h|_x\colon E|_x \to \{x\} \times \mathbb{R}^n \simeq \mathbb{R}^n$ é um isomorfismo linear.

O espaço $\boldsymbol{E}$ é o *espaço fibrado*, o espaço $\boldsymbol{X}$ é a *base*, o espaço $\mathbb{R}^n$ é a *fibra* e a função $\mathrm{p}\colon E \to X$ é a *projeção fibrada* de $\boldsymbol{E}$. Cada par (A, h) como acima é uma *trivialização local*. Para cada $x \in X$, a *fibra de E sobre x* é o espaço $E|_x = \mathrm{p}^{-1}(\{x\})$.

Um *fibrado vetorial diferencial* é um fibrado vetorial em que $\boldsymbol{E}$ e $\boldsymbol{X}$ são variedades diferenciais, p é diferenciável e existem cartas fibradas (A, h) como acima tais que h é difeomorfismo.

A seguinte proposição é evidente.

$\vdash$ **Proposição 22.3.** *Sejam $\boldsymbol{X}$ um espaço topológico (variedade diferencial), $n \in \mathbb{N}$ e $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial (diferencial) de $\mathbb{R}^n$ sobre $\boldsymbol{X}$. O par $(\boldsymbol{E}, \mathrm{p})$ é um fibrado topológico de $\mathbb{R}^n$ sobre $\boldsymbol{X}$.*

$\vdash$ **Proposição 22.4** (Fibrado Tangente)**.** *Sejam $\boldsymbol{V}$ uma variedade diferencial d-dimensional. O par $(\mathrm{T}\boldsymbol{V}, \mathrm{p})$, em que $\mathrm{T}\boldsymbol{V}$ é o fibrado tangente, $\mathrm{p}\colon V \to \mathrm{T}V$ é a projeção canônica e $\mathrm{T}V|_p$ tem a estrutura linear canônica, é um fibrado vetorial diferencial de posto d sobre $\boldsymbol{V}$.*

$\square$ *Demonstração.* Sejam $p \in V$ e (A, x) uma carta de $\boldsymbol{V}$ em p. Definimos a função

$$\begin{aligned} h\colon \mathrm{T}V|_A &\longrightarrow A \times \mathbb{R}^d \\ \mathop{+}_{i \in [d]} v^i \left.\frac{\partial}{\partial x^i}\right|_p &\longmapsto (p, (v^0, \ldots, v^{d-1})). \end{aligned}$$

Claramente $\mathrm{p}_A \circ h = \mathrm{p}$, e $h|_p\colon \mathrm{T}V|_p \to \mathrm{T}V|_p$ é linear. Seja $(\mathrm{T}V|_A, \varphi)$ a carta de $\mathrm{T}V$ induzida de (A, x). Então temos que $\varphi = (x, \mathrm{I}) \circ h\colon \mathrm{T}V|_A \to x(A) \times \mathbb{R}^d$. Como φ e (x, I) são difeomorfismos, h também é, o que mostra que (A, h) é uma trivialização local em p. Isso conclui a demonstração. $\blacksquare$

22.2.1 Funções de transição

Relembremos que $\overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ é o grupo de homeomorfismos lineares de $\mathbb{R}^n$ para $\mathbb{R}^n$.

$\vdash$ **Proposição 22.5.** *Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$. Para todas trivializações locais (A, h) e (A', h') de $\boldsymbol{E}$ tais que $A \cap A' \neq \emptyset$, existe função diferenciável*

$$\begin{aligned} T^h_{h'} \colon A \cap A' &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto T^h_{h'}|_p. \end{aligned}$$

tal que, para todos $(p, v) \in (A \cap A') \times \mathbb{R}^n$,

$$h' \circ h^{-1}(p, v) = (p, T^h_{h'}|_p(v)).$$

$\square$ *Demonstração.* Sejam (A, h) e (A', h') trivializações locais de $\boldsymbol{E}$ tais que $A \cap A' \neq \emptyset$. O seguinte diagrama comuta.

$$\begin{tikzcd} (A \cap A') \times \mathbb{R}^n \arrow[dr, "\mathrm{p}_A"'] & E|_{A \cap A'} \arrow[l, "h'"'] \arrow[r, "h"] \arrow[d, "\mathrm{p}"] & (A \cap A') \times \mathbb{R}^n \arrow[dl, "\mathrm{p}_A"] \\ & A & \end{tikzcd}$$

Segue que $\mathrm{p}_A \circ (h' \circ h^{-1}) = \mathrm{p}_A$, o que implica que existe função diferenciável $\sigma \colon (A \cap A') \times \mathbb{R}^n \to \mathbb{R}^n$ tal que, para todo $(p, v) \in$

$$h' \circ h^{-1}(p, v) = (p, \sigma(p, v)).$$

Para todo $p \in A \cap A'$, $T^h_{h'}|_p := \sigma(p, \cdot) \colon \mathbb{R}^n \to \mathbb{R}^n$ é um homeomorfismo linear. Assim, a função

$$\begin{aligned} T^h_{h'} \colon A \cap A' &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto T^h_{h'}|_p \end{aligned}$$

é diferenciável e satisfaz o enunciado. $\blacksquare$

$\vdash$ **Definição 22.3.** Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$. Para todas trivializações locais (A, h) e (A', h') de $\boldsymbol{E}$ tais que $A \cap A' \neq \emptyset$, a *função de transição* de (A, h) para (A', h') é a função

$$\begin{aligned} T^h_{h'} \colon A \cap A' &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto T^h_{h'}|_p. \end{aligned}$$

No caso do fibrado tangente $\mathrm{T}\boldsymbol{V}$ de uma variedade diferencial $\boldsymbol{V}$, dadas cartas (A, x) e (A', x') de $\boldsymbol{V}$ tais que $A \cap A' \neq 0$, a função de transição das cartas $(\mathrm{T}V|_A, \varphi)$ e $(\mathrm{T}V|_{A'}, \varphi')$ de $\mathrm{T}\boldsymbol{V}$ induzidas por (A, x) e (A', x'), respectivamente, é a função

$$\begin{aligned} \mathrm{D}(x' \circ x^{-1})\colon A \cap A' &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto \mathrm{D}(x' \circ x^{-1})|_p. \end{aligned}$$

$\vdash$ **Proposição 22.6** (Atlas Fibrado)**.** *Sejam $\boldsymbol{V}$ uma variedade diferencial d-dimensional, $(E|_p)_{p \in V}$ uma família de espaços lineares reais n-dimensionais, $E := \bigsqcup_{p \in V} E|_p$, $\mathrm{p}\colon E \to V$ uma função tal que $\mathrm{p}(E|_p) = \{p\}$ e $\mathcal{A}$ um conjunto de pares (A, h) tais que*

1. *$(A)_{(A,h) \in \mathcal{A}}$ é uma cobertura aberta de V;*
2. *Para toda $(A, h) \in \mathcal{A}$, com $E|_A := \mathrm{p}^{-1}(A)$, $h\colon E|_A \to A \times \mathbb{R}^n$ é uma bijeção tal que $h \lfloor_{E|_p}\colon E|_p \to \{p\} \times \mathbb{R}^n \simeq \mathbb{R}^n$ é um isomorfismo linear;*
3. *Para todas $(A, h), (A', h') \in \mathcal{A}$, existe função diferenciável $T^h_{h'}\colon A \cap A' \to \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ tal que, para todo $(p, v) \in (A \cap A') \times \mathbb{R}^n$,*

$$h' \circ h^{-1}(p, v) = (p, T^h_{h'}|_p(v)).$$

Então existe única estrutura diferencial sobre E que o torna uma variedade diferencial $(d+n)$-dimensional de modo que $(\boldsymbol{E}, \mathrm{p})$ é um fibrado vetorial de $\mathbb{R}^n$ sobre $\boldsymbol{V}$.

$\square$ *Demonstração.* Seja $p \in V$ e $(A, h) \in \mathcal{A}$ tal que $p \in A$. Seja (U_p, x_p) uma carta de V em p tal que $U_p \subseteq A$. Defina a função

$$\bar{x}_p := (x_p \times \mathrm{I}_{\mathbb{R}^n}) \circ h\colon E|_{U_p} \to x_p(U_p) \times \mathbb{R}^n \subseteq \mathbb{R}^d \times \mathbb{R}^n.$$

Mostraremos que

$$\{(E|_{U_p}, \bar{x}_p)\}_{p \in V}$$

é um atlas diferencial $(d+n)$-dimensional de E.

1. (Cartas) Como (U_p, x_p) é carta, x_p é injetivo, portanto $(x_p \times \mathrm{I}_{\mathbb{R}^n})$ é injetivo; como h é injetivo, segue que $\bar{x}_p$ é injetivo.
   Como (U_p, x_p) é carta, $x_p(U_p)$ é aberto, logo

$$\bar{x}_p(E|_{U_p}) = (x_p \times \mathrm{I}_{\mathbb{R}^n}) \circ h(E|_{U_p}) = (x_p \times \mathrm{I}_{\mathbb{R}^n})(U_p \times \mathbb{R}^n) = x_p(U_p) \times \mathbb{R}^n$$

   é aberto.
2. (Cobertura) Como $p \in U_p$, $E|_p \subseteq E|_{U_p} \subseteq E$, logo

$$E = \bigcup_{p \in V} E|_p \subseteq \bigcup_{p \in V} E|_{U_p} \subseteq E.$$

3. (Compatibilidade) Sejam $p, p' \in V$. Então

$$E|_{U_p} \cap E|_{U_{p'}} = \mathrm{p}^{-1}(U_p) \cap \mathrm{p}^{-1}(U_{p'}) = \mathrm{p}^{-1}(U_p \cap U_{p'}) = E|_{U_p \cap U'_p}.$$

Como (U_p, x_p) e (U_p, x_p) são compatíveis, então $x_p(U_p \cap U_{p'})$ e $x_{p'}(U_p \cap U_{p'})$ são abertos, logo

$$\bar{x}_p(E|_{U_p} \cap E|_{U_{p'}}) = \bar{x}_p(E|_{U_p \cap U'_p}) = x_p(U_p \cap U_{p'}) \times \mathbb{R}^n$$

e

$$\bar{x}_{p'}(E|_{U_p} \cap E|_{U'_p}) = \bar{x}_{p'}(E|_{U_p \cap U'_p}) = x_{p'}(U_p \cap U_{p'}) \times \mathbb{R}^n$$

são abertos.
Como $x_{p'}$ e x_p são difeomorfismos, $(x_{p'} \times \mathrm{I}_{\mathbb{R}^n})$ e $(x_p \times \mathrm{I}_{\mathbb{R}^n})$ são difeomorfismos; como $h' \circ h^{-1}$ é difeomorfismo, segue que

$$\bar{x}_{p'} \circ \bar{x}_p{}^{-1} = (x_{p'} \times \mathrm{I}_{\mathbb{R}^n}) \circ h' \circ h^{-1} \circ (x_p \times \mathrm{I}_{\mathbb{R}^n})^{-1}$$

é difeomorfismo.
Isso mostra que $\{(E|_{U_p}, \bar{x}_p)\}_{p \in V}$ é um atlas diferencial de E. Se $\boldsymbol{V}$ é segundo-contável, tem subatlas enumerável, e podemos tomar U_p nesse subatlas e $\{(E|_{U_p}, \bar{x}_p)\}_{p \in V}$ será um atlas enumerável, o que mostra que $\boldsymbol{E}$ é segundo contável. Se $\boldsymbol{V}$ é separado (T_2), note que pontos e, e' estão em um mesmo espaço $E|_p$ estão em uma mesma carta, enquanto que se esses pontos estão respectivamente em espaços $E|_p$ e $E|_{p'}$ distintos, podemos tomar U_p e $U_{p'}$ disjuntos, de modo que

$$E|_{U_p} \cap E|_{U_{p'}} = E|_{U_p \cap U_{p'}} = \emptyset,$$

que são vizinhanças abertas de e, e', o que mostra que E é separado.

A atlas maximal desse atlas define uma estrutura diferencial sobre E e temos uma variedade diferencial $\boldsymbol{E}$.

Para todo $(A, h) \in \mathcal{A}$, a função $h\colon E|_{U_p} \to U_p \times \mathbb{R}^n$ é um difeomorfismo, pois sua representação coordenada com respeito às cartas $(E|_{U_p}, \bar{x}_p)$ de E e $(U_p \times \mathbb{R}^n, x_p \times \mathrm{I}_{\mathbb{R}^n})$ de $U_p \times \mathbb{R}^n$ é a identidade

$$(x_p \times \mathrm{I}_{\mathbb{R}^n}) \circ h \circ \bar{x}_p{}^{-1} = \mathrm{I}_{x_p(U_p) \times \mathbb{R}^n} := x_p(U_p) \times \mathbb{R}^n \to x_p(U_p) \times \mathbb{R}^n,$$

um difeomorfismo.

A função $\mathrm{p}\colon E \to V$ é diferenciável, pois sua representação coordenada com respeito às cartas $(E|_{U_p}, \bar{x}_p)$ de E e (U_p, x_p) de V é a projeção

$$x_p \circ \mathrm{p} \circ \bar{x}_p{}^{-1} = \mathrm{p}_{x_p(U_p)}\colon x_p(U_p) \times \mathbb{R}^n \to x_p(U_p),$$

uma função diferenciável.

Mostremos que vale $\mathrm{p}_{U_p} \circ h = \mathrm{p}$. Seja $e \in E|_{U_p}$. Então existe $p' \in U_p$ tal que $e \in E|_{p'}$. Por hipótese, $\mathrm{p}(e) = p'$. Por outro lado, como $h(E|_p) = \{p\} \times \mathbb{R}^n$, existe $v \in \mathbb{R}^n$ tal que $h(e) = (p', v)$, portanto

$$\mathrm{p}_{U_p} \circ h(e) = p' = \mathrm{p}(e).$$

Por hipótese, $h \restriction_{E|_p} : E|_p \to \{p\} \times \mathbb{R}^n \simeq \mathbb{R}^n$ é um isomorfismo linear. Isso mostra que h é uma trivialização local e, como $(A)_{(A,h)\in\mathcal{A}}$ cobre V, segue que $(\boldsymbol{E}, \pi)$ é um fibrado vetorial de $\mathbb{R}^n$ sobre $\boldsymbol{V}$.

A unicidade segue da estrutura diferencial segue do fato de que as funções $(h)_{(A,h)\in\mathcal{A}}$ devem ser difeomorfismos, portanto qualquer outra estrutura deve conter as cartas construídas e portanto será igual a essa estrutura. ■

As funções de transição $T^h_{h'}$ satisfazem as seguintes propriedades.

⊢ **Proposição 22.7.** *Seja* $(\boldsymbol{E}, \mathrm{p})$ *um fibrado vetorial diferencial de* $\mathbb{R}^n$ *sobre uma variedade diferencial* $\boldsymbol{V}$.

1. *(Cociclicidade) Para todas trivializações locais* (A, h), (A', h') *e* (A'', h'') *de* $\boldsymbol{E}$ *e todo* $p \in A \cap A' \cap A''$,

$$T^h_{h''}|_p = T^{h'}_{h''}|_p \circ T^h_{h'}|_p.$$

2. *Para toda trivialização local* (A, h) *de* $\boldsymbol{E}$ *e todo* $p \in A$,

$$T^h_h|_p = \mathrm{I}_{\mathbb{R}^n};$$

3. *Para todas trivializações locais* (A, h) *e* (A', h') *de* $\boldsymbol{E}$ *e todo* $p \in A \cap A'$,

$$T^{h'}_h|_p = {T^h_{h'}|_p}^{-1};$$

□ *Demonstração.* 1. Como $h'' \circ h^{-1} = (h'' \circ h'^{-1}) \circ (h' \circ h^{-1})$, segue que, para todo $v \in \mathbb{R}^n$,

$$\begin{aligned}(p, T^h_{h''}|_p(v)) &= h'' \circ h^{-1}(p, v)\\ &= (h'' \circ h'^{-1}) \circ (h' \circ h^{-1})(p, v)\\ &= h'' \circ h'^{-1}(p, T^h_{h'}|_p(v))\\ &= (p, T^{h'}_{h''}|_p \circ T^h_{h'}|_p(v)),\end{aligned}$$

o que mostra que $T^h_{h''}|_p(v) = T^{h'}_{h''}|_p \circ T^h_{h'}|_p(v)$, logo $T^h_{h''}|_p = T^{h'}_{h''}|_p \circ T^h_{h'}|_p$.

2. Do item anterior segue que

$$T_h^h|_p = T_h^h|_p \circ T_h^h|_p,$$

o que mostra que

$$T_h^h|_p = T_h^h|_p \circ (T_h^h|_p)^{-1} = \mathrm{I}_{\mathbb{R}^n}.$$

■

Reciprocamente, a partir de uma família de mapas como acima, podemos construir o fibrado vetorial. No entanto, essa construção não é necessariamente única.

⊢ **Proposição 22.8** (Construção por Funções de Transição)**.** *Sejam* $\boldsymbol{V}$ *uma variedade diferencial,* $\mathcal{C} = (A_i)_{i\in I}$ *uma cobertura aberta de* V *e* $\{T_{i'}^i\}_{(i,i')\in I^2}$ *um conjunto de funções diferenciáveis*

$$\begin{aligned} T_{i'}^i \colon A_i \cap A_{i'} &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto T_{i'}^i|_p \end{aligned}$$

tais que

1. *(Cociclicidade) Para todos* $i, i', i'' \in I$ *e todo* $p \in A_i \cap A_{i'} \cap A_{i''}$,

$$T_{i''}^i|_p = T_{i''}^{i'}|_p \circ T_{i'}^i|_p.$$

Existem variedade diferencial $\boldsymbol{E}$ *e* $\mathrm{p}\colon E \to V$ *função diferenciável sobrejetiva tal que* $(\boldsymbol{E}, \mathrm{p})$ *é um fibrado vetorial diferencial de* $\mathbb{R}^n$ *sobre* $\boldsymbol{V}$.

□ *Demonstração.* A construção é simples de se acompanhar, mas tem várias etapas. Construiremos o espaço E, os espaços $E|_p$, a projeção $\mathrm{p}\colon E \to V$ e as funções h e usaremos a proposição 22.6 para mostrar a existência de estrutura diferencial para termos um fibrado vetorial.

Por simplicidade, para todos $i, i' \in I$, $p \in A_i \cap A_{i'}$ e $v \in \mathbb{R}^n$, denotaremos $T_{i'}^i|_p(v)$ por $T_{i'}^i|_p v$.

1. (Construir o espaço E) Definimos

$$E' := \bigsqcup_{i\in I}(A_i \times \mathbb{R}^n) = \{(i,(p,v)) \mid i \in I \ \text{ e } \ (p,v) \in A_i \times \mathbb{R}^n\}.$$

Por simplicidade, denotaremos $(i,(p,v)) \in E'$ por (i,p,v). Definimos a relação $\sim$ em E' por

$$(i,p,v) \sim (i',p',v') \quad \Longleftrightarrow \quad p = p' \ \text{ e } \ v' = T_{i'}^i|_p v.$$

Mostremos que $\sim$ é relação de equivalência.

1.1. (Reflexividade) Primeiro notemos que, pelo mesmo argumento de 22.7, para todo $i \in I$ vale que $T_i^i|_p = \mathrm{I}_{\mathbb{R}^n}$. Seja $(i,p,v) \in E$. Disso segue que $T_i^i|_p v = v$, portanto $(i,p,v) \sim (i,p,v)$.

1.2. (Simetria) Sejam (i,p,v), $(i',p',v') \in E'$ tais que $(i,p,v) \sim (i',p',v')$. Então $p = p'$ e $v' = T_{i'}^i|_p v$. Da cociclicidade de $\{T_{i'}^i\}_{(i,i')\in I^2}$ segue que

$$v = T_i^i|_p v = T_i^{i'}|_p \circ T_{i'}^i|_p v = T_i^{i'}|_p v'.$$

Como $p' = p$, segue que $(i',p',v') \sim (i,p,v)$.

1.3. (Transitividade) Sejam (i,p,v), (i',p',v'), $(i'',p'',v'') \in E'$ tais que $(i,p,v) \sim (i',p',v')$ e $(i',p',v') \sim (i'',p'',v'')$. Então $p = p'$, $p' = p''$, $v' = T_{i'}^i|_p v$ e $v'' = T_{i''}^{i'}|_p v'$. Da cociclicidade de $\{T_{i'}^i\}_{(i,i')\in I^2}$ segue que

$$T_{i''}^i|_p v = T_{i''}^{i'}|_p \circ T_{i'}^i|_p v = T_{i''}^{i'}|_p v' = v''.$$

Como $p'' = p$, segue que $(i,p,v) \sim (i'',p'',v'')$.

Definimos então o espaço

$$E := {E'}\big/{\sim} = \{[(i,p,v)] \mid i \in I \ \text{ e } \ (p,v) \in A_i \times \mathbb{R}^n\}.$$

Por simplicidade, denotaremos $[(i,p,v)] \in E$ por $[i,p,v]$.

2. (Construir os espaços $E|_p$) Para todo $p \in V$, definimos

$$E|_p := \{[i,q,v] \in E \mid q = p\}.$$

Notemos que, para todos $p, p' \in V$, se $E|_p \cap E|_{p'} \neq \emptyset$, então $p = p'$, pois se $[i,q,v] \in E|_p \cap E|_{p'}$, então $q = p$ e $q = p'$, logo $p = p'$. Além disso, como $(A_i)_{i\in I}$ é cobertura de V,

$$E = \{[i,p,v] \mid i \in I \ \text{ e } \ (p,v) \in A_i \times \mathbb{R}^n\} = \bigcup_{p\in V} E|_p,$$

o que implica que

$$E \simeq \bigsqcup_{p\in V} E|_p.$$

Para todo $[j,p,u] \in E|_p$ e todo $i \in I$ tal que $p \in A_i$, existe único $v \in \mathbb{R}^n$ tal que $(i,p,v) \in [j,p,u]$. Para mostrar a existência, basta tomar $v = T_i^j|_p u$ e temos que

$$(j,p,u) \sim (i,p,T_i^j|_p u) = (i,p,v).$$

Agora, suponha que existem $v, v' \in \mathbb{R}^n$ tais que $(i,p,v), (i,p,v') \in [j,p,u]$. Então $(i,p,v) \sim (i,p,v')$, portanto $v' = T_i^i|_p v = v$. Sendo assim, a partir de agora podemos sempre tomar um representante de $[j,p,u] \in E|_p$ com um $i \in I$ conveniente, contanto que $p \in A_i$.

Vamos dar estrutura de grupo para $E|_p$.

2.1. (+) A soma em $E|_p$ é a função

$$\begin{aligned} +\colon E|_p \times E|_p &\longrightarrow E|_p \\ ([i,p,v],[i,p,v']) &\longmapsto [i,p,v+v'], \end{aligned}$$

que está bem definida pois se $(i,p,v) \sim (j,p,u)$ e $(i,p,v') \sim (j,p,u')$, então $u = T^i_j|_p v$ e $u' = T^i_j|_p v'$, e segue da linearidade de $T^i_j|_p$ que

$$u + u' = T^i_j|_p v + T^i_j|_p v' = T^i_j|_p (v + v'),$$

o que implica que

$$[j,p,u+u'] = [j,p,T^i_j|_p(v+v')] = [i,p,v+v'].$$

A soma é associativa: para todos $[i,p,v],[i',p,v'],[i'',p,v''] \in E|_p$,

$$\begin{aligned} ([i,p,v]+[i,p,v']) + [i,p,v''] &= [i,p,v+v'] + [i,p,v''] \\ &= [i,p,v+v'+v''] \\ &= [i,p,v] + [i,p,v'+v''] \\ &= [i,p,v] + ([i,p,v']) + [i,p,v'']) \end{aligned}$$

A soma é comutativa: para todos $[i,p,v],[i,p,v'] \in E|_p$,

$$[i,p,v]+[i,p,v'] = [i,p,v+v'] = [i,p,v'+v] = [i,p,v'] + [i,p,v].$$

2.2. (0) A identidade em $E|_p$ é

$$0 := [i,p,0],$$

que está bem definida pois, para todo $i' \in I$ tal que $p \in A_{i'}$, segue da linearidade de $T^i_{i'}|_p$ que

$$(i,p,0) \sim (i',p,T^i_{i'}|_p 0) = (i',p,0).$$

Ela é identidade de +: para todo $[i,p,v] \in E|_p$,

$$0 + [i,p,v] = [i,p,0] + [i,p,v] = [i,p,0+v] = [i,p,v].$$

2.3. (−) A inversa em $E|_p$ é a função

$$\begin{aligned} -\colon E|_p &\longrightarrow E|_p \\ [i,p,v] &\longmapsto [i,p,-v], \end{aligned}$$

que está bem definida pois, se $(i,p,v) \sim (j,p,u)$, então $u = T^i_j|_p v$, e segue da linearidade de $T^i_j|_p$ que

$$-u = -T^i_j|_p v = T^i_j|_p(-v),$$

o que implica que

$$[j,p,-u] = [j,p,T^i_j|_p(-v)] = [i,p,-v].$$

Ela é inversa com respeito a + e 0: para todo $[(i,(p,v))] \in E|_p$,

$$-[i,p,v] + [i,p,v] = [i,p,-v] + [i,p,v] = [i,p,-v+v] = [i,p,0] = 0.$$

Isso mostra que $\boldsymbol{E|_p} := (E|_p, +, -, 0)$ é um grupo comutativo. Agora vamos construir a estrutura de espaço linear, ou seja, a ação de corpo de $\mathbb{R}$ sobre $E|_p$.

2.1. $(\cdot)$ A ação de $\mathbb{R}$ sobre $E|_p$ é a função

$$\begin{aligned} \cdot\colon \mathbb{R} \times E|_p &\longrightarrow E|_p \\ (c,[i,p,v]) &\longmapsto [i,p,cv], \end{aligned}$$

que está bem definida pois se $(i,p,v) \sim (i',p,v')$, então $v' = T^i_{i'}|_p v$ e segue da linearidade de $T^i_{i'}|_p$ que

$$(i',p,cv') = (i',p,cT^i_{i'}|_p v) = (i',p,T^i_{i'}|_p(cv)) \sim (i,p,cv).$$

Para mostrar que $\cdot$ é ação de corpo, devemos mostrar o seguinte.

2.1.1. Seja $c \in \mathbb{R}$. Para todos $[i,p,v],[i,p,v'] \in E|_p$,

$$\begin{aligned} c([i,p,v]+[i,p,v']) &= c[i,p,v+v'] \\ &= [i,p,c(v+v')] \\ &= [i,p,cv+cv'] \\ &= [i,p,cv]+[i,p,cv'] \\ &= c[i,p,v]+c[i,p,v']. \end{aligned}$$

2.1.2. Para todos $c,c' \in \mathbb{R}$ e $[i,p,v] \in E|_p$,

$$\begin{aligned} (c+c')[i,p,v] &= [i,p,(c+c')v] \\ &= [i,p,cv+c'v] \\ &= [i,p,cv]+[i,p,c'v] \\ &= c[i,p,v]+c'[i,p,v], \end{aligned}$$

$$\begin{aligned}(cc')[i,p,v] &= [i,p,(cc')v]\\ &= [i,p,c(c'v)]\\ &= c[i,p,c'v]\\ &= c(c'[i,p,v])\end{aligned}$$

e

$$1[i,p,v] = [i,p,1v] = [i,p,v].$$

Isso mostra que $(\boldsymbol{E|_p}, \cdot)$ é um espaço linear sobre $\mathbb{R}$. Esse espaço tem dimensão n, pois $([i,p,b_k])_{k\in[n]}$ é uma base de $E|_p$, em que $(b_k)_{k\in[n]}$ é uma base de $\mathbb{R}^n$. Isso ocorre pois, se $[i,p,v] \in E|_p$, então como $v \in \mathbb{R}^n$ e $(b_k)_{k\in[n]}$ gera $\mathbb{R}^n$, existe $(v^k)_{k\in[n]} \in \mathbb{R}^n$ tal que

$$v = \mathop{+}_{k\in[n]} v^k b_k,$$

logo

$$[i,p,v] = [i,p,\mathop{+}_{k\in[n]} v^k b_k] = \mathop{+}_{k\in[n]} v^k [i,p,b_k],$$

o que mostra que $([i,p,b_k])_{k\in[n]}$ gera $E|_p$; se $(c^k)_{k\in[n]} \in \mathbb{R}^n$ são não nulos, então $+_{k\in[n]} c^k b_k \neq 0$, pois $(b_k)_{k\in[n]}$ é linearmente independente, logo

$$\mathop{+}_{k\in[n]} c^k[(i,(p,b_k))] = [(i,(p,\mathop{+}_{k\in[n]} c^k b_k))] \neq 0,$$

o que mostra que $([i,p,b_k])_{k\in[n]}$ é linearmente independente, portanto base.

3. (Construir a projeção p) Definimos função

$$\begin{aligned}\mathrm{p}\colon E &\longrightarrow V\\ [i,p,v] &\longmapsto p,\end{aligned}$$

que está bem definida pois se $(i',p',v') \sim (i,p,v) \in E$, então $p = p'$. Claramente $\mathrm{p}(E|_p) = \{p\}$.

4. (Construir as trivializações locais h_i) Definimos, para todo $C \subseteq V$, $E|_C := \mathrm{p}^{-1}(C)$. Notemos que, para todo $p \in V$, $E|_p := E|_{\{p\}}$.
Para todo $i \in I$, definimos a função

$$\begin{aligned}h_i\colon E|_{A_i} &\longrightarrow A_i \times \mathbb{R}^n\\ [i,p,v] &\longmapsto (p,v).\end{aligned}$$

Note que, para todo $[j,p,u] \in E|_{A_i}$, temos $p \in A_i$, o que garante que existe representante $(i,p,v) \in [j,p,u]$. Essa definição não é independente de representante.

A função h_i tem inversa

$$\begin{aligned} {h_i}^{-1}\colon A_i \times \mathbb{R}^n &\longrightarrow E|_{A_i} \\ (p,v) &\longmapsto [i,p,v], \end{aligned}$$

pois, para todo $[i,p,v] \in E|_{A_i}$ e todo $(p,v) \in A_i \times \mathbb{R}^n$,

$${h_i}^{-1} \circ h_i[i,p,v] = {h_i}^{-1}(p,v) = [i,p,v]$$

e

$$h_i \circ {h_i}^{-1}(p,v) = h_i[i,p,v] = (p,v).$$

Mostremos que, para todo $i \in I$ e todo $p \in A_i$, a função $h_i \restriction_{E|_p}\colon E|_p \to \{p\} \times \mathbb{R}^n$ é isomorfismo linear. Primeiro notemos que o contradomínio de $h_i \restriction_{E|_p}$ é $\{p\} \times \mathbb{R}^n$, pois para todo $[i,p,v] \in E|_p$,

$$h_i[i,p,v] = (p,v) \in \{p\} \times \mathbb{R}^n.$$

Além disso, ${h_i}^{-1} \restriction_{\{p\}\times\mathbb{R}^n}\colon \{p\} \times \mathbb{R}^n \to E|_p$ é a inversa de $h_i \restriction_{E|_p}$. Notemos que o contradomínio de ${h_i}^{-1} \restriction_{\{p\}\times\mathbb{R}^n}$ é $E|_p$, pois

$${h_i}^{-1}(p,v) = [i,p,v] \in E|_p.$$

Claramente ${h_i}^{-1} \restriction_{\{p\}\times\mathbb{R}^n} = {h_i \restriction_{E|_p}}^{-1}$.
Para mostrar a linearidade, sejam $[i,p,v], [i,p,v'] \in E|_p$ e $c \in \mathbb{R}$. Então

$$\begin{aligned} h_i(c[i,p,v] + [i,p,v']) &= h_i([i,p,cv+v']) \\ &= (p, cv+v') \\ &= c(p,v) + (p,v') \\ &= ch_i[i,p,v] + h_i[i,p,v']. \end{aligned}$$

5. Agora notemos que, para todos $i, i' \in I$ e $p \in A_i \cap A_{i'}$, temos que $(i,p,v) \sim (i', p, T^i_{i'}|_p v)$, logo

$$h_{i'} \circ {h_i}^{-1}(p,v) = h_{i'}([i,p,v]) = h_{i'}([i',p,T^i_{i'}|_p v]) = (p, T^i_{i'}|_p(v)),$$

o que mostra que $T^i_{i'}$ é a função de transição de (A_i, h_i) para $(A_{i'}, h_{i'})$.

Pela proposição 22.6, essa construção mostra que E tem única estrutura diferencial tal que $(\boldsymbol{E}, \mathrm{p})$ é um fibrado vetorial de $\mathbb{R}^n$ sobre $\boldsymbol{V}$. ■

Embora a proposição 22.6 garanta que a estrutura diferencial sobre E é única, aqui não necessariamente garantimos a unicidade da construção do fibrado porque não garantimos que as construções de E, p ou h_i são únicas. Mantendo as construções de E e p, mas tomando $L \in \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ e definindo

$$\begin{aligned} h_i \colon E|_{A_i} &\longrightarrow A_i \times \mathbb{R}^n \\ [i, p, v] &\longmapsto (p, Lv) \end{aligned}$$

com inversa

$$\begin{aligned} {h_i}^{-1} \colon A_i \times \mathbb{R}^n &\longrightarrow E|_{A_i} \\ (p, v) &\longmapsto [i, p, L^{-1}v], \end{aligned}$$

ainda teríamos h_i uma bijeção que restrita às fibras é isomorfismo linear, e teríamos funções de transição

$$\begin{aligned} L \circ T^i_{i'} \circ L^{-1} \colon A_i \cap A_{i'} &\longrightarrow \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n) \\ p &\longmapsto L \circ T^i_{i'}|_p \circ L^{-1}, \end{aligned}$$

que é diferenciável porque $T^i_{i'}$ e L são, e $L \circ T^i_{i'}|_p \circ L^{-1} \in \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ pois $T^i_{i'}|_p \in \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ e $L \in \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$.

▷ **Exercício 22.1.** *Verifique que a construção da proposição anterior com essa h_i que usa $L \in \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ também gera uma estrutura de fibrado vetorial* $(\boldsymbol{E}, \mathrm{p})$.

Essa estrutura alternativa é de fato equivalente à que construímos, num sentido de equivalência de fibrados que estudaremos adiante.

22.2.2 Seções locais e globais

:⊢ **Definição 22.4.** Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$. Uma *seção (global)* de $\boldsymbol{E}$ é uma função $s \colon V \to E$ tal que

$$\mathrm{p} \circ s = \mathrm{I}_V.$$

Isso é equivalente a dizer que

$$\begin{aligned} s \colon V &\longrightarrow E \\ p &\longmapsto s|_p \end{aligned}$$

satisfaz, para todo $p \in V$, $s|_p \in E|_p$. O conjunto das seções diferenciáveis de $\boldsymbol{E}$ é denotado $\Gamma(E)$.

Seja $A \subseteq V$. Uma *seção local* de $\boldsymbol{E}$ sobre A é uma função $s \colon A \to E$ tal que $\mathrm{p} \circ s = \mathrm{I}_A$. O conjunto das seções locais diferenciáveis de $\boldsymbol{E}$ sobre A é denotado $\Gamma(E)|_A$.

:⊢ **Definição 22.5.** Sejam $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$ e $s \in \Gamma(E)$ uma seção. O *suporte* de s é o conjunto

$$\operatorname{supp}(s) := \overline{\{p \in V \mid s|_p \neq 0\}}.$$

O conjunto $\Gamma(E)$ é um espaço linear sobre $\mathbb{R}$ com a adição e o multiplicação por escalar induzidos pontualmente pela adição, zero, inversa da adição e multiplicação por escalar de $E|_p$. Além dessa operações, podemos multiplicar uma seção por uma função escalar diferenciável pontualmente e dar a $\Gamma(E)$ uma estrutura de módulo sobre $\mathscr{C}^\infty(V)$.

:⊢ **Definição 22.6.** Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$. A *adição* em $\Gamma(E)$ é

$$\begin{aligned} +\colon \Gamma(E) \times \Gamma(E) &\longrightarrow \Gamma(E) \\ (s, s') &\longmapsto s + s' \colon V \longrightarrow E \\ &\qquad\qquad p \longmapsto s|_p + s'|_p, \end{aligned}$$

a *seção zero* em $\Gamma(E)$ é a seção

$$\begin{aligned} 0\colon V &\longrightarrow E \\ p &\longmapsto 0_{E|_p}, \end{aligned}$$

a *inversa aditiva* de $\Gamma(E)$ é

$$\begin{aligned} -\colon \Gamma(E) &\longrightarrow \Gamma(E) \\ s &\longmapsto -s \colon V \longrightarrow E \\ &\qquad\qquad p \longmapsto -s|_p \end{aligned}$$

e o *produto por escalar* em $\Gamma(E)$ é

$$\begin{aligned} \cdot\colon \mathbb{R} \times \Gamma(E) &\longrightarrow \Gamma(E) \\ (c, s) &\longmapsto cs \colon V \longrightarrow E \\ &\qquad\qquad p \longmapsto cs|_p. \end{aligned}$$

A *multiplicação por função*

$$\begin{aligned} \cdot\colon \mathscr{C}^\infty(V) \times \Gamma(E) &\longrightarrow \Gamma(E) \\ (f, s) &\longmapsto fs \colon V \longrightarrow E \\ &\qquad\qquad p \longmapsto f(p)s|_p. \end{aligned}$$

Isso é uma ação do anel $\mathscr{C}^\infty(V)$ que dá ao espaço $\Gamma(E)$ uma estrutura de módulo.

▷ **Exercício 22.2.** *Seja* $(\boldsymbol{E}, \mathrm{p})$ *um fibrado vetorial diferencial de* $\mathbb{R}^n$ *sobre uma variedade diferencial* $\boldsymbol{V}$.

1. *O espaço de seções* $\Gamma(E)$ *é um espaço linear sobre* $\mathbb{R}$ *com respeito à adição e multiplicação por escalar pontuais.*
2. *O espaço de seções* $\Gamma(E)$ *é um módulo sobre* $\mathscr{C}^\infty(V)$ *com respeito à adição e multiplicação por função pontuais.*

22.2.3 Grupo estrutural

:⊢ **Definição 22.7.** Sejam $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$ e $\boldsymbol{G} \subseteq \overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)$ um grupo diferencial. Uma G-estrutura sobre $(\boldsymbol{E}, \mathrm{p})$ é uma coleção de funções de transição $\{T^i_{i'}\}_{(i,i')\in I^2}$ cujos domínios cobrem V tais que, para todos $i, i' \in I$ e $p \in A_i \cap A_{i'}$, $T^i_{i'}|_p \in G$. Um *grupo estrutural* de $\boldsymbol{E}$ é um grupo diferencial $\boldsymbol{G}$ para o qual existe uma G-estrutura.

► **Exemplo 22.2.** *Seja* $(\boldsymbol{E}, \mathrm{p})$ *um fibrado vetorial diferencial de* $\mathbb{R}^n$ *sobre uma variedade diferencial* $\boldsymbol{V}$.

1. *Se* $G = \{\mathrm{I}\}$ *é um grupo estrutural de* $\boldsymbol{E}$*, então* $\boldsymbol{E}$ *é trivial;*
2. *Uma* $\overleftrightarrow{\mathscr{L}}(\mathbb{R}^n)^+$*-estrutura sobre* $\mathrm{T}V$ *é equivalente a uma orientação de* $\boldsymbol{V}$*;*
3. *Uma* $\overleftrightarrow{\mathscr{L}}_{\|\|}(\mathbb{R}^n)$*-estrutura (também denotado* $O(n)$ *ou* $O_n(\mathbb{R})$*) sobre* $\boldsymbol{E}$ *é equivalente a um produto interno em cada fibra* $E|_p$*, dado por*

$$\langle v, v' \rangle |_p := v'^* T^i_{i'}|_p v.$$

22.2.4 Referenciais locais e globais

:⊢ **Definição 22.8.** Sejam $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$ e $A \subseteq V$ um aberto. Uma k-tupla de seções locais sobre A *linearmente independente* é uma k-tupla $(s_i)_{i\in[k]}$ de seções locais $s_i \colon A \to E$ de $\boldsymbol{E}$ sobre A tais que, para todo $p \in A$, $(s_i|_p)_{i\in[k]}$ é linearmente independente em $E|_p$.

Similarmente, uma k-tupla de seções locais sobre A que *gera* $\boldsymbol{E}$ é uma k-tupla $(s_i)_{i\in[k]}$ de seções locais $s_i \colon A \to E$ de $\boldsymbol{E}$ sobre A tais que, para todo $p \in A$, $(s_i|_p)_{i\in[k]}$ gera $E|_p$.

Um *referencial local* de $\boldsymbol{E}$ sobre A é uma k-tupla de seções locais de $\boldsymbol{E}$ sobre A linearmente independente que *gera* $\boldsymbol{E}$. Um *referencial (global)* de $\boldsymbol{E}$ é um referencial local de $\boldsymbol{E}$ sobre V.

Lembremos que $(e_i)_{i\in[n]}$ é a base canônica de $\mathbb{R}^n$.

$\vdash$ **Proposição 22.9.** *Sejam* $(\boldsymbol{E}, \mathrm{p})$ *um fibrado vetorial diferencial de* $\mathbb{R}^n$ *sobre uma variedade diferencial* $\boldsymbol{V}$ *e* (A, h) *uma trivialização local. A funções*

$$\begin{aligned} s_i \colon A &\longrightarrow E \\ p &\longmapsto h^{-1}(p, e_i). \end{aligned}$$

são seções locais de $\boldsymbol{E}$ *e* $(s_i)_{i\in[n]}$ *é um referencial local de* $\boldsymbol{E}$ *sobre* A.

$\square$ *Demonstração.* As funções s_i são diferenciáveis porque h é difeomorfismo e de $\mathrm{p}_A \circ h = \mathrm{p}$ segue que, para todo $p \in A$,

$$\mathrm{p} \circ s_i(p) = \mathrm{p} \circ h^{-1}(p, e_i) = \mathrm{p}_A(p, e_i) = p,$$

logo $\mathrm{p} \circ s_i = \mathrm{I}_A$, o que mostra que s_i é uma seção local.

Para vermos que $(s_i)_{i\in[n]}$ é um referencial local, notemos que $h \colon E|_p \to \{p\} \times \mathbb{R}^n$ é um isomorfismo e $h(s_i|_p) = (p, e_i)$, então leva $(s_i)_{i\in[n]}$ para a base canônica de $\{p\} \times \mathbb{R}^n$. $\blacksquare$

$:\vdash$ **Definição 22.9.** Sejam $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{V}$ e (A, h) uma trivialização local. O *referencial local associado a* h é o referencial local $(s_i)_{i\in[n]}$ de $\boldsymbol{E}$ sobre A, em que

$$\begin{aligned} s_i \colon A &\longrightarrow E \\ p &\longmapsto h^{-1}(p, e_i). \end{aligned}$$

$\triangleright$ **Exercício 22.3.** *Todo referencial local está associado a uma trivialização local.*

22.3 Fibrados principais

Lembremos o que é uma ação à direita de um grupo. Sejam X um conjunto e $\boldsymbol{G}$ um grupo. Uma ação à direita $X \curvearrowleft \boldsymbol{G}$ de G em X é uma função

$$\begin{aligned} A \colon X \times G &\longrightarrow X \\ (x, g) &\longmapsto x \centerdot g \end{aligned}$$

tal que

1. (Identidade) Para todo $x \in X$,

$$x \centerdot 1 = x;$$

2. (Compatibilidade) Para todos $g, g' \in G$ e $x \in X$,

$$x \centerdot (gg') = (x \centerdot g) \centerdot g';$$

Além disso, consideraremos também a seguinte propriedade da ação.

1. (Livre) Para todo $g \in G$, se existe $x \in X$ tal que $x \centerdot g = x$, então $g = 1$;

:⊢ **Definição 22.10.** Sejam $\boldsymbol{V}$ uma variedade diferencial e $\boldsymbol{G}$ um grupo diferencial. Um $\boldsymbol{G}$-fibrado principal diferencial sobre $\boldsymbol{V}$ é uma tripla $(\boldsymbol{E}, \pi, \centerdot)$ em que $\boldsymbol{P}$ é uma variedade diferencial, $\centerdot\colon E \times G \to E$ é uma ação à direita diferenciável e livre, $V \simeq E/G$ e, para todo $x \in V$, existem vizinhança $A \subseteq V$ de x e difeomorfismo $\phi\colon E|_A := \mathrm{p}^{-1}(A) \to A \times G$ tais que

1. $\mathrm{p}_A \circ \phi = \mathrm{p}$ (o diagrama comuta).

$$\begin{tikzcd}
E|_A \arrow[r, "\phi"] \arrow[d, "\mathrm{p}"'] & A \times G \arrow[dl, "\mathrm{p}_A"] \\
A &
\end{tikzcd}$$

2. (G-equivariância) Se $\phi(e) = (p, h)$, então $\phi(e \centerdot g) = (p, hg)$ — ou, equivalentemente, $\mathrm{p}_G \circ \phi \circ \centerdot g = \centerdot g \circ \mathrm{p}_G \circ \phi$ (o diagrama comuta).

$$\begin{tikzcd}
E|_A \arrow[r, "\centerdot g"] \arrow[d, "\mathrm{p}_G \circ \phi"'] & E|_A \arrow[d, "\mathrm{p}_G \circ \phi"] \\
G \arrow[r, "\centerdot g"'] & G
\end{tikzcd}$$

$$\phi^{-1}(p, h) \centerdot g = \phi^{-1}(p, hg)$$
$$(\centerdot g) \circ \phi^{-1}(p, h) = \phi^{-1}(p, hg)$$
$$(\centerdot g) \circ \phi^{-1}(p, h) = \phi^{-1}(p, (\centerdot g) \circ \mathrm{p}_G(e))$$

⊢ **Proposição 22.10.** *Seja $(\boldsymbol{E}, \pi, \centerdot)$ um fibrado principal diferencial de um grupo diferencial $\boldsymbol{G}$ sobre uma variedade diferencial $\boldsymbol{V}$. Para todas trivializações locais (A, ϕ) e (A', ϕ') de $\boldsymbol{E}$, a função de transição delas é a multiplicação à esquerda pela valor dela na identidade do grupo:*

$$\begin{aligned} T^{\phi}_{\phi'}|_p = E_{T^{\phi}_{\phi'}|_p 1}\colon G &\longrightarrow G \\ g &\longmapsto (T^{\phi}_{\phi'}|_p 1) g. \end{aligned}$$

□ *Demonstração.* Sejam $g, h \in G$ e $e := \phi^{-1}(p, h)$. Da definição de função de transição,

$$\phi'(e) = \phi' \circ \phi^{-1}(p, h) = (p, T^{\phi}_{\phi'}|_p(h))$$

e, da equivariância,

$$(p, T^{\phi}_{\phi'}|_p(hg)) = \phi' \circ \phi^{-1}(p, hg) = \phi'(e \centerdot g) = (p, (T^{\phi}_{\phi'}|_p(h))g),$$

o que implica que

$$T^{\phi}_{\phi'}|_p(hg) = (T^{\phi}_{\phi'}|_p h)g.$$

Tomando $h = 1$, segue que

$$T^{\phi}_{\phi'}|_p(g) = (T^{\phi}_{\phi'}|_p 1)g.$$

■

Por simplicidade, identificaremos $T^{\phi}_{\phi'}|_p$ com $T^{\phi}_{\phi'}|_p 1$, de modo a termos $T^{\phi}_{\phi'}|_p \in G$. As funções de transição, que são da forma

$$T^{\phi}_{\phi'} \colon A \cap A' \to \overleftrightarrow{\mathscr{C}}^{\infty}(G),$$

sob essa identificação podem ser entendidas como funções da forma

$$T^{\phi}_{\phi'} \colon A \cap A' \to G.$$

22.4 Conexões em fibrados vetoriais

22.4.1 Espaços verticais e horizontais

Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{B}$. Tomemos carta (A, x) de $\boldsymbol{B}$ e trivialização local (A, ϕ) de $\boldsymbol{E}$. Elas são funções

$$\begin{aligned} x \colon A &\longrightarrow x(A) \subseteq \mathbb{R}^d \\ b &\longmapsto x(b) \end{aligned}$$

e

$$\begin{aligned} \phi \colon E|_A &\longrightarrow A \times \mathbb{R}^n \\ e &\longmapsto (b, v). \end{aligned}$$

A função

$$\begin{aligned} \bar{x} = (x \times \mathrm{I}_{\mathbb{R}^n}) \circ \phi \colon E|_A &\longrightarrow x(A) \times \mathbb{R}^n \subseteq \mathbb{R}^d \times \mathbb{R}^n \\ e &\longmapsto (x(b), v), \end{aligned}$$

em que $\phi(e) = (b, v)$, é uma carta de $\boldsymbol{E}$. O referencial coordenado de $\mathrm{T}E|_A$ é

$$\left(\left.\frac{\partial}{\partial \bar{x}^i}\right|_e\right)_{i\in[d+n]}.$$

Mas notemos que, para $i \in [d]$,

$$\bar{x}^i = \mathrm{p}_i \circ \bar{x} = \mathrm{p}^i \circ x = x^i$$

e, para $j \in d + [n]$,

$$\bar{x}^i = \mathrm{p}_i \circ \bar{x} = \mathrm{p}^i \circ \phi = \phi^i,$$

portanto

$$\left(\left.\frac{\partial}{\partial \bar{x}^i}\right|_e\right)_{i\in[d+n]} = \left(\left.\frac{\partial}{\partial x^i}\right|_e, \left.\frac{\partial}{\partial \phi^j}\right|_e\right)_{i\in[d],j\in d+[n]}.$$

Consideremos a função diferenciável

$$\mathrm{D}\mathrm{p}|_e \colon \mathrm{T}E|_e \to \mathrm{T}B|_b.$$

22.4.2 Conexão/derivada covariante

Definamos

$$\Omega^k_B(E) = \Omega^k(B) \otimes \Gamma(E).$$

Em particular,

$$\Omega^0_B(E) = \mathscr{C}^\infty(B) \otimes \Gamma(E) = \Gamma(E)$$

e

$$\Omega^1_B(E) = \Omega^1(B) \otimes \Gamma(E).$$

:⊢ **Definição 22.11.** Seja $(\boldsymbol{E}, \mathrm{p})$ um fibrado vetorial diferencial de $\mathbb{R}^n$ sobre uma variedade diferencial $\boldsymbol{B}$. Uma *derivada covariante* em $\boldsymbol{E}$ é uma função linear sobre $\mathbb{R}$

$$\nabla \colon \Omega^0_B(E) \to \Omega^1_B(E).$$

tal que, para todas $f \in \mathscr{C}^\infty(B)$ e $s \in \Gamma(E)$,

$$\nabla(fs) = \mathrm{d}f \otimes s + f\nabla s.$$

Equivalentemente, ∇ é uma função

$$\nabla \colon \Gamma(E) \to \Omega^1(B) \otimes \Gamma(E)$$

que toma seções e retorna 1-formas a valores em seções.

Estendemos a derivada covariante $\nabla\colon \Omega^0_B(E) \to \Omega^1_B(E)$ para uma função[1] linear sobre $\mathbb{R}$

$$\nabla\colon \Omega^k_B(E) \to \Omega^{k+1}_B(E)$$

para todo k tal que, para todas $\omega \in \Omega^k(B)$ e $s \in \Gamma(E)$,

$$\nabla(\omega \otimes s) = \mathrm{d}\omega \otimes s + (-1)^k \omega \otimes \nabla s.$$

[1]Geralmente essa derivada estendida é denotada d_∇ e chamada de derivada exterior covariante, mas aqui adotaremos a simplicidade de manter ∇ para todos os casos.

A derivada covariante é a função $\mathbb{R}$-linear

$$\nabla\colon \Omega_B^k(E) \to \Omega_B^{k+1}(E).$$

Localmente, temos

$$\nabla = \mathrm{d} + A.$$

A *curvatura* é a função $\mathscr{C}^\infty(B)$-linear

$$\nabla^2\colon \Omega_B^k(E) \to \Omega_B^{k+2}(E).$$

Localmente, temos

$$\nabla^2 = \mathrm{d}A + A \wedge A.$$

Capítulo 23

Grupos diferenciais

23.1 Definições básicas

Lembremos que denotamos um grupo pela notação multiplicativa: o grupo é denotado $(G, \times, \diagup, 1)$, sendo

$$\begin{aligned} \times \colon G^2 &\longrightarrow G \\ (g, g') &\longmapsto g \times g = gg' \end{aligned}$$

a operação binária, $1 \in G$ a identidade e

$$\begin{aligned} \diagup \colon G &\longrightarrow G \\ g &\longmapsto \diagup g = g^{-1} \end{aligned}$$

a inversa.

No caso de um grupo comutativo, às vezes usa-se a notação aditiva $(G, +, -, 0)$, sendo

$$\begin{aligned} + \colon G^2 &\longrightarrow G \\ (g, g') &\longmapsto g + g' \end{aligned}$$

a operação binária, $0 \in G$ a identidade e

$$\begin{aligned} - \colon G &\longrightarrow G \\ g &\longmapsto -g \end{aligned}$$

a inversa. Para a teoria geral, usaremos a notação multiplicativa.

Denotamos as translações à esquerda e à direita e a conjugação por um elemento $g \in G$ respectivamente por

$$\begin{aligned} E_g \colon G &\longrightarrow G \\ h &\longmapsto gh \end{aligned}$$

$$\begin{aligned} D_g\colon G &\longrightarrow G \\ h &\longmapsto hg \end{aligned}$$

$$\begin{aligned} C_g\colon G &\longrightarrow G \\ h &\longmapsto ghg^{-1}. \end{aligned}$$

:⊢ **Definição 23.1.** Um *grupo diferencial*[1] é uma lista $\boldsymbol{G} = ((G, \mathcal{A}), \times, \diagup, 1)$ em que $(G, \mathcal{A})$ é uma variedade diferencial, $(G, \times, \diagup, 1)$ é um grupo e

$$\times\colon G^2 \longrightarrow G \qquad \text{e} \qquad \diagup\colon G \longrightarrow G$$

são diferenciáveis[2].

Diferentemente do caso de grupos topológicos, em que é necessário exigir que a operação binária e a inversa do grupo sejam contínuas, para que um grupo com estrutura de variedade seja um grupo diferencial basta que sua operação binária seja diferenciável, não é necessário assumir a diferenciabilidade da inversa $\diagup\colon G \longrightarrow G$. Para demonstrar isso, primeiro deve-se notar que as translações à esquerda e à direita e a conjugação são difeomorfismos se a operação binária é diferenciável. A seguir, a diferencial parcial de $\times\colon G \times G \longrightarrow G$ com relação à entrada 0 é denotada D_0 e a diferencial parcial com relação à entrada 1 é denotada D_1.

▷ **Exercício 23.1.** *Seja $\boldsymbol{G} = ((G, \mathcal{A}), \times, \diagup, 1)$ uma lista em que $(G, \mathcal{A})$ é uma variedade diferencial, $(G, \times, \diagup, 1)$ é um grupo e a operação binária $\times\colon G^2 \longrightarrow G$ é diferenciável. Então, para todo $g \in G$, as translações à esquerda e à direita E_g, D_g e a conjugação C_g são difeomorfismos e*

1. $\mathrm{D}_0(\times)|_{(g,g')} = \mathrm{D}(D_{g'})|_g$;
2. $\mathrm{D}_1(\times)|_{(g,g')} = \mathrm{D}(E_g)|_{g'}$;

⊢ **Proposição 23.1.** *Seja $\boldsymbol{G} = ((G, \mathcal{A}), \times, \diagup, 1)$ uma lista em que $(G, \mathcal{A})$ é uma variedade diferencial, $(G, \times, \diagup, 1)$ é um grupo e a operação binária $\times\colon G^2 \longrightarrow G$ é diferenciável. Então $\boldsymbol{G}$ é um grupo diferenciável, a inversa $\diagup\colon G \longrightarrow G$ é um difeomorfismo e, para todo $g \in G$,*

$$\mathrm{D}(\diagup)|_g = -\mathrm{D}(E_{g^{-1}})|_1 \circ \mathrm{D}(D_{g^{-1}})|_g.$$

[1] Grupos diferenciais são comumente chamados de grupos de Lie, em homenagem ao matemático Sophus Lie.

[2] Assumiremos aqui, como de costume, que o grau de diferenciabilidade é $\mathscr{C}^\infty$, mas na maioria dos casos só é necessário $\mathscr{C}^2$.

□ *Demonstração.* Para todo $(g, g') \in G \times G$, a diferencial parcial de $\times$ relativa à entrada 1 é

$$\mathrm{D}_1(\times)|_{(g,g')} = \mathrm{D}(E_g)|_{g'}.$$

Como E_g é difeomorfismo, $\mathrm{D}(E_g)|_{g'}$ é bijeção, portanto sobrejetiva. Como $g \times g^{-1} = 1$, pelo teorema da função implícita existem vizinhanças abertas $A \subseteq G$ de g e $A' \subseteq G$ de g^{-1} e função diferenciável $f\colon A \longrightarrow A'$ tal que $g \times f(g) = 1$. Da unicidade da inversa segue que $f = \diagup$, portanto que $\diagup$ é diferenciável em g, logo diferenciável. Como $\diagup$ é sua própria função inversa ($\diagup \circ \diagup = \mathrm{I}$), segue que é difeomorfismo.

O teorema da função implícita implica ainda que

$$\mathrm{D}(\diagup)|_g = -\mathrm{D}_1(\times)|_{(g,g^{-1})} \circ \mathrm{D}_0(\times)|_{(g,g^{-1})}.$$

Como $(\mathrm{D}(E_g)|_{g^{-1}})^{-1} = \mathrm{D}(g^{-1}\times)|_1$, segue que

$$\begin{aligned}\mathrm{D}(\diagup)|_g &= -\mathrm{D}_1(\times)|_{(g,g^{-1})} \circ \mathrm{D}_0(\times)|_{(g,g^{-1})} \\ &= -(\mathrm{D}(E_g)|_{g^{-1}})^{-1} \circ \mathrm{D}(D_{g^{-1}})|_g \\ &= -\mathrm{D}(g^{-1}\times)|_1 \circ \mathrm{D}(D_{g^{-1}})|_g.\end{aligned}$$

■

Em particular, isso implica que $\mathrm{D}(\diagup)|_1 = -\mathrm{I}|_1\colon \mathrm{T}G|_1 \longrightarrow \mathrm{T}G|_1$, pois $1^{-1} = 1$ e $E_1 = D_1 = \mathrm{I}\colon G \longrightarrow G$, logo

$$\mathrm{D}(\diagup)|_1 = -\mathrm{D}(E_{1^{-1}})|_1 \circ \mathrm{D}(D_{1^{-1}})|_1 = -\mathrm{I}|_1 \circ \mathrm{I}| - 1 = -\mathrm{I}|_1.$$

Denotamos o fibrado tangente de G por $\mathrm{T}G$ e a projeção canônica por $\mathrm{p}\colon \mathrm{T}G \longrightarrow G$.

⊢ **Proposição 23.2.** *Seja* $\boldsymbol{G}$ *um grupo diferencial. Os fibrados tangente* $\mathrm{T}G$ *e cotangente* T^*G *de* G *são triviais.*

□ *Demonstração.* • ($\mathrm{T}G$) Basta mostrar que a função

$$\begin{aligned}\Phi\colon G \times \mathrm{T}G|_1 &\longrightarrow \mathrm{T}G \\ (g, v) &\longmapsto \mathrm{D}(E_g)|_1 v\end{aligned}$$

é um difeomorfismo. Como, para todo $g \in G$, $E_g\colon G \longrightarrow G$ é difeomorfismo, sua diferencial $\mathrm{D}(E_g)|_1\colon \mathrm{T}G|_1 \longrightarrow \mathrm{T}G|_g$ é um isomorfismo linear, portanto Φ é bijeção. Notemos que $\Phi(g, v) = \mathrm{D}_1(\times)|_{(g,1)}v$, portanto é diferenciável. A inversa de Φ é dada por

$$\begin{aligned}\Phi^{-1}\colon \mathrm{T}G &\longrightarrow G \times \mathrm{T}G|_1 \\ v &\longmapsto (\mathrm{p}(v), \mathrm{D}(E_g)|_{\mathrm{p}(v)^{-1}}v)\end{aligned}$$

que também é diferenciável, portanto Φ é um difeomorfismo.

- (T^*G) Analogamente, basta mostrar que a função

$$\begin{aligned}\Psi\colon G \times \mathrm{T}^*G|_1 &\longrightarrow \mathrm{T}G\\ (g,\alpha) &\longmapsto \mathrm{D}(E_g)|_1^*\alpha\end{aligned}$$

é um difeomorfismo.

■

23.1.1 Álgebra infinitesimal de um grupo diferencial

23.1.1.1 Campos vetoriais invariantes

:⊢ **Definição 23.2.** Seja $\boldsymbol{G}$ um grupo diferencial. Um *campo vetorial invariante à esquerda* em $\boldsymbol{G}$ é um campo vetorial $v \in \mathfrak{T}^1((G,\mathcal{A}))$ tal que, para todo $g \in G$,

$$\mathrm{D}(E_g)_\star v = v,$$

ou seja, para todo $g, g' \in G$,

$$\mathrm{D}(E_g)|_{g'}(v|_{g'}) = v|_{gg'}.$$

O conjunto desses campos é denotado $\mathfrak{T}^1_E(\boldsymbol{G})$. Um *campo vetorial invariante à direita* em $\boldsymbol{G}$ é um campo vetorial $v \in \mathfrak{T}^1((G,\mathcal{A}))$ tal que, para todo $g \in G$,

$$\mathrm{D}(D_g)_\star v = v,$$

ou seja, para todo $g, g' \in G$,

$$\mathrm{D}(D_g)|_{g'}(v|_{g'}) = v|_{g'g}.$$

O conjunto desses campos é denotado $\mathfrak{T}^1_D(\boldsymbol{G})$.

Os espaços dos campos invariantes são subespaços lineares de $\mathfrak{T}((G,\mathcal{A}))$

Para definir um campo invariante, basta definirmos seu valor $v \in \mathrm{T}G|_1$ em $1 \in G$, pois todos os outros valores podem ser obtidos empurrando v pela diferencial da translação. Para campos invariante à esquerda,

$$v|_p := \mathrm{D}(E_g)|_1(v),$$

e para campos invariantes à direita,

$$v|_g = \mathrm{D}(D_g)|_1(v).$$

Denotaremos esses campos por v_E e v_D, respectivamente. Eles induzem isomorfismos de espaços lineares

$$\begin{aligned}\Phi_E\colon \mathrm{T}G|_1 &\longrightarrow \mathfrak{T}^1_E(G)\\ v &\longmapsto v_E\colon G \longrightarrow \mathrm{T}G\\ &\qquad\qquad g \longmapsto v_E|_g := \mathrm{D}(E_g)|_1(v)\end{aligned}$$

$$\begin{aligned}\Phi_D\colon \mathrm{T}G|_1 &\longrightarrow \mathfrak{T}^1_D(G)\\ v &\longmapsto v_D\colon G \longrightarrow \mathrm{T}G\\ &\qquad\qquad g \longmapsto v_E|_g := \mathrm{D}(D_g)|_1(v).\end{aligned}$$

23.1.1.2 Álgebra infinitesimal

Lembremos que o espaço de campos vetoriais $\mathfrak{T}^1(G,\mathcal{A})$ da variedade $(G,\mathcal{A})$ é uma álgebra derivativa[3] sobre $\mathbb{R}$ cujo produto é o colchete de campos vetoriais $[\cdot,\cdot]$. Os espaços de campos vetoriais invariantes à esquerda e à direita são subálgebras da álgebra de campos vetoriais, pois são subespaços lineares e o colchete é fechado. Esses colchetes, dos campos invariantes à esquerda e à direita, induzem colchetes em $\mathrm{T}G|_1$ pelos isomorfismos Φ_E e Φ_D anteriormente enunciados.

⊢ **Proposição 23.3.** *Seja $\boldsymbol{G}$ um grupo diferencial.*

1. $(\mathfrak{T}^1_E(\boldsymbol{G}),+,-,0,[\cdot,\cdot],\cdot)$ *é subálgebra de* $(\mathfrak{T}^1(G,\mathcal{A}),+,-,0,[\cdot,\cdot],\cdot)$ *(sobre* $\mathbb{R}$*);*
2. $(\mathfrak{T}^1_D(\boldsymbol{G}),+,-,0,[\cdot,\cdot],\cdot)$ *é subálgebra de* $(\mathfrak{T}^1(G,\mathcal{A}),+,-,0,[\cdot,\cdot],\cdot)$ *(sobre* $\mathbb{R}$*).*

□ *Demonstração.* A demonstração de que são subespaços lineares é direta. Para vermos que o colchete é fechado, basta notarmos que, como $E_g\colon G \longrightarrow G$ é difeomorfismo, segue que, para todos campos vetoriais $v, v' \in \mathfrak{T}^1_E(\boldsymbol{G})$,

$$\mathrm{D}(E_g)\,[v,v'] = [\mathrm{D}(E_g)_\star v, \mathrm{D}(E_g)_\star v'] = [v,v']\,,$$

portanto se $[v,v'] \in \mathfrak{T}^1_E(\boldsymbol{G})$. A demonstração é análoga para campos invariantes à direita. ■

:⊢ **Definição 23.3.** Seja $\boldsymbol{G}$ um grupo diferencial. O *colchete* (à esquerda) em $\mathrm{T}G|_1$ é

$$\begin{aligned}[\cdot,\cdot]_E\colon \mathrm{T}G|_1 \times \mathrm{T}G|_1 &\longrightarrow \mathrm{T}G|_1\\ (v,v') &\longmapsto [v_E, v'_E]\,|_1.\end{aligned}$$

O *colchete* (à direita) em $\mathrm{T}G|_1$ é

$$\begin{aligned}[\cdot,\cdot]_D\colon \mathrm{T}G|_1 \times \mathrm{T}G|_1 &\longrightarrow \mathrm{T}G|_1\\ (v,v') &\longmapsto [v_D, v'_D]\,|_1.\end{aligned}$$

[3] Álgebra de Lie.

:⊢ **Definição 23.4.** Seja $\boldsymbol{G}$ um grupo diferencial. A *álgebra infinitesimal* de $\boldsymbol{G}$ é a álgebra derivativa

$$\mathbf{T}\boldsymbol{G}|_{\mathbf{1}} := (\mathrm{T}G|_1, [\cdot, \cdot]_E).$$

Parte 4

Dinâmica

Capítulo 24

Sistemas dinâmicos

24.1 Definições gerais

24.1.1 Sistemas dinâmicos discretos e contínuos

:⊢ **Definição 24.1.** Um *sistema dinâmico discreto* é uma dupla (X, f) em que

1. X é um conjunto, o *espaço de fase* do sistema;
2. $f\colon X \longrightarrow X$ é uma função, a *dinâmica* do sistema.

Um sistema dinâmico discreto *invertível* (ou *revertível*) é um sistema dinâmico (X, f) cuja dinâmica f é invertível.

Além de sistemas dinâmicos discretos, podemos também considerar sistemas dinâmicos contínuos, que são definidos a partir de um semifluxo f no espaço de fase X da forma

$$\begin{aligned} f\colon \mathbb{R}_{\geq 0} \times X &\longrightarrow X \\ (t, x) &\longmapsto f^t(x), \end{aligned}$$

que também pode ser entendido como uma função $f\colon \mathbb{R}_{\geq 0} \longrightarrow \mathcal{F}(X, X)$ que é uma ação do monoide aditivo $\mathbb{R}_{\geq}$ sobre o espaço X, ou seja, um homomorfismo de monoides entre o monoide aditivo $\mathbb{R}_{\geq 0}$ e o monoide de composição de funções no espaço de funções $\mathcal{F}(X, X)$.

:⊢ **Definição 24.2.** Um *sistema dinâmico contínuo* é uma dupla (X, f) em que

1. X é um conjunto, o *espaço de fase* do sistema;
2. $f\colon \mathbb{R}_{\geq 0} \times X \longrightarrow X$ é um semifluxo, a *dinâmica* do sistema.

Esse semifluxo pode ser também um fluxo completo, no sentido que de estar definido para todo $\mathbb{R}$, não somente em $\mathbb{R}_{\geq 0}$. Nesse caso, o fluxo seria uma função $f\colon \mathbb{R} \times X \longrightarrow X$, uma ação do grupo aditivo $\mathbb{R}$, o sistema seria *invertível* (ou *revertível*).

Essas duas definições estão relacionadas. No caso de um sistema dinâmico contínuo, podemos considerar o mapa de tempo 1 do (semi)fluxo, $f^1\colon X \longrightarrow X$, que nos dá um sistema dinâmico discreto em X. Por outro lado, a partir de um sistema discreto cuja dinâmica é $f\colon X \longrightarrow X$, podemos considerar as composições sucessivas da dinâmicas consigo mesma, denotadas $f^n = f \circ \cdots \circ f$, de modo que teríamos uma ação de $\mathbb{Z}_{\geq 0}$ sobre X (ou de $\mathbb{Z}$ caso o sistema seja invertível). Sob essa perspectiva, podemos definir uma noção mais geral de sistema dinâmico.

:⊢ **Definição 24.3.** Seja $\boldsymbol{M}$ um monoide. Um *sistema $\boldsymbol{M}$-dinâmico* é um par (X, f) em que

1. X é um conjunto, o *espaço de fase* do sistema;
2. $f\colon \boldsymbol{T} \curvearrowright X$ é uma ação de monoides em X, a *dinâmica* do sistema.

Por essa definição, sistemas dinâmicos discretos são *sistemas $\mathbb{Z}_{\geq 0}$-dinâmicos* (ou *$\mathbb{Z}$-dinâmicos* no caso invertível) e sistemas dinâmicos contínuos são *sistemas $\mathbb{R}_{\geq 0}$-dinâmicos* (ou *$\mathbb{R}$-dinâmicos* no caso invertível). Pode-se se considerar a ação de monoides e grupos mais gerais — às vezes $\mathbb{Z}^d$, $\mathbb{R}^d$ ou até algum grupo abstrato — mas isso não será feito aqui. Estudaremos principalmente sistemas dinâmicos discretos e, ocasionalmente, sistemas contínuos.

24.1.2 Órbitas

A partir de agora, apresentaremos as definições apenas para sistemas dinâmicos discretos.

:⊢ **Definição 24.4** (Iterados e conjunto órbita)**.** Sejam (X, f) um sistema dinâmico discreto, $x \in X$ e $i \in \mathbb{Z}_{\geq 0}$ (ou $i \in \mathbb{Z}$ caso o sistema seja invertível). O *i-ésimo iterado de x* é o ponto $f^i(x)$.

Seja $I \subseteq \mathbb{Z}_{\geq}$ (ou $I \subseteq \mathbb{Z}$ caso o sistemas seja invertível). O *conjunto de I-iterados de x* é

$$f^I(x) := \left\{ f^i(x) \;\middle|\; i \in I \right\}.$$

O *conjunto órbita* de um ponto x é o conjunto $\mathcal{O}(x) := f^{\mathbb{Z}_{\geq 0}}(x)$ (caso o sistema seja invertível, o *conjunto órbita* de x é $\mathcal{O}(x) := f^{\mathbb{Z}}(x)$ e o *conjunto de semiórbita positiva* de x é o conjunto $\mathcal{O}^+(x) := f^{\mathbb{Z}_{\geq 0}}(x)$); .

A mesma definição vale para sistemas contínuos.

Na definição anterior, consideramos o conjunto dos iterados de um ponto x, ou seja, os ponto pelos quais o ponto x passará ou, no caso invertível, pelos quais ele passou. No entanto, podemos também considerar a órbita como uma sequência de pontos de modo a ter informação sobre em qual tempo o ponto original passou por cada ponto de sua órbita. Consideraremos essa definição a seguir. Para isso, devemos usar sequências de pontos do espaço de fase, isto é, elementos $(x_i)_{i\in\mathbb{Z}}$ de $X^{\mathbb{Z}}$.

:⊢ **Definição 24.5** (Órbitas e espaço de órbitas)**.** Seja (X, f) um sistema dinâmico discreto. Uma *órbita* de (X, f) é uma sequência $(x_i)_{i\in\mathbb{Z}} \in X^{\mathbb{Z}}$ tal que, para todo $i \in \mathbb{Z}$, $f(x_i) = x_{i+1}$. O *espaço de órbitas de* (X, f) é

$$\widetilde{X} := \left\{(x_i)_{i\in\mathbb{Z}} \in X^{\mathbb{Z}} \;\middle|\; \forall_{i\in\mathbb{Z}}\ f(x_i) = x_{i+1}\right\},$$

o conjunto das órbitas do sistema. A *dinâmica do espaço de órbitas* é a função

$$\begin{aligned} \widetilde{f}\colon \widetilde{X} &\longrightarrow \widetilde{X} \\ (x_i)_{i\in\mathbb{Z}} &\longmapsto (f(x_i))_{i\in\mathbb{Z}}. \end{aligned}$$

Note que, para toda órbita $\widetilde{x} \in \widetilde{X}$, vale $(f(x_i))_{i\in\mathbb{Z}} = (x_{i+1})_{i\in\mathbb{Z}}$, pois por definição de órbita vale $f(x_i) = x_{i+1}$.

Note que a dinâmica $\widetilde{f}$ pode ser definida no espaço produto $X^{\mathbb{Z}}$

$$\begin{aligned} \widetilde{f}\colon X^{\mathbb{Z}} &\longrightarrow X^{\mathbb{Z}} \\ (x_i))_{i\in\mathbb{Z}} &\longmapsto (f(x_i))_{i\in\mathbb{Z}}. \end{aligned}$$

Como $\widetilde{X} \subseteq X^{\mathbb{Z}}$, a projeção natural

$$\begin{aligned} \mathrm{p}_0\colon X^{\mathbb{Z}} &\longrightarrow X \\ (x_i)_{i\in\mathbb{Z}} &\longmapsto x_0 \end{aligned}$$

do espaço produto pode ser restrita ao espaço de órbitas. Nesse caso, vale que, para toda órbita $(x_i))_{i\in\mathbb{Z}} \in \widetilde{X}$,

$$f \circ \mathrm{p}_0((x_i))_{i\in\mathbb{Z}}) = f(x_0) = x_1 = \mathrm{p}_0 \circ \widetilde{f}((x_i))_{i\in\mathbb{Z}}),$$

o que mostra que $f \circ \mathrm{p}_0 = \mathrm{p}_0 \circ \widetilde{f}$. O espaço de órbitas pode ser alternativamente descrito como

$$\widetilde{X} = \left\{x \in X^{\mathbb{Z}} \;\middle|\; f \circ \mathrm{p}_0(x) = \mathrm{p}_0 \circ \widetilde{f}(x)\right\}.$$

24.1.3 Conjuntos invariantes

24.1.3.1 Invariância positiva e negativa

O estudo de órbitas de sistemas dinâmicos pode ser especificado a partir do estudo de conjunto invariantes, que são subconjuntos do espaço de fase que são invariantes pela dinâmica em dois sentidos distintos: invariância positiva e invariância negativa. Esses conjuntos representam pedaços do espaço que estão de certa forma isolados do resto do sistema e podem ser analisados como subsistemas.

:⊢ **Definição 24.6.** Seja (X, f) um sistema dinâmico discreto.

1. Um conjunto *positivamente invariante* de (X, f) é um conjunto $C \subseteq X$ tal que $f(C) \subseteq C$;
2. Um conjunto *negativamente invariante* de (X, f) é um conjunto $C \subseteq X$ tal que $f^{-1}(C) \subseteq C$;
3. Um conjunto *invariante* de (X, f) é um conjunto $C \subseteq X$ que é positivamente e negativamente invariante.

Essas definições são equivalentes às seguintes propriedades em termos de todas as imagens positivas ou negativas do conjunto, que é um exercício simples.

▷ **Exercício 24.1.** *Sejam (X, f) um sistema dinâmico discreto e $C \subseteq X$.*

1. *O conjunto C é positivamente invariante se, e somente se, para todo $i \in \mathbb{Z}_{\geq 0}$ vale que $f^i(C) \subseteq C$, ou seja,*

$$\bigcup_{i \in \mathbb{Z}_{\geq 0}} f^i(C) \subseteq C;$$

2. *O conjunto C é negativamente invariante se, e somente se, para todo $i \in \mathbb{Z}_{\geq 0}$ vale que $f^{-i}(C) \subseteq C$, ou seja*

$$\bigcup_{i \in \mathbb{Z}_{\geq 0}} f^{-i}(C) \subseteq C;$$

Esa proposição sugere como definir invariância para sistemas dinâmicos contínuos: exigindo que todos iterados do conjunto estejam contidos no conjunto.

Para interpretar o significado dos conjuntos positivamente e negativamente invariantes, precisamos interpretar, para um conjunto $C \subseteq X$, o que a imagem direta $f(C)$ e imagem inversa $f^{-1}(C)$ significam. A imagem direta representa o conjunto dos 1-ésimos iterados dos pontos de C e a imagem inversa significa o conjunto de pontos cujos 1-ésimos iterados estão em C. No caso em que f é um invertível, a imagem inversa é simplesmente a imagem direta da dinâmica inversa f^{-1}, portanto o conjunto dos pontos onde os pontos de C estavam no iterado

anterior, mas quando a dinâmica não é invertível um ponto de C pode ter mais de um ponto que chega nela.

Por consequência, concluímos que $\bigcup_{i\in\mathbb{Z}_{\geq 0}} f^i(C)$ é o conjunto das semiórbitas positivas dos pontos de C, enquanto que o conjunto $\bigcup_{i\in\mathbb{Z}_{\geq 0}} f^{-i}(C)$ é o conjunto dos pontos cujas semiórbitas positivas passam por C. Isso fica claro a partir do seguinte exercício simples de igualdade de conjuntos.

▷ **Exercício 24.2.** *Sejam (X, f) um sistema dinâmico discreto e $C \subseteq X$.*

1. $\bigcup_{i\in\mathbb{Z}_{\geq 0}} f^i(C) = \bigcup_{x\in C} f^{\mathbb{Z}_{\geq 0}}(x)$;
2. $\bigcup_{i\in\mathbb{Z}_{\geq 0}} f^{-i}(C) = \left\{x \in X \;\middle|\; f^{\mathbb{Z}_{\geq 0}}(x) \cap C \neq \emptyset\right\}$.

A partir disso, podemos concluir que um conjunto positivamente invariante é um conjunto C de pontos do sistema cuja semiórbita positiva está contida em C, enquanto que um conjunto negativamente invariante é um conjunto C que contém todos os pontos do sistema cuja semiórbita positiva passa por C.

Ainda de forma mais simplificada, um conjunto positivamente invariante é aquele que não tem órbitas saindo dele, enquanto que um conjunto negativamente invariante é aquele que não tem órbitas entrando nele. Além disso, fica também claro que existe uma dualidade entre invariância positiva e negativa, como mostra a proposição seguinte.

⊢ **Proposição 24.1.** *Sejam (X, f) um sistema dinâmico discreto e $C \subseteq X$. O conjunto C é positivamente invariante se, e somente se, seu complementar $C^{\complement}$ é negativamente invariante.*

Uma definição comum de conjunto invariante é um conjunto $C \subseteq X$ tal que $f^{-1}(C) = C$. Essa definição é equivalente à que apresentamos aqui. Uma outra definição usa $f(C) = C$, mas essa definição nem sempre é equivalente, somente quando o sistema é invertível. Isso também é um exercício simples de conjuntos.

▷ **Exercício 24.3.** *Sejam (X, f) um sistema dinâmico discreto e $C \subseteq X$.*

1. *C é invariante se, e somente se $f^{-1}(C) = C$;*
2. *Se f é sobrejetiva e C é invariante, então $f(C) = C$;*
3. *Se f é injetiva e $f(C) = C$, então C é invariante.*

24.1.3.2 Subsistemas

:⊢ **Definição 24.7.** Um *conjunto invariante mínimo* é um conjunto invariante que não tem subconjuntos invariantes próprios e não vazios.

O espaço de fase sempre pode ser particionado em partes invariantes, embora o conjunto dessas partes possa ter qualquer cardinalidade. Para ver que isso é verdade, basta considerar o conjunto de órbitas $\{\mathcal{O}(x) \mid x \in X\}$. Cada órbitas é um conjunto invariante.

24.1.3.3 Núcleo de recorrência

Ressaltamos na proposição a seguir algumas equivalências para um conjunto bem importante na teoria de sistemas dinâmicos.

⊢ **Proposição 24.2.** *Sejam $\mathcal{S} = (X, f)$ um sistema dinâmico discreto e $C \subseteq X$.Para todo $x \in X$, $x \in \overline{\lim}_{t\in\mathbb{Z}_{\geq 0}} f^{-t}(C)$ se, e somente se, alguma uma sequência $(t_n)_{n\in\mathbb{N}}$ de inteiros positivos estritamente crescente ($t_0 < t_1 < \cdots$) satisfaz que, para todo $n \in \mathbb{N}$, $f^{t_n}(x) \in C$.*

□ *Demonstração.* O conjunto $\overline{\lim}_{t\in\mathbb{Z}_{\geq 0}} f^{-t}(C)$ é o conjunto dos pontos que pertencem a infinitos dos conjuntos $f^{-t}(C)$, ou seja, os pontos $x \in X$ para os quais existem uma sequência $(t_n)_{n\in\mathbb{N}}$ de inteiros positivos estritamente crescente ($t_0 < t_1 < \cdots$) tal que, para todo $n \in \mathbb{N}$, $f^{t_n}(x) \in C$. ■

24.1.4 Periodicidade

:⊢ **Definição 24.8** (Ponto periódico)**.** Seja (X, f) um sistema dinâmico discreto. Um *ponto periódico* é um ponto $x \in X$ que satisfaz, para algum $i \in \mathbb{Z}_{>0}$, $f^i(x) = x$; nesse caso, i é um *período* de x e o ponto é também dito i-periódico. Um *ponto fixo* é um ponto 1-periódico.

O conjunto dos pontos periódicos é

$$\Pi(f) := \left\{x \in X \;\middle|\; \exists_{i\in\mathbb{Z}_{>0}}\ f^i(x) = x\right\}$$

e, para cada $i \in \mathbb{Z}_{>0}$, o conjunto dos pontos i-periódicos é denotado $\Pi^i(f)$.

$$\Pi^i(f) := \left\{x \in X \;\middle|\; f^i(x) = x\right\}.$$

Por consequência da definição, vale

$$\Pi(f) = \bigcup_{i\in\mathbb{Z}_{\geq 0}} \Pi^i(f).$$

Na proposição a seguir usamos a relação de divisão $\preceq$ entre números inteiros — um número inteiro $i \in \mathbb{Z}_{>0}$ divide outro $i' \in \mathbb{Z}_{>)}$, denotado $i \preceq i'$, quando para algum $q \in \mathbb{Z}_{>)}$ vale $iq = i'$.

⊢ **Proposição 24.3.** *Sejam (X, f) um sistema dinâmico discreto e $i \in \mathbb{Z}_{>0}$.*

1. *Para todo $i' \in \mathbb{Z}_{>0}$ tal que $i \preceq i'$, vale $\Pi^i(f) \subseteq \Pi^{i'}(f)$;*
2. *A dinâmica f restrita a $\Pi^i(f)$ é invertível e $f^{-1} = f^{i-1}$;*
3. *O conjunto $\Pi^i(f)$ é positivamente invariante.*

□ *Demonstração.* 1. Se $i \preceq i'$, para algum $q \in \mathbb{Z}_{>0}$ vale que $iq = i'$. Seja $x \in \Pi^i$, de modo que vale $f^i(x) = x$. Mostraremos por indução em q que $x \in \Pi^{i'}(f)$. Para $q = 1$, vale $f^{i'}(x) = f^{i1}(x) = f^i(x) = x$. Agora suponhamos que a proposição vale para q e mostremos que vale para $q + 1$. Por hipótese, vale $f^{iq}(x) = x$. Mas então

$$f^{i(q+1)}(x) = f^{iq} \circ f^i(x) = f^{iq}(x) = x.$$

2. Seja $x \in \Pi^i(f)$. Então vale $f^i(x) = x$, o que implica que $f^{i-1}(f(x)) = f(f^{i-1}(x)) = f^i(x) = x$, portanto $f^{i-1} = f^{-1}$ em $\Pi^i(f)$.
3. Devemos mostrar que $f(\Pi^i(f)) \subseteq \Pi^i(f)$. Seja $x' \in f(\Pi^i(f))$ e $x \in \Pi^i(f)$ tal que $x' = f(x)$. Então vale $f^i(x) = x$, logo

$$f^i(f(x)) = f^{i+1}(x) = f(f^i(x)) = f(x).$$

Isso mostra que $x' = f(x) \in \Pi^i(f)$, o que implica que $f(\Pi^i(f)) \subseteq \Pi^i(f)$. ■

Em particular, essa proposição mostra que qualquer iterado de um ponto periódico é periódico e tem o mesmo período do ponto original.

Além de considerar diferentes períodos de um ponto, podemos também considerar o menor desses valores.

:⊢ **Definição 24.9.** Sejam (X, f) um sistema dinâmico discreto e $x \in X$. O *período mínimo* de x é o menor $i \in \mathbb{Z}_{>0}$ tal que $x \in \Pi^i(f)$.

Esse mínimo a princípio é dado em relação à ordem natural de $\mathbb{Z}$, mas, como vale que se $i \preceq i'$ então $\Pi^i(f) \subseteq \Pi^{i'}(f)$ — ou seja, se um número é período de um ponto, qualquer múltiplo dele também é período desse ponto — podemos também considerar a ordem de divisão em $\mathbb{Z}_{\geq 0}$ e o mínimo dessa ordem, que é o máximo divisor comum, denotado $\curlywedge$. Isso inclusive nos permite, ao definir o período mínimo, considerar o 0 como um possível valor para o período. Claramente, todos pontos têm 'período' 0, pois satisfazem $f^0(x) = x$, mas, sob a relação de ordem de divisão, o 0 é o *maior* número inteiro, pois qualquer outro inteiro o divide, e assim ao definir o período mínimo de um ponto como o menor período usando o máximo divisor comum, obteríamos que os pontos cujo período mínimo é 0 são os pontos que não são periódicos.

24.2 Conjugação de sistemas dinâmicos

:⊢ **Definição 24.10.** Sejam $\mathcal{S}_1 = (X_1, f_1)$ e $\mathcal{S}_2 = (X_2, f_2)$ sistemas dinâmicos. Uma *semiconjugação* (morfismo de sistemas dinâmicos) de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma função sobrejetiva $\phi : X_1 \to X_2$ que satisfaz $\phi \circ f_1 = f_2 \circ \phi$ (o diagrama comuta).

$$\begin{tikzcd} X_1 \arrow[r, "f_1"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\ X_2 \arrow[r, "f_2"'] & X_2 \end{tikzcd}$$

Denota-se $\phi : \mathcal{S}_1 \to \mathcal{S}_2$. Uma *conjugação* (isomorfismo de sistemas dinâmicos) de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma semiconjugação $\phi : \mathcal{S}_1 \to \mathcal{S}_2$ que é invertível.

Sempre que temos uma função sobrejetiva $\phi : X_1 \to X_2$, para todo $y \in X_2$ existe $x \in X_1$ tal que $y = \phi(x)$. Assim podemos induzir a dinâmica de (X_1, f_1) em X_2, definindo $f_2(y) := \phi(f_1(x))$, de modo que a dinâmica f_2 é uma função de X_2 para X_2 e vale $f_2 \circ \phi = \phi \circ f_1$. Isso faz com que os sistemas (X_1, f_1) e (X_2, f_2) sejam semiconjugados.

A partir das propriedade da conjugação, vamos deduzir alguns resultados imediatos. Podemos notar que as funções compostas ${f_1}^n$ e ${f_2}^n$ ainda satisfazem a propriedade comutativa que f_1 e f_2 satisfazem. Isso basicamente diz que em

sistemas discretos, preservar a função f é preservar a ação da dinâmica. Segue a proposição.

⊢ **Proposição 24.4.** *Sejam (X_1, f_1) e (X_2, f_2) sistemas dinâmicos e $\phi : X_1 \to X_2$ uma semiconjugação. Então, para todo natural $n \geq 1$, vale $\phi \circ {f_1}^n = {f_2}^n \circ \phi$ (o diagrama comuta).*

$$\begin{tikzcd} X_1 \arrow[r, "{f_1}^n"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\ X_2 \arrow[r, "{f_2}^n"'] & X_2 \end{tikzcd}$$

□ *Demonstração.* Vamos provar a proposição acima usando indução. Primeiro, notamos que a propriedade vale para $n = 1$ por definição. Agora, supondo que valha $\phi \circ {f_1}^{n-1} = {f_2}^{n-1} \circ \phi$, temos que

$$\begin{aligned} \phi \circ {f_1}^n &= \phi \circ ({f_1}^{n-1} \circ f_1) \\ &= (\phi \circ {f_1}^{n-1}) \circ f_1 \\ &= ({f_2}^{n-1} \circ \phi) \circ f_1 \\ &= {f_2}^{n-1} \circ (\phi \circ f_1) \\ &= {f_2}^{n-1} \circ (f_2 \circ \phi) \\ &= ({f_2}^{n-1} \circ f_2) \circ \phi \\ &= {f_2}^n \circ \phi, \end{aligned}$$

o que conclui a indução. ■

Ainda, podemos ver que a semiconjugação preserva órbitas periódicas, no sentido em que, se algum elemento do conjunto inicial tem órbita periódica, então sua imagem pela semiconjugação terá, também, órbita periódica.

⊢ **Proposição 24.5** (Semiconjugação preserva peridiocidade da órbita.)**.** *Sejam (X_1, f_1) e (X_2, f_2) sistemas dinâmicos e $\phi : X_1 \to X_2$ uma semiconjugação. Para todo $x \in X$, se x é um ponto n-periódico, então $\phi(x)$ é um ponto n-periódico. Ainda, o período mínimo de $\phi(x)$ divide o período mínimo de x.*

□ *Demonstração.* Seja $x \in X$ tal que $x = {f_1}^n(x)$. Pela propriedade da semiconjugação,

$$\begin{aligned}\phi(x) &= \phi({f_1}^n(x)) \\ &= (\phi \circ {f_1}^n)(x) \\ &= ({f_2}^n \circ \phi)(x) \\ &= {f_2}^n(\phi(x)).\end{aligned}$$

Como $\phi(x)$ é n-periódico, segue que seu período mínimo divide n. ■

24.3 Translação em grupos

Sejam $\boldsymbol{G}$ um grupo e $\alpha \in G$. A translação à esquerda por α é a função

$$\begin{aligned}E_\alpha \colon G &\longrightarrow G \\ g &\longmapsto \alpha g.\end{aligned}$$

24.3.1 Translação no $\mathbb{T}^1$

:⊢ **Definição 24.11.** Seja $[\alpha] \in \mathbb{T}$. A *rotação* por $[\alpha]$ em $\mathbb{T}^1$ é a função

$$\begin{aligned}R_\alpha \colon \mathbb{T}^1 &\longrightarrow \mathbb{T}^1 \\ [x] &\longmapsto [x + \alpha].\end{aligned}$$

Uma *rotação racional* é uma rotação cuja constante α é racional e uma *rotação irracional* é uma rotação cuja constante α é irracional.

⊢ **Proposição 24.6.** *A dupla* $(\mathbb{T}^1, R_\alpha)$ *é um sistema dinâmico discreto revertível para todo* $\alpha \in \mathbb{R}$ *e*

1. *se* α *é racional, todos os pontos de* $\mathbb{T}^1$ *são periódicos;*
2. *se* α *é irracional, nenhum ponto de* $\mathbb{T}^1$ *é periódico.*

□ *Demonstração.* Por definição a dupla é um sistema dinâmico discreto e, como $R_{-\alpha} = {R_\alpha}^{-1}$, segue que o sistema é invertível.

1. Seja $[x] \in \mathbb{T}^1$. Para mostrar que $[x]$ é periódico, devemos achar um natural n tal que ${R_\alpha}^n([x]) = [x]$. A constante α é um número racional e, por isso, pode ser escrita em forma de fração como $\alpha = \frac{p}{q}$, sendo p e q inteiros e q não nulo. Tomando $n = q$, e notando que $q\alpha = p$ é um inteiro, segue que

   $$ {R_\alpha}^q([x]) = [x + q\alpha] = [x] $$

   o que mostra que $[x]$ é q-periódico.

2. Seja $[x] \in \mathbb{T}^1$. Podemos provar a afirmação acima por contradição, notando que, se houvesse algum inteiro n tal que ${R_\alpha}^n([x]) = [x]$, como ${R_\alpha}^n([x]) = [x + n\alpha]$, teríamos $n\alpha$ inteiro. Mas como n é inteiro e α irracional, isso é impossível. ■

24.3.2 Translação no $\mathbb{T}^d$

Um caso mais geral que a rotação no círculo $\mathbb{T}^1$ é o caso da translação no toro $\mathbb{T}^d$. Ela é basicamente uma rotação em cada entrada e é chamada de translação pois é a transformação induzida no espaço quociente $\mathbb{T}^d$ de uma translação em $\mathbb{R}^d$. Ela é também um caso específico de translações em grupos topológicos.

:⊢ **Definição 24.12.** Seja $[\alpha] \in \mathbb{T}^d$. A *translação* por $[\alpha]$ em $\mathbb{T}^d$ é a função

$$\begin{aligned} T_\alpha \colon \mathbb{T}^d &\longrightarrow \mathbb{T}^d \\ [x] &\longmapsto [x + \alpha]. \end{aligned}$$

24.4 Expansão no $\mathbb{T}^1$

:⊢ **Definição 24.13.** Seja $N \in \mathbb{Z}_{\geq 0}$. A *expansão por* N em $\mathbb{T}^1$ é a função

$$\begin{aligned} E_N \colon \mathbb{T}^1 &\longrightarrow \mathbb{T}^1 \\ [x] &\longmapsto [Nx]. \end{aligned}$$

Note que a função está bem definida porque N é um número inteiro.

24.5 Deslocamentos simbólicos

Adotaremos a convenção de que o natural $N \in \mathbb{N}$ é o conjunto dos naturais menores que ele, seguindo a construção de ordinais do Von Neumann. Isso simplificará a notação. Assim temos que $N = \{0, \cdots, N-1\}$. Nesta seção consideraremos duas dinâmicas semelhantes, pois ambas agem em sequências, mas diferentes, pois uma considera sequências unilaterais e outra, bilaterais.

24.5.1 Deslocamento unilateral

Consideremos o conjunto

$$N^{\mathbb{Z}_{\geq 0}} = \left\{ (x_i)_{i \in \mathbb{Z}_{\geq 0}} \;\middle|\; \forall i \in \mathbb{Z}_{\geq 0} \quad x_i \in N \right\}$$

das *sequências unilaterais* em N símbolos, as funções do tipo

$$\begin{aligned} x\colon \mathbb{Z}_{\geq 0} &\longrightarrow N \\ i &\longmapsto x_i. \end{aligned}$$

:⊢ **Definição 24.14.** O *deslocamento unilateral* em $N^{\mathbb{Z}_{\geq 0}}$ é a função

$$\begin{aligned} \sigma\colon N^{\mathbb{Z}_{\geq 0}} &\longrightarrow N^{\mathbb{Z}_{\geq 0}} \\ (x_i)_{i\in\mathbb{Z}_{\geq 0}} &\longmapsto (x_{i+1})_{i\in\mathbb{Z}_{\geq 0}}. \end{aligned}$$

Essa dinâmica é conhecida como *deslocamento de Bernoulli* [1]. Basicamente, o deslocamento desloca cada entrada de uma sequência $(x_0, x_1, x_2, \ldots)$ à esquerda, levando-a para sequência $(x_1, x_2, x_3, \ldots)$. Claramente essa função não é invertível, logo o sistema dinâmico formado por ela não é invertível. Ressaltamos isso na proposição a seguir.

⊢ **Proposição 24.7.** *A dupla $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ é um sistema dinâmico discreto irrevertível.*

⊢ **Proposição 24.8.** *Seja $k \in \mathbb{Z}_{\geq 0}$. Uma sequência unilateral $(x_i)_{i\in\mathbb{Z}_{\geq 0}} \in N^{\mathbb{Z}_{\geq 0}}$ é um ponto k-periódico de $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ se, e somente se, para todo $i \in \mathbb{Z}_{\geq 0}$, vale $x_i = x_{i+k}$.*

□ *Demonstração.* Se vale $x_i = x_{i+k}$ para todo $i \in \mathbb{Z}_{\geq 0}$, a sequência $(x_i)_{i\in\mathbb{Z}_{\geq 0}}$ claramente será k-periódica, pois

$$\sigma^k((x_i)_{i\in\mathbb{Z}_{\geq 0}}) = (x_{i+k})_{i\in\mathbb{Z}_{\geq 0}} = (x_i)_{i\in\mathbb{Z}_{\geq 0}}.$$

Por outro lado, se $(x_i)_{i\in\mathbb{Z}_{\geq 0}}$ tiver órbita periódica, valerá

$$\sigma^k((x_i)_{i\in\mathbb{Z}_{\geq 0}}) = (x_i)_{i\in\mathbb{Z}_{\geq 0}}$$

para algum k. Nesse caso, teremos que $x_i = x_{i+k}$ para todo $i \in \mathbb{Z}_{\geq 0}$. ■

⊢ **Proposição 24.9.** *A função*

$$\begin{aligned} \phi\colon N^{\mathbb{Z}_{\geq 0}} &\longrightarrow \mathbb{T}^1 \\ (x_i)_{i\in\mathbb{Z}_{\geq 0}} &\longmapsto \left[\mathop{+}_{i\in\mathbb{Z}_{\geq 0}} \frac{x_i}{N^{i+1}}\right] \end{aligned}$$

é uma semiconjugação de $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ para $(\mathbb{T}^1, E_N)$.

[1] *Jacob Bernoulli* (1655 – 1705), matemático suíço.

□ *Demonstração.* Seja $(x_i)_{i\in\mathbb{Z}_{\geq 0}} \in N^{\mathbb{Z}_{\geq 0}}$. Então

$$\begin{aligned}
\phi \circ \sigma((x_i)_{i\in\mathbb{Z}_{\geq 0}}) &= \phi((x_{i+1})_{i\in\mathbb{Z}_{\geq 0}}) \\
&= \left[\mathop{+}_{i\in\mathbb{Z}_{\geq 0}} \frac{x_{i+1}}{N^{i+1}} \right] \\
&= \left[N \mathop{+}_{i\in\mathbb{Z}_{\geq 0}} \frac{x_{i+1}}{N^{i+2}} \right] \\
&= \left[N \left(\frac{x_0}{N} + \mathop{+}_{i\in\mathbb{Z}_{>0}} \frac{x_i}{N^{i+1}} \right) \right] \\
&= \left[N \mathop{+}_{i\in\mathbb{Z}_{\geq 0}} \frac{x_i}{N^{i+1}} \right] \\
&= E_N \circ \phi((x_i)_{i\in\mathbb{Z}_{\geq 0}}).
\end{aligned}$$

Somamos x_0 na quarta linha pois é um inteiro, não altera a classe de equivalência. ■

Notemos que a função ϕ não é uma conjugação porque não é invertível — existe mais de uma representação N-ária para alguns números reais em $[0, 1]$, portanto mais de uma sequência em $N^{\mathbb{Z}_{\geq 0}}$ que o representa. No entanto, veremos mais à frente que isso não será problema quando considerarmos sistemas dinâmicos de medida, pois o conjunto desses pontos problemáticos terá medida nula e, portanto, a conjugação poderá ser definida. Isso será importante pois significa que os dois sistemas são o mesmo sob a ótica de sistemas dinâmicos de medida.

24.5.2 Deslocamento bilateral

Consideremos o conjunto

$$N^{\mathbb{Z}} = \{(x_i)_{i\in\mathbb{Z}} \mid \forall i \in \mathbb{Z} \quad x_i \in N\}$$

das *sequências bilaterais* em N símbolos, as funções do tipo

$$\begin{aligned}
x \colon \mathbb{Z} &\longrightarrow N \\
i &\longmapsto x_i.
\end{aligned}$$

:⊢ **Definição 24.15.** O *deslocamento bilateral (à esquerda)* em $N^{\mathbb{Z}}$ é a função

$$\begin{aligned}
\sigma \colon N^{\mathbb{Z}} &\longrightarrow N^{\mathbb{Z}} \\
(x_i)_{i\in\mathbb{Z}} &\longmapsto (x_{i+1})_{i\in\mathbb{Z}}.
\end{aligned}$$

O *deslocamento bilateral à direita* é σ^{-1}.

⊢ **Proposição 24.10.** *A dupla $(N^{\mathbb{Z}}, \sigma)$ é um sistema dinâmico discreto revertível.*

24.5.3 Subdeslocamentos

Os deslocamentos definidos anteriormente são deslocamentos completos, no sentido de que todas as sequências são consideradas no sistema. É possível, porém, selecionar subconjuntos dessas sequências tal que o deslocamento ainda está bem definido nelas. Definimos que cada entrada de uma sequência $(x_i)_{i\in\mathbb{Z}} \in N^{\mathbb{Z}}$ tem entradas em $\{0, \ldots, N-1\}$. Esses subconjuntos são escolhidos a partir de uma regra que indica qual entrada de uma sequência pode ser seguida de qual outra entra — por exemplo, para $N = 9$, sequências em que uma entrada 4 só pode ser seguida de uma entrada 2 ou uma entrada 8. Para ter toda a informação necessária, essas sequências são definidas a partir de uma *matriz de adjacência*, uma matriz que indica qual entrada pode ser adjacente a qual outra.

:⊢ **Definição 24.16.** Uma *matriz de adjacência* de ordem N é uma matriz $A \in \mathbb{M}_N(\mathbb{Z})$ de entradas 0 ou 1 tal que $a_{i,j} = 0$ se uma entrada com valor i não pode ser seguida de uma entrada com valor j e $a_{i,j} = 1$ se uma entrada de valor i pode ser seguida de uma entrada de valor j.

:⊢ **Definição 24.17.** Sejam $N \in \mathbb{Z}_{>0}$ e A uma matriz de adjacência de ordem N. O conjunto das sequências de $N^{\mathbb{Z}}$ *admissíveis* por A é o conjunto

$$N^{\mathbb{Z}}|_A := \left\{(x_i)_{i\in\mathbb{Z}} \;\middle|\; \forall i \in \mathbb{Z}, a_{x_i,x_{i+1}} = 1\right\}.$$

O conjunto das sequências de $N^{\mathbb{Z}_{\geq 0}}$ restritas por A é o conjunto

$$N^{\mathbb{Z}_{\geq 0}}|_A := \left\{(x_i)_{i\in\mathbb{Z}_{\geq 0}} \;\middle|\; \forall i \in \mathbb{Z}_{\geq 0}, a_{x_i,x_{i+1}} = 1\right\}.$$

O *subdeslocamento bilateral* com *matriz de adjacência* A é a função $\sigma|_A := \sigma|_{N^{\mathbb{Z}}|_A}$ e o *subdeslocamento unilateral* com *matriz de adjacência* A é a função $\sigma|_A := \sigma|_{N^{\mathbb{Z}_{\geq 0}}|_A}$. Os sistemas $(N^{\mathbb{Z}}|_A, \sigma|_A)$ e $(N^{\mathbb{Z}_{\geq 0}}|_A, \sigma|_A)$ são *sistemas de deslocamento parcial* e a matriz A é a *matriz de adjacência* dos sistemas.

Esses sistemas também são conhecidos como *cadeias de Markov* [2].

⊢ **Proposição 24.11.** *Sejam $N \in \mathbb{Z}_{>0}$ e A uma matriz de adjacência de ordem N. A inclusão $\iota\colon N^{\mathbb{Z}}|_A \to N^{\mathbb{Z}}$ é uma semiconjugação do sistema $(N^{\mathbb{Z}}|_A, \sigma|_A)$ para o sistema $(N^{\mathbb{Z}}|_A, \sigma|_A)$.*

□ *Demonstração.* Queremos mostrar que $\iota \circ \sigma|_A = \sigma \circ \iota$. Seja $(x_i)_{i\in\mathbb{Z}} \in N^{\mathbb{Z}}|_A$.

[2] *Andrei Andreyevich Markov* (1856 – 1922), matemático russo.

Então

$$\begin{aligned}
\iota \circ \sigma|_A((x_i)_{i\in\mathbb{Z}}) &= \iota(\sigma|_A((x_i)_{i\in\mathbb{Z}})) \\
&= \iota((x_{i+1})_{i\in\mathbb{Z}}) \\
&= (x_{i+1})_{i\in\mathbb{Z}} \\
&= \sigma((x_i)_{i\in\mathbb{Z}}) \\
&= \sigma(\iota((x_i)_{i\in\mathbb{Z}})) \\
&= \sigma \circ \iota((x_i)_{i\in\mathbb{Z}}).
\end{aligned}$$

■

O mesmo vale para os sistemas unilaterais. A partir da matriz da matriz de adjacência podemos determinar quantos pontos periódicos de um determinado período existem.

⊢ **Proposição 24.12.** *Sejam $p \in \mathbb{Z}_{>0}$ e $(N^{\mathbb{Z}}|_A, \sigma|_A)$ um sistema de subdeslocamento com matriz de adjacência A. A quantidade de pontos p-periódicos de $(N^{\mathbb{Z}}|_A, \sigma|_A)$ é* $\operatorname{tr}(A^p)$.

□ *Demonstração.* Sabemos que $(x_i)_{i\in\mathbb{Z}}$ é p-periódico se, e somente se, para todo $i \in \mathbb{Z}$ vale $x_i = x_{i+p}$. Sendo assim, procuramos os pontos da forma

$$(\ldots, x_{i_0}, x_{i_1}, \ldots, x_{i_{p-2}}, x_{p-1}, \ldots)$$

e portanto devemos achar os vetores $(i_0, \ldots, i_{p-1}) \in N^p$ tais que a sequência de entradas

$$i_0 \to i_1 \to \cdots i_{p-2} \cdots i_{p-1}$$

é uma sequência permitida pela matriz de adjacência. Isso é equivalente a termos

$$\bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} = a_{i_0,i_1} a_{i_1,i_2} \cdots a_{i_{p-3},i_{p-2}} a_{i_{p-2},i_{p-1}} = 1,$$

pois esse produto é 1 se, e somente se todos os fatores são 1. Assim, a quantidade de pontos periódicos é a soma de todos esses produtos

$$\begin{aligned}
|\Pi^p(\sigma_A)| &= \bigplus_{(i_0,\ldots,i_{p-1})\in N^p} \bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} \\
&= \bigplus_{i_0=0}^{N-1} \bigplus_{(i_1,\ldots,i_{p-1})\in N^{p-1}} \bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} \\
&= \bigplus_{i_0=0}^{N-1} (A^p)_{i_0,i_0} = \operatorname{tr}(A^p).
\end{aligned}$$

■

24.5.4 Itinerários

:⊢ **Definição 24.18.** Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X ($N \in \mathbb{N}$) e $x \in X$. O *itinerário* de x com respeito a $\mathcal{P}$ é a sequência $(x_t)_{t\in T} \in N^T$ tal que $x_t = k$ se, e somente se, $f^t(x) \in P_k$.

⊢ **Proposição 24.13.** *Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X e $\phi : X \to N^T$ a função que leva $x \in X$ em seu itinerário com respeito a $\mathcal{P}$. Então*

$$\phi \circ f = \sigma \circ \phi.$$

□ *Demonstração.* Sejam $x \in X$ e $\phi(x) = (x_t)_{t\in T}$. Então $\phi(f(x)) = (x_{t+1})_{t\in T}$, pois $f^t(f(x)) = f^{t+1}(x)$. Portanto

$$\phi \circ f(x) = \phi(f(x)) = (x_{t+1})_{t\in T} = \sigma((x_t)_{t\in T}) = \sigma(\phi(x)) = \sigma \circ \phi(x),$$

o mostra que $\phi \circ f = \sigma \circ \phi$. ■

A imagem $\phi(X)$ é invariante por σ, pois

$$\sigma(\phi(X)) = \phi(f(X)) = \phi(X),$$

de modo que ϕ é uma semiconjugação de XX para $\phi(X)$, pois satisfaz a propriedade comutativa e é sobrejetiva por definição. Nem sempre tem-se ϕ injetiva, isso depende da partição e do sistema que se estuda. Ainda, nem sempre ϕ preserva as estruturas adicionais de X. Nas seções posteriores serão apresentadas partições que preservam tais estruturas.

Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X e $(x_t)_{t\in T} \in N^T$. O conjunto

$$\bigcap_{t\in T} f^{-t}(P_{x_t}).$$

é o conjunto dos pontos $x \in X$ tal que $f^t(x) \in P_{x_t}$ para todo $t \in T$, pois $f^t(x) \in P_{x_t}$ se, e somente se, $x \in f^{-t}(P_{x_t})$; ou seja, o conjunto de todos os pontos tais que $\phi(x) = (x_t)_{t\in T}$.

24.5.5 Automorfismos hiperbólicos no $\mathbb{T}^2$

Após o estudo de duas dinâmicas envolvendo o círculo, a dinâmica de rotação na e a de expansão, voltamos nossa atenção nesta seção a um novo objeto geométrico: o toro. O sistema dinâmico a ser estudado nesta seção tem como conjunto o toro $\mathbb{T}^2$, descrito acima, e a dinâmica que age nele é chamada de automorfismo hiperbólico. Essa função é dada por uma matriz de dimensão 2×2 com algumas características

específicas. Como a próxima definição dependente muito da utilização de matrizes, convém definir algumas coisas sobre elas antes de abordarmos a dinâmica no toro. Denotaremos uma matriz sempre com uma letra maiúscula e seus elementos com a respectiva letra minúscula acompanhados dos índices; escrevemos $A = [a_{i,j}]$. A dimensão de uma matriz será explicitada somente quando necessário e será simplificada para matrizes quadradas, como em $A_{2x2} = A_2$. Denotaremos o determinante de uma matriz A por $|A|$ ou $\det(A)$ e sua matriz inversa por A^{-1}. Convenções esclarecidas, definimos o automorfismo.

:⊢ **Definição 24.19.** Seja $A \in \mathrm{SL}_2(\mathbb{Z})$ uma matriz cujos autovalores têm valor absoluto não unitário. O *automorfismo hiperbólico* no toro é definido a partir de A como

$$\begin{aligned} [A] \colon \mathbb{T}^2 &\longrightarrow \mathbb{T}^2 \\ [x] &\longmapsto [Ax]. \end{aligned}$$

Para justificar a nomenclatura dessa classe de funções como automorfismos hiperbólicos, devemos demonstrar algumas propriedades tanto da matriz A como da função $[A]$. A matriz é dada por

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

As entradas devem ser inteiras para que a operação $[A](x)$ esteja bem definida no toro. Isso ocorre pois, para todo $y \in [x]$, temos $y = x + k$, com $k \in \mathbb{Z}^2$, e segue que

$$[Ay] = [A(x + k)] = [Ax + Ak] = [Ak],$$

pois $Ak \in \mathbb{Z}^2$.

Devemos justificar por que essa função é um automorfismo. Claramente, a função $[A]$ é um endomorfismo por ter como domínio e contradomínio o mesmo conjunto. Para ser um automorfismo, deve ser também um isomorfismo. Isso ocorre porque o determinante da matriz A é igual a 1, o que faz com que sua matriz inversa

$$A^{-1} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

seja também uma matriz com entradas inteiras, determinante unitário e autovalores distintos, e a função inversa $L_{A^{-1}} := [A^{-1}x]$ esteja bem definida no toro. O automorfismo $[A]$ recebe o nome de hiperbólico porque matrizes cujos autovalores têm valor absoluto diferente de 1 recebem esse nome. Vale notar, por fim, que para todo inteiro k, vale

$$[A]^k = [A^k].$$

A partir da definição da matriz A, podemos deduzir algumas características imediatas dos autovalores, autovetores e determinante da matriz.

⊢ **Proposição 24.14.** *Seja $A \in \mathrm{SL}_2(\mathbb{Z})$ uma matriz cujos autovalores têm valor absoluto não unitário. Então*

1. *Os autovalores de A são*

$$\lambda_+ := \frac{1}{2}\left(a_{11} + a_{22} + \sqrt{(a_{11} + a_{22})^2 - 4}\right)$$
$$\lambda_- := \frac{1}{2}\left(a_{11} + a_{22} - \sqrt{(a_{11} + a_{22})^2 - 4}\right)$$

   e as entradas de A satisfazem $a_{11} + a_{22} \neq \pm 2$, $a_{12} \neq 0$ e $a_{21} \neq 0$.
2. *Os autovalores de λ_+ e λ_- são irracionais;*
3. *Os autovetores de A são*

$$v_{\lambda_+} := \left(1, \frac{\lambda_+ - a_{11}}{a_{12}}\right)$$
$$v_{\lambda_-} := \left(1, \frac{\lambda_- - a_{11}}{a_{12}}\right).$$

□ *Demonstração.* 1. Primeiro, notemos que o polinômio característico de A será

$$\chi_A(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + 1.$$

Assim, isolando λ pela fórmula quadrática, obtemos λ_+ e λ_- como no enunciado. Para termos $\lambda_+ \neq \lambda_-$, devemos ter $\sqrt{(a_{11} + a_{22})^2 - 4} \neq 0$, o que implica $a_{11} + a_{22} \neq \pm 2$. Por fim, se $a_{12} = 0$ ou $a_{21} = 0$, temos $a_{12}a_{21} = 0$, o que implica $\det A = a_{11}a_{22} = 1$. Teríamos, portanto, $a_{22} = \frac{1}{a_{11}}$, mas como todas entradas são inteiras, isso implicaria $a_{11} = a_{22} = \pm 1$, e teríamos a contradição $a_{11} + a_{22} = \pm 2$.

2. Vamos demonstrar o lema por absurdo. Primeiramente notamos que os autovalores de A devem ser reais pois, caso contrário, seriam complexos conjugados. Então seu produto seria $\lambda_+\lambda_- = |\lambda_+|^2 = |\lambda_-|^2 = \pm 1$. Mas isso é absurdo, pois os módulos dos autovalores são diferentes de 1. Consideremos o autovalor λ_+ e suponhamos que ele é um número racional. Ao isolar a raiz na expressão de λ_+ obtemos

$$\sqrt{(a_{11} + a_{22})^2 - 4} = 2\lambda_+ - (a_{11} + a_{22}).$$

Mas $(a_{11} + a_{22})^2 - 4$ é inteiro e sabemos que a raiz de um número inteiro é inteira ou irracional. Devemos ter $2\lambda_+ - (a_{11} + a_{22})$ inteiro, portanto. Fazendo $x = a_{11} + a_{22}$ e $y = 2\lambda_+ - (a_{11} + a_{22})$, devemos ter x e y inteiros e, rearrumando a equação,

$$x^2 - y^2 = 4.$$

De fato, para achar uma solução inteira da equação acima, notamos que $x^2 - y^2 = (x+y)(x-y)$ e temos três opções, já que $4 = 2 \cdot 2 = 4 \cdot 1$. A primeira é $x + y = 2$ e $x - y = 2$. Isso implica $x + y = x - y$, que implica $y = 0$ e, por fim, $a_{11} + a_{22} = x = 2$, o que contradiz o item anterior da proposição. As outras duas opções são $x + y = 4$ e $x - y = 1$ ou $x + y = 1$ e $x - y = 4$. Ambas, no entanto, implicam $x = \frac{5}{2}$, o que é um absurdo porque $x = a_{11} + a_{22}$ é um número inteiro. Assim, concluímos que o autovalor λ_+ é irracional. Como $\lambda_- = \frac{1}{\lambda_+}$, ambos autovalores são ambos irracionais.

3. Para calcular os autovalores de A, devemos resolver o sistema de equações $(A - \lambda I)v = 0$, em que λ representa qualquer um dos dois autovalores. De fato,

$$(A - \lambda I) = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}.$$

Notemos que $a_{11} \neq \lambda$, pois a_{11} é um inteiro e os autovalores de A são irracionais. Assim, concluímos que $a_{11} \neq \lambda$, seja $\lambda = \lambda_+$ ou $\lambda = \lambda_-$. Podemos, portanto, resolver o sistema inicial $(A - \lambda I)v = 0$ usando a matriz M a seguir, obtida a partir da matriz $(A - \lambda I)$ ao se dividir a primeira linha por $a_{11} - \lambda$ e subtraí-la, multiplicada por a_{21}, da segunda linha. Nota-se que, ao se dar o trabalho de calcular o descrito anteriormente, a expressão que aparece na segunda entrada da linha debaixo da matriz M vale 0, pois o numerador é o determinante de $(A - \lambda I)$.

$$M = \begin{bmatrix} 1 & \frac{a_{12}}{a_{11} - \lambda} \\ 0 & 0 \end{bmatrix}.$$

Ao resolvermos $Mv = 0$ para $v = (x, y)$, obtemos a expressão $x + \frac{a_{12}}{a_{11} - \lambda} y = 0$. Por fim, notemos que $a_{12} \neq 0$, e temos $y = \frac{\lambda - a_{11}}{a_{12}} x$. Tomando $x = 1$, resultam os autovetores como no enunciado.

■

:⊢ **Definição 24.20.** Um *ponto racional* de $\mathbb{T}^2$ é um ponto $[r] \in \mathbb{T}^2$ tal que $r \in \mathbb{Q}^2$.

De fato, esses pontos estão bem definidos pois, para todo $s \in [r]$, temos $s = r + m$, com $m \in \mathbb{Z}^2$. Logo temos $y \in \mathbb{Q}^2$, já que a soma de um racional com um inteiro ainda é racional.

⊢ **Proposição 24.15.** *O conjunto $Per([A])$ é o conjunto dos pontos racionais de $\mathbb{T}^2$.*

□ *Demonstração.* Seja r um ponto racional de $\mathbb{T}^2$ da forma $r := [(\frac{a}{k}, \frac{b}{k})]$, em que a, b e k são naturais, k não nulo. Notemos que, ao aplicarmos $[A]$ em r, temos

$$[A](r) = \left[\left(\frac{(a_{11}a + a_{12}b)}{k}, \frac{(a_{21}a + a_{22}b)}{k}\right)\right],$$

que é da mesma forma que r, pois $(a_{11}a + a_{12}b)$ e $(a_{21}a + a_{22}b)$ são inteiros. Assim concluímos que pontos da forma $[(\frac{a}{k}, \frac{b}{k})]$ continuam dessa forma após aplicarmos $[A]$ e, de fato, $[A]$ simplesmente permuta pontos dessa forma. Como, para termos classes de equivalência diferentes, há k escolhas de valores tanto para a como para b, existem k^2 pontos do toro dessa forma. Pelo princípio das gavetas, deve haver inteiros i e j tais que $[A]^i(r) = [A]^j(r)$ e $|i - j| \leq k^2$. Sem perda de generalidade, supomos $j \geq i$ e, aplicando $[A]^{-i}$ ao dois lados dessa igualdade, temos $r = [A]^{j-i}(r)$; ou seja, r tem órbita periódica de período menor ou igual a k^2. ■

Capítulo 25

Sistemas dinâmicos lineares

:⊢ **Definição 25.1.** Um *sistema dinâmico linear* é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é um espaço linear, (X, f) é um sistema dinâmico e f é uma ação linear em $\boldsymbol{X}$.

Estudaremos somente sistemas dinâmicos lineares em espaços normados de dimensão finita.

25.1 Sistemas dinâmicos lineares hiperbólicos

:⊢ **Definição 25.2.** Um *sistema dinâmico linear hiperbólico* é um sistema dinâmico linear revertível $\mathcal{S} = (\boldsymbol{E}, f)$, em que $\boldsymbol{E}$ é um espaço normado de dimensão finita[1], para o qual existem $c \in]0, \infty[$, $\lambda \in]0, 1[$ e uma decomposição

$$E = E^+ \oplus E^-$$

em subespaços lineares invariantes E^+ e E^- de E tais que, para todo $t \in \mathbb{I}^+$,

$$\left\| f^t|_{E^+} \right\| \leq c\lambda^t \qquad \text{e} \qquad \left\| f^{-t}|_{E^-} \right\| \leq c\lambda^t.$$

O espaço E^+ é o subespaço *estável*[2] e o espaço E^- é o subespaço *instável* do sistema.

Note que, como $c \in]0, \infty[$ e $\lambda \in]0, 1[$, então $c' := \log(c) \in]-\infty, \infty[$ e $\lambda' := \log(\lambda) \in]-\infty, 0[$, logo $c\lambda^t = \exp(c' + \lambda' t) = \mathrm{e}^{c' + \lambda' t}$, o que explicita o fato de que as normas vão a zero exponencialmente. Ainda, note que E^+ ou E^- podem ser $\{0\}$. Além disso, como as normas são equivalentes em $\boldsymbol{E}$, a definição de hiperbolicidade não depende da norma.

[1] Mais geralmente, espaços normados completos podem ser considerados.

[2] Aqui, os superíndices podem novamente gerar confusão. A mesma definição para os conjuntos estável e instável é usada.

⊢ **Proposição 25.1.** *Seja* $\mathcal{S} = (\boldsymbol{E}, f)$ *um sistema dinâmico linear revertível. O sistema* $\mathcal{S}$ *é hiperbólico se, e somente se, os autovalores de* f *têm valor absoluto diferente de* 1.

As proposições a seguir são enunciadas para E^+, mas proposições análogas valem para E^-.

⊢ **Proposição 25.2.** *Seja* $\mathcal{S} = (\boldsymbol{E}, f)$ *um sistema dinâmico linear hiperbólico.*

1. *Para todo* $t \in \mathbb{I}^+$ *e todo* $v \in E^+$
$$c^{-1}\lambda^{-t}\|v\| \leq \left\|f^{-t}(v)\right\|.$$
2. *Para todo* $v \in E^+$*, quando* $t \to \infty$ *tem-se*
$$\left\|f^{t}(v)\right\| \to 0 \qquad e \qquad \left\|f^{-t}(v)\right\| \to \infty.$$

□ *Demonstração.* 1. Mostraremos primeiro que, para todo $t \in \mathbb{I}^+$, todo $t' \in \mathbb{I}$ e todo $v \in E^+$,
$$\left\|f^{t+t'}(v)\right\| \leq c\lambda^t \left\|f^{t'}(v)\right\|.$$
Como E^+ é invariante, para todo $t' \in \mathbb{I}$ tem-se $f^{t'}(v) \in E^+$, portanto
$$\left\|f^{t+t'}(v)\right\| = \left\|f^{t}(f^{t'}(v))\right\| \leq c\lambda^t \left\|f^{t'}(v)\right\|.$$
Tomando $t' = -t$, segue que $\|v^+\| \leq c\lambda^t \|f^{-t}(v)\|$, logo da positividade de c e λ temos que
$$c^{-1}\lambda^{-t}\|v\| \leq \left\|f^{-t}(v)\right\|.$$
2. Como $\lambda \in]0,1[$, temos que, quando $t \to \infty$, $\lambda^t \to 0$ e $\lambda^{-t} \to \infty$, logo
$$\left\|f^{t}(v)\right\| \to 0 \qquad \text{e} \qquad \left\|f^{-t}(v)\right\| \to \infty. \qquad \blacksquare$$

Definimos as projeções $\mathrm{p}_+ \colon E \to E^+$ e $\mathrm{p}_- \colon E \to E^-$ e, para cada $v \in E$, $v_+ := \mathrm{p}_+(v)$ e $v_- := \mathrm{p}_-(v)$, de modo que $v = v_+ + v_-$.

25.1.1 Cones

:⊢ **Definição 25.3.** Seja $\mathcal{S} = (\boldsymbol{E}, f)$ um sistema dinâmico linear hiperbólico. Para todo $\varepsilon \in]0, \infty[$, o *$\varepsilon$-cone* ao redor de E^+ é o conjunto
$$C_\varepsilon(E^+) := \{v \in E \mid \|\mathrm{p}_-(v)\| \leq \varepsilon \|\mathrm{p}_+(v)\|\}$$
e o *ε-cone* ao redor de E^- é o conjunto
$$C_\varepsilon(E^-) := \{v \in E \mid \|\mathrm{p}_+(v)\| \leq \varepsilon \|\mathrm{p}_-(v)\|\}.$$

Vale notar que, se definíssemos os cones para $\varepsilon = 0$, teríamos $C_0(E^+) = E^+$ e $C_0(E^-) = E^-$. Também por isso, temos que $E^+ \subseteq C_\varepsilon(E^+)$ e $E^- \subseteq C_\varepsilon(E^-)$ para todo ε. Ainda, obtemos a seguinte proposição.

$\vdash$ **Proposição 25.3.** *Seja $\mathbf{S} = (\boldsymbol{E}, f)$ um sistema dinâmico linear hiperbólico. A decomposição hiperbólica $E = E^+ + E^-$ é única e*

$$E^+ = \left\{ v \in E \;\middle|\; \exists_{\varepsilon \in]0,\infty[} \forall_{t \in \mathbb{I}^+} f^t(v) \in C_\varepsilon(E^+) \right\}$$

e

$$E^- = \left\{ v \in E \;\middle|\; \exists_{\varepsilon \in]0,\infty[} \forall_{t \in \mathbb{I}^+} f^{-t}(v) \in C_\varepsilon(E^-) \right\}.$$

A constante c da definição de hiperbolicidade pode ser omitida se mudarmos a norma de $\boldsymbol{E}$ para uma norma adequada.

$\vdash$ **Proposição 25.4.** *Seja $\mathbf{S} = (\boldsymbol{E}, f)$ um sistema dinâmico linear hiperbólico, $\|\cdot\|$ a norma de E. Existem norma $\|\cdot\|'$ de E e $\lambda' \in]0,1[$ tais que, para todo $t \in \mathbb{I}^+$,*

$$\left\| f^t|_{E^+} \right\|' \leq (\lambda')^t \qquad e \qquad \left\| f^{-t}|_{E^-} \right\|' \leq (\lambda')^t.$$

Essa norma é chamada de *norma adaptada* do sistema (E, f).

Capítulo 26

Sistemas dinâmicos topológicos

:⊢ **Definição 26.1.** Um *sistema dinâmico topológico* é uma dupla $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é um espaço topológico, (X, f) é um sistema dinâmico e f é uma ação contínua em $\boldsymbol{X}$.

26.1 Conjugação topológica

:⊢ **Definição 26.2.** Sejam $\boldsymbol{\mathcal{S}_1} = (\boldsymbol{X_1}, f_1)$ e $\boldsymbol{\mathcal{S}_2} = (X_2, f_2)$ sistemas dinâmicos topológicos. Uma *semiconjugação topológica* (morfismo de sistemas dinâmicos topológico) de $\boldsymbol{\mathcal{S}_1}$ para $\boldsymbol{\mathcal{S}_2}$ é uma semiconjugação $\phi : \mathcal{S}_1 \to \mathcal{S}_2$ que é uma função contínua de $\boldsymbol{X_1}$ para $\boldsymbol{X_2}$. Denota-se $\phi : \boldsymbol{\mathcal{S}_1} \to \boldsymbol{\mathcal{S}_2}$. Uma *conjugação topológica* (isomorfismo de sistemas dinâmicos topológicos) de $\boldsymbol{\mathcal{S}_1}$ para $\boldsymbol{\mathcal{S}_2}$ é uma semiconjugação topológica $\phi : \boldsymbol{\mathcal{S}_1} \to \boldsymbol{\mathcal{S}_2}$ que é um homeomorfismo de $\boldsymbol{X_1}$ para $\boldsymbol{X_2}$.

26.2 Transitividade topológica e minimalidade

⊢ **Proposição 26.1.** *Seja $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico. São equivalentes*

1. *Não existe partição $X = P \cup P'$ cujas partes têm interior vazio e umas delas é positivamente invariante;*
2. *Para todos abertos não vazios $A, A' \subseteq X$, existe $t \in \mathbb{Z}_{\geq 0}$ tal que $f^t(A) \cap A' \neq \emptyset$;*
3. *Para todo aberto não vazio $A \subseteq X$,*
$$\overline{\mathcal{O}^+(A)} = X;$$
4. *Para todo aberto não vazio $A \subseteq X$,*
$$\overline{\mathcal{O}^-(A)} = X.$$

Essas propriedades são importantes e recebem o nome de transitividade topológica.

:⊢ **Definição 26.3.** Um sistema dinâmico *topologicamente transitivo* é um sistema dinâmico topológico $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ tal que, para todos abertos não vazios $A, A' \subseteq X$, existe $t \in \mathbb{Z}_{\geq 0}$ tal que $f^t(A) \cap A' \neq \emptyset$.

⊢ **Proposição 26.2.** *Seja $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico revertível. Então $(\boldsymbol{X}, f)$ é topologicamente transitivo se, e somente se, $(\boldsymbol{X}, f^{-1})$ o é.*

□ *Demonstração.* Segue diretamente das últimas equivalências da proposição anterior. ■

⊢ **Proposição 26.3.** *Sejam $\boldsymbol{X}$ um espaço métrico sem pontos isolados e $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico.*

1. *Se $x \in X$ tem órbita densa, então, para todo $t \in \mathbb{Z}_{\geq 0}$, $f^t(x)$ também tem;*
2. *Se existe $x \in X$ cuja órbita é densa em X, então $\boldsymbol{\mathcal{S}}$ é topologicamente transitivo.*

A propriedade de órbita densa nem sempre é equivalente à transitividade topológica, e recebe o nome de transitividade por órbitas.

:⊢ **Definição 26.4.** Um sistema dinâmico *transitivo por órbitas* é um sistema dinâmico topológico $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ em que existe ponto $x \in X$ cuja órbita é densa em X.

Em um caso clássico, pode-se mostrar que transitividade por órbitas implica transitividade topológica. Esse teorema foi provado por Birkhoff para $\mathbb{R}^d$, mas pode ser enunciado como segue.

⊢ **Proposição 26.4.** *Sejam $\boldsymbol{X}$ um espaço métrico completo separável e $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico. Se $\boldsymbol{\mathcal{S}}$ é topologicamente transitivo, então é transitivo por órbitas.*

:⊢ **Definição 26.5.** Um sistema dinâmico *mínimo* é um sistema dinâmico topológico $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ cujos pontos têm órbita densa em X.

⊢ **Proposição 26.5.** *O sistema de rotação $(\mathbb{T}^1, R_\alpha)$ é mínimo se α é irracional.*

⊢ **Proposição 26.6.** *Seja $\boldsymbol{G}$ um grupo topológico, $\alpha \in G$ e $E_\alpha \colon G \to G$ a translação à esquerda por α. Se o sistema $(\boldsymbol{G}, E_\alpha)$ é topologicamente transitivo, então ele é mínimo.*

26.3 Pontos errantes, conjuntos limites e pontos recorrentes

26.3.1 Pontos errantes

:⊢ **Definição 26.6.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico. Um ponto *errante* do sistema é um ponto $p \in X$ para o qual existe vizinhança $V \subseteq X$ de p satisfazendo

$$V \cap \bigcup_{t \in \mathbb{Z}_{>0}} f^t(V) = \emptyset.$$

Um ponto *não-errante* do sistema é um ponto $p \in X$ que não é errante: para toda vizinhança $V \subseteq X$ de p,

$$V \cap \bigcup_{t \in \mathbb{Z}_{>0}} f^t(V) \neq \emptyset.$$

O conjunto dos pontos não-errantes do sistema $\mathcal{S}$ é denotado $\Omega(\mathcal{S})$ (ou Ω, ou ainda $\Omega(f)$, de acordo com o que for mais relevante ressaltar).

Um ponto errante é um ponto para o qual, em alguma vizinhança, nenhuma semiórbita positiva da vizinhança retorna para a vizinhança, e um não-errante é um ponto para o qual, em toda vizinhança, alguma semiórbita positiva da vizinhança retorna para a vizinhança.

Essa propriedade é, de certa forma, um tipo de periodicidade topológica. Por isso, pode-se esperar que pontos periódicos sejam não-errantes.

⊢ **Proposição 26.7.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico. Todo ponto periódico do sistema é não-errante:*

$$\Pi(f) \subseteq \Omega(f).$$

□ *Demonstração.* Seja $p \in X$ um ponto t-periódico. Para toda vizinhança de V de p, segue que $V \cap f^t(V) \neq \emptyset$, pois $p = f^t(p) \in f^t(V)$, logo p é não-errante. ■

⊢ **Proposição 26.8.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico.*

1. *O conjunto dos pontos não-errantes Ω é fechado;*
2. *O conjunto dos pontos não-errantes Ω é positivamente invariante ($f(\Omega) \subseteq \Omega$);*
3. *Se $\mathcal{S}$ é revertível, $\Omega(f^{-1}) = \Omega(f)$ e Ω é invariante ($f^{-1}(\Omega) = \Omega$).*

□ *Demonstração.* 1. O conjunto de pontos errantes é aberto, pois se p é um ponto não-errante, existe vizinhança V de p tal que

$$V \cap \bigcup_{t \in \mathbb{Z}_{>0}} f^t(V) = \emptyset,$$

logo todo $q \in V^\circ$ também é não-errante. Isso implica que o complementar desse conjunto, o conjunto Ω, é fechado.

2. Sejam $p \in \Omega$. Para toda vizinhança V de $f(p)$, $V' := f^{-1}(V)$ é vizinhança de p. Como p é não-errante, existe $t \in \mathbb{Z}_{>0}$ tal que $V' \cap f^t(V') \neq \emptyset$. Disso segue que
$$\emptyset \neq f(V' \cap f^t(V')) \subseteq f(V') \cap f(f^t(V')) = f(V') \cap f^t(f(V')) \subseteq V \cap f^t(V),$$
portanto $f(p)$ é não-errante, o que mostra que $f(\Omega) \subseteq \Omega$.
3. Seja $p \in \Omega(f)$. Para toda vizinhança V de p, existe $t \in \mathbb{Z}_{>0}$ tal que $V \cap f^t(V) \neq \emptyset$. Assim, segue da injetividade de f(usada na segunda contenção) que
$$\emptyset \neq f^{-t}(f^t(V) \cap V) \subseteq f^{-t}(f^t(V)) \cap f^{-t}(V) \subseteq V \cap f^{-t}(V),$$
logo $p \in \Omega(f^{-1})$. A contenção contrária segue ao trocar os papéis de f e f^{-1}, portanto $\Omega(f^{-1}) = \Omega(f)$. Dessa igualdade, tem-se do item 1 que $f(\Omega) \subseteq \Omega$ e $f^{-1}(\Omega) \subseteq \Omega$. Usando essa segunda contenção, segue da sobrejetividade de f que
$$\Omega = f(f^{-1}(\Omega)) \subseteq f(\Omega),$$
portanto $\Omega = f(\Omega)$, e conclui-se da injetividade de f que Ω é invariante. ■

26.3.2 Pontos limite

:⊢ **Definição 26.7.** Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico e $p \in X$. O *conjunto limite positivo*[1] de p é o conjunto de pontos limites[2] de $\mathcal{O}^+(p)$, dado por
$$\mathrm{L}_f^+(p) := \bigcap_{t \in \mathbb{Z}_{\geq 0}} \overline{\{f^{t'}(p) \mid t' > t\}}.$$
Se $\mathcal{S}$ é revertível, então o *conjunto limite negativo*[3] de p é o conjunto $\mathrm{L}_f^-(p) := \mathrm{L}_{f^{-1}}^+(p)$.

O *conjunto limite positivo* do sistema é o conjunto
$$\mathrm{L}^+(f) := \overline{\bigcup_{p \in X} \mathrm{L}_f^+(p)},$$
o *conjunto limite negativo* é o conjunto $\mathrm{L}^-(f) := \mathrm{L}^+(f^{-1})$ e o *conjunto limite* é o conjunto
$$\mathrm{L}(f) := \mathrm{L}^+(f) \cup \mathrm{L}^-(f).$$

Ou seja, os pontos de $\mathrm{L}_f^+(p)$ são aqueles que são limites de $(f^{t_n}(p))_{n \in \mathbb{N}}$ para alguma sequência $(t_n)_{n \in \mathbb{N}} \in (\mathbb{Z}_{\geq 0})^{\mathbb{N}}$ que tende ao infinito.

[1] Chamado comumente de ω-limite e denotado $\omega_f(p)$.
[2] O conjunto de pontos limites de um conjunto é às vezes chamado de o *derivado* do conjunto.
[3] Chamado comumente de α-limite e denotado $\alpha_f(p)$.

⊢ **Proposição 26.9.** *Sejam* $\mathcal{S} = (\boldsymbol{X}, f)$ *um sistema dinâmico topológico e* $p \in X$*. Então*

1. *O conjunto limite positivo* $\mathrm{L}^+(p)$ *é não-vazio, fechado e positivamente invariante*

⊢ **Proposição 26.10.** *Seja* $\mathcal{S} = (\boldsymbol{X}, f)$ *um sistema dinâmico topológico. Todo ponto periódico do sistema é um ponto limite e todo ponto limite é um ponto não-errante:*

$$\Pi(f) \subseteq \mathrm{L}(f) \subseteq \Omega(f).$$

□ *Demonstração.* Para a primeira inclusão, seja $p \in X$ um ponto periódico. Então existe $t \in \mathbb{Z}_{\geq 0}$ tal que $f^t(p) = p$. Tomando a sequência $(nt)_{n\in\mathbb{N}}$, segue que $f^{nt}(p) = p$, logo $f^{nt}(p) \to p$, o que mostra que $p \in \mathrm{L}(f)$.

Agora, seja $q \in \mathrm{L}(f)$. Consideraremos o caso em que $q \in \mathrm{L}^+(f)$. Existe $p \in X$ tal que $q \in \mathrm{L}^+_f(p)$. Assim, para toda vizinhança V de q, existem $t, t' \in \mathbb{Z}_{\geq 0}$, $t' > t$, tais que $f^t(p) \in V$ e $f^{t'}(p) \in V$. Como $f^{t'-t}(f^t(p)) = f^t(p)$, isso implica que

$$V \cap f^{t'-t}(V) \neq \emptyset,$$

logo $q \in \Omega(f)$. ■

⊢ **Proposição 26.11.** *Seja* $\mathcal{S} = (\boldsymbol{X}, f)$ *um sistema dinâmico topológico.*

1. *O conjunto dos pontos limite* $\mathrm{L}(f)$ *é fechado;*
2. *O conjunto dos pontos limite* L *é positivamente invariante (*$f(\mathrm{L}) \subseteq \mathrm{L}$*);*
3. *Se* $\mathcal{S}$ *é revertível,* L *é invariante (*$f^{-1}(\mathrm{L}) = \mathrm{L}$*).*

26.3.3 Pontos recorrentes

:⊢ **Definição 26.8.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico. Um ponto *(positivamente) recorrente* é um ponto $p \in X$ tal que $p \in \mathrm{L}^+(p)$. O conjunto dos pontos recorrentes do sistema é denotado $\mathrm{R}(f)$.

Sejam X um espaço topológico com base enumerável $(U_i)_{i\in\mathbb{N}}$ e $f : X \to X$ uma função contínua. O conjunto dos pontos de U_i que retornam para U_i finitas vezes é

$$\tilde{U}_i := U_i \cap \varliminf_{n\in\mathbb{N}} f^{-n}({U_i}^{\complement}).$$

O limite inferior é o conjunto de pontos do espaço que passam por U_i finitas vezes, que é dado por

$$\varliminf_{n\in\mathbb{N}} f^{-n}({U_i}^{\complement}) = \bigcup_{n\in\mathbb{N}} \bigcap_{m\in\mathbb{N}} f^{-n+m}({U_i}^{\complement}).$$

Esse conjunto é o complementar do conjunto de pontos do espaço que passam por U_i infinitas vezes, dado pelo limite superior

$$\overline{\lim_{n\in\mathbb{N}}} f^{-n}(U_i) = \bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} f^{-n+m}(U_i).$$

Agora, notemos que o complementar da união dos $\tilde{U}_i$ é o conjunto $\mathrm{R}(f)$ dos pontos recorrentes de f. Para isso, primeiro notemos que

$$\begin{aligned} R &:= \left(\bigcup_{i\in\mathbb{N}} \tilde{U}_i\right)^{\complement} \\ &= \bigcap_{i\in\mathbb{N}} \tilde{U}_i^{\complement} \\ &= \bigcap_{i\in\mathbb{N}} \left(U_i \cap \underline{\lim}_{n\in\mathbb{N}} f^{-n}(U_i^{\complement})\right)^{\complement} \\ &= \bigcap_{i\in\mathbb{N}} \left(U_i^{\complement} \cup \overline{\lim_{n\in\mathbb{N}}} f^{-n}(U_i)\right). \end{aligned}$$

$\vdash$ **Proposição 26.12.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico, $\boldsymbol{X}$ com base enumerável $(U_i)_{i\in\mathbb{N}}$.*

$$\mathrm{R}(f) = \bigcap_{i\in\mathbb{N}} \left(U_i^{\complement} \cup \overline{\lim_{n\in\mathbb{N}}} f^{-n}(U_i)\right).$$

$\square$ *Demonstração.* ($R \subseteq \mathrm{R}(f)$) Sejam $x \in R$ e U vizinhança de x. Como $(U_i)_{i\in\mathbb{N}}$ é base, seja U_i tal que $x \in U_i \subseteq U$. Como $x \in R$, então $x \notin \tilde{U}_i$, portanto x retorna a U_i infinitas vezes, o que mostra que $x \in \mathrm{R}(f)$.

($\mathrm{R}(f) \subseteq R$) Sejam $x \in \mathrm{R}(f)$ e U_i aberto básico. Existem dois casos: $x \notin U_i$ ou $x \in U_i$. No segundo caso, como x é ponto recorrente, x retorna infinitamente a U_i, portanto $x \in \overline{\lim}_{n\in\mathbb{N}} f^{-n}(U_i)$. Assim, em ambos os casos, $x \in U_i^{\complement} \cup \overline{\lim}_{n\in\mathbb{N}} f^{-n}(U_i)$, o que mostra que $x \in R$. ■

26.4 Conjuntos estável e instável

$\vdash$ **Definição 26.9.** Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico e $C \in X$ um conjunto positivamente invariante. O *conjunto estável* de C é o conjunto

$$W_f^+(C) := \left\{p \in X \;\middle|\; f^t(p) \to C,\ t \to \infty\right\}.$$

Se o sistema é revertível, o *conjunto instável* de C é $W_f^-(C) := W_{f^{-1}}^+(C)$. Quando $p \in X$ é um ponto fixo, denotam-se $W_f^+(p) := W_f^+(\{p\})$ e $W_f^-(p) := W_f^-(\{p\})$.

Os superíndices '+' e '−' podem gerar certa confusão, pois o índice '+' poderia tanto indicar que a *distância* entre os pontos q e o ponto p *aumenta* ou que no tempo *futuro* os pontos q tendem ao ponto p; problema análogo ocorre para o índice '− '. Escolhemos aqui a segunda opção porque ela segue o padrão que de outras notações, em que '+' indica a evolução positiva (futura) do sistema e '−' indica sua evolução negativa (passada).

⊢ **Proposição 26.13.** *Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico e $C \subseteq X$ um conjunto positivamente invariante. Então $W^+(C)$ é invariante.*

□ *Demonstração.* Seja $p \in W^+(C)$. Para todo $s \in T^+$, temos que $f^t(f^s(p)) = f^{t+s}(p) \to C$ quando $t \to \infty$, portanto $f^s(p) \in W^+(C)$, o que mostra que o conjunto é positivamente invariante. Para mostrar que é negativamente invariante, basta notar que se $q \in f^{-s}(p)$, então $f^s(q) = p$, logo $f^t(q) \to C$. ■

26.5 Exemplos

26.5.1 Topologia dos deslocamentos

Queremos definir uma topologia para os espaços de sequências $N^{\mathbb{Z}_{\geq 0}}$ e $N^{\mathbb{Z}}$ de modo que os deslocamentos σ neles definidos sejam contínuos. Para isso, vamos induzir o espaço produto $N^{\mathbb{Z}}$ com a topologia discreta de N, pois essa é uma abordagem bem natural. A princípio, não sabemos se isso garantirá que o deslocamento σ é contínuo, mas veremos que ele de fato será. Construiremos a seguir a topologia de $N^{\mathbb{Z}}$, mas o caso de $N^{\mathbb{Z}_{\geq 0}}$ é análogo.

A topologia discreta em N é a topologia em que todo subconjunto de N é aberto. Uma base para essa topologia é o conjunto dos conjuntos da forma $\{k\}$, com $k \in N = \{0, \ldots, N-1\}$. Essa topologia é o conjunto $\wp(N)$. A topologia induzida em $N^{\mathbb{Z}}$ é a topologia puxada pelas projeções na t-ésima entrada $\pi_t : N^{\mathbb{Z}} \to N$ (também conhecida como topologia produto ou topologia inicial com respeito às projeções), ou seja,

$$\left\langle \bigcup_{t \in \mathbb{Z}} {\pi_t}^{\star}(\wp(N)) \right\rangle.$$

Uma sub-base da topologia produto de $N^{\mathbb{Z}}$ são os conjuntos

$$C_t[k] := {\pi_t}^{-1}(\{k\}) = \left\{ x \in N^{\mathbb{Z}} \;\middle|\; x_t = k \right\}$$

para cada $t \in \mathbb{Z}$ e $k \in N$ (chamados às vezes de *cilindros abertos*). Uma base para essa topologia são interseções finitas de conjuntos da sub-base, ou seja, conjuntos da forma

$$C_{t_0,\ldots,t_n}(k_0, \ldots, k_n) := \bigcap_{i=0}^{n} C_{t_i}[k_i] = \left\{ x \in N^{\mathbb{Z}} \;\middle|\; x_{t_0} = k_0, \ldots, x_{t_n} = k_n \right\}$$

para cada $t_0, \ldots, t_n \in \mathbb{Z}$ e $k_0, \ldots, k_n \in N$, chamados de *cilindros*.

Estando assim definida a topologia de $N^{\mathbb{Z}}$, vamos agora mostrar que σ é contínua. Seja $C_t[k]$ um cilindro sub-básico. Então

$$\sigma^{-1}(C_t[k]) = \sigma^{-1}({\pi_t}^{-1}(\{k\})) = (\pi_t \circ \sigma)^{-1}(\{k\}) = {\pi_{t-1}}^{-1}(\{k\}) = C_{t-1}[k],$$

que é aberto pois é um cilindro sub-básico. Isso mostra que σ é contínua. Está provado, assim, a seguinte proposição.

⊢ **Proposição 26.14.** *Os sistemas* $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ *e* $(N^{\mathbb{Z}}, \sigma)$ *são sistemas dinâmicos topológicos.*

26.5.2 Itinerários topológicos

Lembremos que se $\mathcal{P} = (P_i)_{i \in N}$ é uma partição finita de X, então o itinerário de um ponto $x \in X$ com respeito a $\mathcal{P}$ é a sequência $(x_t)_{t \in T} \in N^T$ tal que $x_t = k$ se, e somente se, $f^t(x) \in P_k$. Denotamos a função que leva x em seu itinerário por $\phi : X \to N^T$. Queremos investigar as condições em $\mathcal{P}$ para que ϕ seja uma função contínua.

Notemos que $x_t = \pi_t(\phi(x))$, portanto podemos escrever que $\pi_t(\phi(x)) = k$ se, e somente se, $f^t(x) \in P_k$. Isso é o mesmo que dizer que

$$(\pi_t \circ \phi)^{-1}(k) = f^{-t}(P_k).$$

Essa identidade é a chave para a demonstração da proposição a seguir.

⊢ **Proposição 26.15.** *Sejam* $(\boldsymbol{X}, f)$ *um sistema dinâmico topológico discreto,* $\mathcal{P} = (P_i)_{i \in N}$ *uma partição aberta finita de* $\boldsymbol{X}$ *e* $\phi : X \to N^T$ *a função que leva* $x \in X$ *em seu itinerário com respeito a* $\mathcal{P}$*. Então* ϕ *é uma função contínua.*

□ *Demonstração.* Para isso, seja $C_t[k]$ um cilindro sub-básico de $N^{\mathbb{T}}$. Nesse caso, segue que

$$\phi^{-1}(C_t[k]) = \phi^{-1}({\pi_t}^{-1}(k)) = (\pi_t \circ \phi)^{-1}(k) = f^{-t}(P_k),$$

que é aberto pois f é contínua e P_k é aberto. ■

Note que se existe uma partição aberta não trivial ($\mathcal{P} = (X)$) de $\boldsymbol{X}$, o espaço $\boldsymbol{X}$ é não conexo, pois as partes das partições são abertas e disjuntas. Em particular isso não é muito desejável, pois queremos estudar espaços topológicos conexos, nos quais por definição não existe partição aberta. Para que essa dificuldade seja superada, consideraremos mais a frente outro tipo de partição, as partições de medida, em que a exigência de que as partes sejam disjuntas será enfraquecida para que sejam quase disjuntas.

26.5.3 Automorfismos no toro

Após essa discussão inicial sobre o sistema $(\mathbb{T}^2, [A])$, seguimos para abordar a ideia de variedades estável e instável de um ponto do toro. Para um ponto p do toro, definimos a variedade estável de p como a projeção no toro da reta em $\mathbb{R}^2$ que passa por p e é paralela ao autovetor v_{λ_-} de A associado ao autovalor $\lambda_- < 1$. Analogamente, definimos a variedade instável de p como a projeção no toro da reta em $\mathbb{R}^2$ que passa por p e é paralela ao autovetor v_{λ_+} de A associado ao autovalor $\lambda_+ > 1$. Segue a definição rigorosa.

:⊢ **Definição 26.10.** Seja $[p] \in \mathbb{T}^2$. A *variedade estável* de $[p]$ é o conjunto

$$W^-([p]) := \left\{ [p + \alpha v_{\lambda_-}] \;\middle|\; \alpha \in \mathbb{R} \right\},$$

Analogamente, a *variedade instável* de $[p]$ é o conjunto

$$W^+([p]) := \left\{ [p + \alpha v_{\lambda_+}] \;\middle|\; \alpha \in \mathbb{R} \right\}.$$

Note que as variedades estável e instável estão bem definidas pois, para todo $q \in [p]$, temos $q = p+m$, com $m \in \mathbb{Z}^2$. Assim, $[q+\alpha v_{\lambda_-}] = [p+m+\alpha v_{\lambda_-}] = [p+\alpha v_{\lambda_-}]$ e $[q + \alpha v_{\lambda_+}] = [p + m + \alpha v_{\lambda_+}] = [p + \alpha v_{\lambda_+}]$, e segue que $W^-([q]) = W^-([p])$ e $W^+([q]) = W^+([p])$. A próxima proposição justifica a nomeação dos conjuntos $W^-([p])$ e $W^+([p])$.

⊢ **Teorema 26.16.** *Seja $[p]$ um ponto do toro. Então, para um ponto $[p'] \in W^s([p])$,*

$$\lim_{n\to\infty} d_{\mathbb{T}^2}([A]^n([p']), [A]^n([p])) = 0.$$

Da mesma forma, para um ponto $[p'] \in W^u([p])$,

$$\lim_{n\to\infty} d_{\mathbb{T}^2}([A]^{-n}([p']), [A]^{-n}([p])) = 0.$$

□ *Demonstração.* Vamos demonstrar a proposição para a variedade instável. A outra demonstração é análoga. Se $[p'] \in W^s([p])$, ele é da forma $[p'] = [p + \alpha v_{\lambda_-}]$, para $\alpha \in \mathbb{R}$. Assim,

$$[A]([p']) = [A]([p + \alpha v_{\lambda_-}]) = [A^n(p + \alpha v_{\lambda_-})] = [A^n p + \alpha(\lambda_-)^n v_{\lambda_-}]$$

e $[A]^n([p]) = [A^n p]$. Notemos que, como $|\lambda_-| \leq 1$, para n suficientemente grande a distância no toro é igual à distância no plano, pois $|\lambda_-|^n \to 0$. Assim, vale

$$d_{\mathbb{T}^2}([A^n p + \alpha(\lambda_-)^n v_{\lambda_-}], [A^n p]) = d_{\mathbb{R}^2}(A^n p + \alpha(\lambda_-)^n v_{\lambda_-}, A^n p) = d_{\mathbb{R}^2}(\alpha(\lambda_-)^n v_{\lambda_-}, 0)$$

e segue o enunciado. ■

⊢ **Proposição 26.17.** *As variedades estável e instável de um ponto* $[p]$ *do toro são invariáveis sob ação do automorfismo* $[A]$ *no sentido que valem* $[A](W^s([p])) = W^s([A]([p]))$ *e* $[A](W^u([p])) = W^u([A]([p]))$.

□ *Demonstração.* Na demonstração, usaremos W para representar qualquer uma das duas variedades estável e instável. Da mesma forma, v para o autovetor e λ para o autovalor associado à respectiva variedade.

Primeiro demonstraremos que

$$[A](W([p])) \subset W([A]([p])).$$

Seja $[x] \in [A](W([p]))$. Então existe $\alpha \in \mathbb{R}$ tal que $[x] = [A]([p + \alpha v])$. Logo

$$[x] = [A(p + \alpha v)] = [Ap + \alpha Av] = [Ap + \alpha\lambda v].$$

Fazendo $\beta := \alpha\lambda$, temos $\beta \in \mathbb{R}$ e $[x] = [Ap + \beta v] = [[A]([p]) + \beta v]$. Portanto $[x] \in W([A]([p])$.

Agora demonstramos que

$$W([A]([p])) \subset [A](W([p])).$$

Tomamos $[x] \in W([A]([p]))$. Então existe $\alpha \in \mathbb{R}$ tal que $[x] = [[A]([p]) + \alpha v]$. Como $\lambda \neq 0$, fazemos $\beta := \frac{\alpha}{\lambda}$. Então temos $\beta \in \mathbb{R}$ e

$$[x] = [Ap + \beta\lambda v] = [Ap + \beta Av] = [A(p + \beta v)] = [A]([p + \beta v])$$

e, portanto, $x \in [A](W^s(p))$. ■

⊢ **Proposição 26.18.** *As variedades estável e instável de um ponto do toro são densas em* $\mathbb{T}^2$.

□ *Demonstração.* Seja $[p] = [(p_1, p_2)]$ um ponto de $\mathbb{T}^2$. Na demonstração, usaremos $W([p])$ para representar qualquer uma das duas variedades de $[p]$. Queremos mostrar que, dado $\varepsilon > 0$ e $[t] = [(t_1, t_2)] \in \mathbb{T}^2$, existe $[w] = [w_1, w_2] \in W([p])$ tal que $d_{\mathbb{T}^2}([t], [w]) \leq \varepsilon$. Sem perda de generalidade, consideraremos que $p, t, w \in S^1 \times S^1$ para facilitar a demonstração.

Consideremos o conjunto $C := \{[(t_1, r)] : r \in \mathbb{R}\}$. Esse conjunto é um subconjunto de $\mathbb{T}^2$ e auxiliará na demonstração. Acharemos o ponto $[w]$ nesse conjunto, considerando o conjunto $W([p]) \cap C$. Esse conjunto interseção é interessante pois a distância entre seus potos pode ser calculada usando a distância do círculo. Tomemos $[c] \in C$, com $c = (t_1, r) \in S^1 \times S^1$ sem perda de generalidade. Notamos

que

$$\begin{aligned}
d_{\mathbb{T}^2}([t],[c]) &= \min\left\{d_{\mathbb{R}^2}(t+m,c+n) \;\middle|\; m,n\in\mathbb{Z}^2\right\}\\
&= \min\left\{\left((t_1+m_1-t_1-n_1)^2+(t_2+m_2-r-n_2)^2\right)^{\frac{1}{2}}\right|\\
&\qquad |m_1,m_2,n_1,n_2\in\mathbb{Z}\}\\
&= \min\left\{((t_2-r-z)^2)^{\frac{1}{2}} \;\middle|\; z\in\mathbb{Z}\right\}\\
&= \min\{|t_2-r-z| \mid z\in\mathbb{Z}\}\\
&= d_{S^1}(t_2,r),
\end{aligned}$$

pois $t_2, r \in S^1$. Isso mostra que a distância de um ponto de C ao ponto $[t]$ é igual à distância de suas segundas coordenadas no círculo.

Agora, definamos $l := \frac{\lambda - a_{11}}{a_{12}}$, sendo λ o autovalor λ_+ ou λ_-, de acordo com qual variedade estamos considerando, e seja $\alpha \in \mathbb{R}$. Os pontos de $W([p]) \cap C$ são da forma

$$[p+\alpha v_\lambda] = [(p_1+\alpha, p_2+\alpha l)],$$

de modo que $p_1+\alpha = t_1+k$, com $k\in\mathbb{Z}$; ou seja, definindo $m := p_2-(p_1-t_1)l \bmod 1$, temos pontos da forma

$$[(t_1, m+kl)].$$

Isso significa que os pontos de $W([p]) \cap C$ podem ser obtido através de uma rotação irracional R_l aplicada em $m \in S^1$. Como sabemos pelo teorema **??** que as rotações irracionais são densas em S^1 para qualquer ponto do círculo, dados $\varepsilon > 0$ e $t_2 \in S^1$, existe inteiro a tal que

$$d_{S^1}(t_2, m+al) = d_{S^1}(t_2, R_l^a(m)) \le \varepsilon.$$

Por fim, tomando $[w] = [t_1, m+al]$, temos $[w] \in W([p]) \cap C$; ou seja, $[w] \in W([p])$, vale

$$d_{\mathbb{T}^2}([t],[w]) = d_{S^1}(t_2, m+al) \le \varepsilon$$

e está provada a proposição. ■

⊢ **Teorema 26.19.** *A dinâmica do sistema* $(\mathbb{T}^2, [A])$ *é topologicamente transitiva.*

□ *Demonstração.* Para demonstrar isso, tomemos dois conjuntos não-vazios abertos quaisquer no toro, U e V. Agora, como a variedade estável é densa no toro, pegamos um ponto $u \in U$ e tomamos um segmento de reta I_U passando por u que esteja contido de U e seja paralelo a v_{λ_-}. Da mesma forma, pegamos um ponto $v \in V$ e tomamos um segmento de reta I_V passando por v que esteja contido em V e seja paralelo a v_{λ_+}. Como as variedades estável e instável são densas em $\mathbb{T}^2$ e o comprimento dos intervalo em U e V ambos aumentam ao aplicarmos neles,

respectivamente, $[A]^-1$ e $[A]$, existem $p, q \in \mathbb{N}$ tais que $[A]^{-p}(I_U) \cap [A]^q(I_V) \neq \emptyset$. Assim, tomamos um ponto w' nesse conjunto. Segue que $[A]^{-q}(w') \in V$ e $[A]^p(w') \in U$. Definindo $w := [A]-q(w')$ e tomando $n = p + q$, temos que $w \in [A]^{-n}(V) \cap U$ e está provado o teorema. ■

Entropia

:⊢ **Definição 26.11.** A *entropia* de um sistema dinâmico (X, f) é definida como

$$h_f := \lim_{n\to\infty} \frac{\#\left(Per_n(f)\right)}{n}$$

em que $Per_n(f) := \{x \in X : f^n(x) = x\}$.

⊢ **Teorema 26.20.** *A entropia de sistema* $(\mathbb{T}^2, [A])$ *associado à matriz* A *é*

$$h_{[A]} = \log|\lambda_+|.$$

□ *Demonstração.* De fato, dos resultados anteriores vem que

$$\begin{aligned} h_{[A]} &= \frac{\log|\det(A^n - I)|}{n} \\ &= \frac{\log|2 - (\lambda_+)^n - (\lambda_-)^n|}{n} \\ &= \frac{\log|(\lambda_+)^n[1 - 2(\lambda_-)^n + (\lambda_-)^{2n}]|}{n} \\ &= \log|\lambda_+| + \frac{\log|1 - 2(\lambda_-)^n + (\lambda_-)^{2n}|}{n} \end{aligned}$$

Daí, segue claramente o resultado procurado, pois $|\lambda_-| < 1$. ■

Capítulo 27

Sistemas dinâmicos métricos

:⊢ **Definição 27.1.** Um sistema dinâmico *métrico* é um par $(\mathcal{S} = (\boldsymbol{X}, f))$ em que $\boldsymbol{X} = (X, |\cdot,\cdot|)$ é um espaço métrico e $\mathcal{S} = ((X, \mathcal{T}), f)$, é um sistema dinâmico topológico.

27.1 Conjuntos estáveis e instáveis

:⊢ **Definição 27.2.** Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico métrico e $p \in X$. O *conjunto estável* de p é o conjunto

$$W^+_{\mathcal{S}}(p) := \left\{ q \in X \;\middle|\; |f^n(p), f^n(q)| \overset{n\to\infty}{\to} 0 \right\}.$$

Se o sistema é revertível, o *conjunto instável* de p é o conjunto $W^-_{(\boldsymbol{X},f)}(p) := W^+_{(\boldsymbol{X},f^{-1})}(p)$. Denota-se também $W^+_f(p)$ e $W^+(p)$.

Mostraremos na proposição seguinte que a definição de conjunto estável não depende da métrica escolhida para gerar a estrutura uniforme de X, ou seja, que se uma métrica é uniformemente equivalente à métrica de $\boldsymbol{X}$, então a variedade estável de pode ser definida com aquela métrica. No entanto, em espaços métricos compactos todas funções contínuas são uniformemente contínuas, o que implica que todas métricas topologicamente equivalente são uniformemente equivalentes, que por sua vez implica que os conjuntos só dependem da estrutura topológica do espaço, isto é, que podem ser definidos para espaços metrizáveis.

Se $\mathcal{S} = ((X, |\cdot,\cdot|), f)$ é um sistema dinâmico métrico e $|\cdot,\cdot|'$ é uma métrica sobre X uniformemente equivalente a $|\cdot,\cdot|$, então $\mathcal{S} = ((X, |\cdot,\cdot|'), f)$ também é um sistema dinâmico métrico, pois f também é contínua com relação a $|\cdot,\cdot|'$.

⊢ **Proposição 27.1.** *Sejam $\mathcal{S} = ((X, |\cdot,\cdot|), f)$ um sistema dinâmico métrico e $|\cdot,\cdot|'$ é uma métrica sobre X uniformemente equivalente a $|\cdot,\cdot|$. Para todo $p \in X$,*

$$W^+_{((X,|\cdot,\cdot|),f)}(p) = W^+_{((X,|\cdot,\cdot|'),f)}(p).$$

□ *Demonstração.* Seja $q \in W^+_{((X,|\cdot,\cdot|),f)}$. Então $|f^n(p), f^n(q)| \overset{n\to\infty}{\to} 0$. Queremos mostrar que $|f^n(p), f^n(q)|' \overset{n\to\infty}{\to} 0$. Seja $\varepsilon\,]0,\infty[$. Como $|\cdot,\cdot|$ e $|\cdot,\cdot|'$ são uniformemente equivalentes, existe $\delta \in]0,\infty[$ tal que, para todo $\in \mathbb{N}$, se $|f^n(p), f^n(q)| < \delta$ então $|f^n(p), f^n(q)|' < \varepsilon$. Como $|f^n(p), f^n(q)| \overset{n\to\infty}{\to} 0$, existe $N \in \mathbb{N}$ tal que, para todo $n \in \mathbb{N}$, $n \geq N$ implica $|f^n(p), f^n(q)| < \delta$. Mas então, para todo $n \geq N$,

$$|f^n(p), f^n(q)| < \delta \Rightarrow |f^n(p), f^n(q)|' < \varepsilon,$$

o que mostra que $|f^n(p), f^n(q)|' \overset{n\to\infty}{\to} 0$. Assim, temos que

$$W^+_{(X,|\cdot,\cdot|),f}(p) \subseteq W^+_{(X,|\cdot,\cdot|'),f}(p).$$

Analogamente, tem-se que $W^+_{((X,|\cdot,\cdot|'),f)}(p) \subseteq W^+_{((X,|\cdot,\cdot|),f)}(p)$. ■

27.2 Conjuntos estáveis e instáveis locais

:⊢ **Definição 27.3.** Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico métrico $p \in X$ e $\varepsilon \in]0,\infty[$. O *conjunto estável ε-local* de p é o conjunto

$$W^+_f(p)_\varepsilon := \{q \in X \mid \forall_{n\in\mathbb{I}^+} |f^n(p), f^n(q)| < \varepsilon\}.$$

Se o sistema é revertível, o *conjunto instável ε-local* de p é o conjunto $W^-_f(p)_\varepsilon := W^+_{f^{-1}}(p)_\varepsilon$.

$W^{+,-}_\varepsilon(p)$ é a componente conexa de $W^{+,-}(p) \cap \mathrm{B}(x;\varepsilon)$.

$W^{+,-}_\varepsilon(p)$ varia continuamente com p e $W^+_\varepsilon(p) \cap W^-_\varepsilon(p) = \{p\}$.

Existem $\varepsilon, \delta \in]0,\infty[$ tais que, para todos $q \in X$, se $|p,q| < \delta$, então $W^+_\varepsilon(p) \cap W^-_\varepsilon(q)$ tem somente um ponto. Esse ponto é definido como

$$[p,q] := W^+_\varepsilon(p) \cap W^-_\varepsilon(q).$$

Isso é chamado de estrutura de produto local.

:⊢ **Definição 27.4.** Um sistema dinâmico *(positivamente) expansivo* é um sistema dinâmico metrizável compacto $\mathcal{S} = (\boldsymbol{X}, f)$ para o qual existe $c \in]0,\infty[$ tal que, para todos $x, x' \in X$ tais que $x \neq x'$, existe $t \in \mathbb{I}^+$ tal que

$$|x, x'| > c.$$

A compacidade implica que a expansividade não depende da distância escolhida.

Essa estrutura de produto local implica que é expansivo.

27.3 Pseudo órbitas

:⊢ **Definição 27.5.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico metrizável. Uma *pseudo órbita (positiva)* do sistema é uma sequência $(x_i)_{i\in\mathbb{I}^+}$ em X para a qual existe $\varepsilon \in]0, \infty[$ tal que, para todo $i \in \mathbb{I}^+$,

$$|f(x_i), x_{i+1}| < \varepsilon.$$

:⊢ **Definição 27.6.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico metrizável. Uma *pseudo órbita* do sistema é uma sequência $x = (x_i)_{i=a}^b$ em X, com $a, b \in \mathbb{I} \cup \{-\infty, \infty\}$ e $a \leq b$, para a qual existe $\varepsilon \in]0, \infty[$ tal que, para todo $i \in [a, b-1]$,

$$|f(x_i), x_{i+1}| < \varepsilon.$$

Uma pseudo órbita *positiva* é uma pseudo órbita tal que $a \geq 0$, uma pseudo órbita *negativa* é uma pseudo órbita tal que $b \leq 0$; uma pseudo *finita* é uma pseudo órbita tal que $a, b \in \mathbb{I}$

27.4 Distância no espaços de símbolos

Uma distância pode ser definida em Σ^+ e Σ de modo a gerar a mesma topologia definida na seção anterior. Com essa distância, esses espaços são compactos.

:⊢ **Definição 27.7.** A *distância* em Σ é a função

$$\begin{aligned} d\colon \Sigma \times \Sigma &\longrightarrow [0, \infty[\\ (x, x') &\longmapsto \mathop{+}_{i\in\mathbb{I}} \frac{|x_i - x'_i|}{2^{|i|}}. \end{aligned}$$

Capítulo 28

Sistemas dinâmicos de medida

28.1 Definições gerais

:⊢ **Definição 28.1.** Um *sistema dinâmico de medida* (ou *sistema que preserva medida*) é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é um espaço de medida, (X, f) é um sistema dinâmico e f é uma ação que preserva medida em $\boldsymbol{X}$.

Como estudaremos somente o caso **discreto**, a função f é no caso um morfismo de medida, uma função que preserva a medida do espaço. Ainda, estudaremos somente **espaços de probabilidade**. Introduzimos agora noções de morfismo de sistemas dinâmicos de probabilidade. Em geral, quando tratamos da teoria da medida, podemos considerar funções que são iguais a menos de um conjunto de medida nula.

28.1.1 Quase invariância

:⊢ **Definição 28.2.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico de medida. Um conjunto *quase invariante* de $\mathcal{S}$ é um conjunto $M \in \mathcal{M}$ tal que $f^{-1}(M) \mathrel{\mathring{=}} M$.

⊢ **Proposição 28.1.** *Sejam $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico de medida e $M \in \mathcal{M}$ um conjunto quase invariante de $\mathcal{S}$. Então existe um conjunto invariante $N \in \mathcal{M}$ tal que $M \mathrel{\mathring{=}} N$.*

□ *Demonstração.* Consideremos o limite superior

$$N := \overline{\lim_{n\in\mathbb{N}}} f^{-n}(M) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} f^{-n}(M),$$

o conjunto dos pontos que pertencem a infinitos dos conjuntos $f^{-n}(M)$. Notemos que, para todo $m \in \mathbb{N}$,

$$M \triangle \bigcup_{n=m}^{\infty} f^{-n}(M) \subseteq \bigcup_{n=m}^{\infty} M \triangle f^{-n}(M),$$

pois, se x é elemento do primeiro, então (1) é elemento de M mas não da união dos $f^{-n}(M)$ ou (2) é elemento da união mas não de M. No primeiro caso, ele não é elemento de nenhum dos $f^{-n}(M)$, logo é elemento de todos $M \triangle f^{-n}(M)$ e, portanto, da união; no segundo caso, é elemento de algum $f^{-n}(M)$ e, portanto, de algum $M \triangle f^{-n}(M)$, logo da união. Ainda, para todo $n \in \mathbb{N}$ vale

$$M \triangle f^{-n}(M) \subseteq \bigcup_{i=0}^{n-1} f^{-i}(M) \triangle f^{-(i+1)}(M).$$

Para ver isso, seja $x \in M \triangle f^{-n}(M)$. Suponhamos que $x \in M$. Então não pertence a $f^{-n}(M)$. Então, se $x \notin f^{-1}(M)$, temos que $x \in M \triangle f^{-1}(M)$ e, se $x \in f^{-1}(M)$, repetimos o processo com $f^{-2}(M)$ chegando, no máximo, até a $x \in f^{-(n-1)}(M)$, o que implicará $x \in f^{-(n-1)}(M) \triangle f^{-n}(M)$; em todos os casos, x pertence à união acima. Se $x \in f^{-n}(M)$, o mesmo processo se aplica, mas na ordem inversa. Isso implica que $m(M \triangle f^{-n}(M)) = 0$, pois o termo à direita tem medida zero.

Definamos $N_m := \bigcup_{n=m}^{\infty} f^{-n}(M)$, e temos que $N = \bigcap_{m=0}^{\infty} N_m$ e, claramente,

$$N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots .$$

Mas de $m(M \cap N_m) = 0$, então, segue que $m(N \triangle M) = 0$, logo que $m(N) = m(M)$. Por fim, notando que

$$f^{-1}(N) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} f^{-(n+1)}(M) = \bigcap_{m=0}^{\infty} \bigcup_{n=m+1}^{\infty} f^{-n}(M) = N,$$

segue que N é invariante. ■

28.2 Conjugação que preserva medida

:⊢ **Definição 28.3.** Sejam $\mathcal{S}_1 = (\boldsymbol{X_1}, f_1)$ e $\mathcal{S}_2 = (\boldsymbol{X_2}, f_2)$ sistemas que preservam medida. Uma *semiconjugação de medida* de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma semiconjugação $\phi : C_1 \to C_2$ que preserva medida, definida em conjuntos mensuráveis $C_1 \in \mathcal{M}_1$ e $C_2 \in \mathcal{M}_2$ que satisfazem $f_1(C_1) \subseteq C_1$, $f_2(C_2) \subseteq C_2$ e $m_1(C_1) = m_2(C_2) = 1$. Nesse caso, o sistema $\mathcal{S}_2$ é um *fator* do sistema $\mathcal{S}_1$. Uma *conjugação de medida* de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma semiconjugação de medida que é uma conjugação.

$$\begin{tikzcd}
X_1 \arrow[r, "f_1"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\
X_2 \arrow[r, "f_2"'] & X_2
\end{tikzcd}$$

Como comentado anteriormente, podemos considerar ϕ como definida de X_1 para X_2 sem problemas, pois só não está definida em um conjunto de medida nula.

28.3 Recorrência

O resultado a seguir fala sobre o conjunto dos pontos de um conjunto M que retornam infinitamente para M. Como comentado anteriormente (24.2), considerando um espaço X e um conjunto M, esse conjunto é dado por

$$M \cap \overline{\lim_{t \in \mathbb{Z}_{\geq 0}}} f^{-t}(M).$$

Lembremos que o limite superior de uma sequência de conjuntos $(A_n)_{n \in \mathbb{N}}$ é o conjunto dos pontos que pertencem a infinitos dos conjuntos A_n. Esse conjunto é dado por

$$\overline{\lim_{n \in \mathbb{N}}} A_n := \bigcap_{N=0}^{\infty} \bigcup_{n=N}^{\infty} A_n.$$

No exemplo acima, um ponto $x \in X$ pertencer a $f^{-t}(M)$ equivale a valer $f^t(x) \in M$, ou seja, x passa por M no instante $t \in \mathbb{Z}_{\geq 0}$. Assim, pertencer ao limite superior $\overline{\lim}_{t \in \mathbb{Z}_{\geq 0}} f^{-t}(M)$ significa passar infinitas vezes por M. Se tomamos a interseção desse limite superior com M, temos os pontos de M que retornam infinitamente a M. Isso equivale a dizer que, para todo conjunto mensurável $M \in \mathcal{M}$ (com $m(M) > 0$) existe um conjunto mensurável $F \subseteq M$ com $m(F) = m(M)$ tal que, se $x \in F$, então existe uma sequência de inteiros positivos $t_0 < t_1 < \cdots$ satisfazendo que, para todo $i \in \mathbb{N}$

$$f^{t_i}(x) \in M.$$

$\vdash$ **Teorema 28.2** (Recorrência)**.** *Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida finita e $M \in \mathcal{M}$ um conjunto mensurável. Então quase todo ponto de M retorna infinitamente a M:*

$$M \cap \overline{\lim_{t \in \mathbb{Z}_{\geq 0}}} f^{-t}(M) \mathrel{\overset{\circ}{=}} M.$$

$\square$ *Demonstração.* Se $M \mathrel{\overset{\circ}{=}} \emptyset$, a proposição é evidente. Suponhamos $M \mathrel{\overset{\circ}{\neq}} \emptyset$. O conjunto $M \cap \overline{\lim}_{t \in \mathbb{Z}_{\geq 0}} f^{-t}(M)$ é conjunto dos pontos de M que retornam infinitamente para M. Queremos provar que esse conjunto tem a mesma medida de M. Consideremos os conjuntos

$$M_N := \bigcup_{n=N}^{\infty} f^{-n}(M).$$

Temos que $f^{-1}(M_N) = M_{N+1}$, e, como f preserva m, segue que $m(M_N) = m(M_{N+1})$. Isso implica que $m(M_0) = m(M_N)$ para todo $N \in \mathbb{N}$. Observe que $M_0 \supset M_1 \supset$

$M_2 \ldots$, então segue, pelo lema 15.18, que

$$m\left(\bigcap_{N=0}^{\infty} M_N\right) = \lim_{N\to\infty} m(M_N) = m(M_0).$$

Notando que $\bigcap_{N=0}^{\infty} M_N = \overline{\lim}_{t\in\mathbb{Z}_{\geq 0}} f^{-t}(M)$, consequentemente segue que

$$m(F) = m\left(M \cap \overline{\lim_{t\in\mathbb{Z}_{\geq 0}}} f^{-t}(M)\right) = m(M \cap M_0) = m(M).$$

A segunda igualdade segue de $\bigcap_{N=0}^{\infty} M_N = M_0$ quase sempre, e a terceira vem de $M \subseteq M_0$. ■

Como mencionado antes, a hipótese de que o sistema tem medida finita é fundamental para usar o lema 15.18.

⊢ **Proposição 28.3.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico, $\boldsymbol{X}$ com base enumerável. O conjunto $\mathrm{R}(f)$ dos pontos recorrentes é mensurável (com respeito à σ-álgebra gerada pela topologia).*

□ *Demonstração.* Basta notar que a igualdade da proposição 26.12 é uma interseção enumerável de conjuntos mensuráveis. ■

⊢ **Teorema 28.4** (Recorrência Topológica)**.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico topológico de medida. Se a topologia de $\boldsymbol{X}$ tem base enumerável, quase todo ponto é recorrente:*

$$\mathrm{R}(f) \mathrel{\overset{\circ}{=}} X.$$

□ *Demonstração.* Seja $(A_i)_{i\in\mathbb{N}}$ uma base enumerável da topologia de $\boldsymbol{X}$. Pelo teorema 28.2, existe $\tilde{A}_i \subseteq A_i$ tal que $m(\tilde{A}_i) = 0$, o conjunto dos $x \in A_i$ que retornam para A_i finitas vezes. Defina

$$X' := X \smallsetminus \bigcup_{i\in\mathbb{N}} \tilde{A}_i.$$

Temos que $m(X') = 1$, pois $m(\bigcup_{i\in\mathbb{N}} \tilde{A}_i) \leq +_{i\in\mathbb{N}} m(\tilde{A}_i) = 0$.

Queremos mostrar que $X' \subseteq \mathrm{R}(f)$. Seja $x \in X'$. Para toda vizinhança A de x, queremos encontrar $n \in \mathbb{N}$ tal que $f^n(x) \in A$. Como $(A_i)_{i\in\mathbb{N}}$ é base, existe $i \in \mathbb{N}$ tal que $x \in A_i \subseteq A$. Como $x \in X'$, então $x \notin \tilde{A}_i$, portanto existe $n \in \mathbb{N}$ tal que $f^n(x) \in A_i$, logo $f^n(x) \in A$. ■

28.4 Teoremas ergódicos

28.4.1 Tempo médio de visita

Começamos definindo o conceito de tempo médio de visita para motivar os enunciados dos teoremas ergódicos.

:⊢ **Definição 28.4.** Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida, $M \subseteq X$ um conjunto mensurável de medida não nula. Para todo $x \in X$, o *tempo médio de visita* de x em M é

$$\tau_M(x) := \lim_{n\to\infty} \frac{1}{n} \# \left(\left\{ t \in [n] \;\middle|\; f^t(x) \in M \right\} \right).$$

No caso de um sistema contínuo, definimos

$$\tau_M(x) := \lim_{n\to\infty} \frac{1}{n} \operatorname{vol}^1 \left(\left\{ t \in [0, n[\;\middle|\; f^t(x) \in M \right\} \right),$$

em que vol^1 é a medida de volume unidimensional em $\mathbb{R}$.

Esses limites não necessariamente existem para todos pontos. Podemos reescrever a equação de $\tau_M(x)$ como

$$\widehat{f}(\mathbf{1}_M)(x) = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} \mathbf{1}_M(f^t(x)) = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} ({f^t}_\star \mathbf{1}_M)(x).$$

Ainda, podemos escrever

$$\widehat{f}(\mathbf{1}_M) = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} \mathbf{1}_{f^{-t}(M)} = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} \mathbf{1}_{f^{-t}(M)}.$$

Podemos considerar casos mais gerais de funções $\phi \in \mathscr{M}(X, \mathbb{R})$ e perguntar se o limite

$$\widehat{f}(\phi) = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} \phi \circ f^t = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} {f^t}_\star \phi.$$

Os teoremas ergódicos vai tratar sobre esses limites.

28.4.2 Média temporal e média espacial

:⊢ **Definição 28.5.** Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida e $\phi \in L^1(X, \mathbb{R})$.

1. A *média temporal* de ϕ em $(\boldsymbol{X}, f)$ é a quase-função

$$\widehat{\phi} := \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} \phi \circ f^t = \lim_{n\to\infty} \frac{1}{n} \mathop{\bigplus}_{t\in[n]} {f^t}_\star(\phi);$$

2. Se m é finita e não nula, a *média espacial* de ϕ em $(\boldsymbol{X}, f)$ é

$$\overline{\phi} := \frac{1}{m(X)} \int \phi.$$

Podemos pensar na média temporal como um operador

$$\widehat{\cdot} := \lim_{n\to\infty} \frac{1}{n} \mathop{+}_{t\in[n]} {f^t}_\star.$$

Mais explicitamente,

$$\begin{aligned} \widehat{\cdot}\colon L^1(X,\mathbb{R}) &\longrightarrow L^1(X,\mathbb{R}) \\ \phi &\longmapsto \lim_{n\to\infty} \frac{1}{n} \mathop{+}_{t\in[n]} {f^t}_\star(\phi). \end{aligned}$$

Esse operador é linear e, como veremos, preserva a norma 1.

28.4.3 Teorema ergódico da média

⊢ **Teorema 28.5** (Ergódico da Média[1]). *Seja $(\boldsymbol{X}, f)$ um sistema dinâmico de medida e* $\mathrm{p}\colon \mathscr{I}^2(X,\mathbb{R}) \to (f_\star - \mathrm{I})^{-1}(0)$ *a projeção no subespaço*

$$(f_\star - \mathrm{I})^{-1}(0) = \left\{ \phi \in \mathscr{I}^2(X,\mathbb{R}) \;\middle|\; f_\star\phi = \phi \circ f = \phi \right\}.$$

Para toda função $\phi \in \mathscr{I}^2(X,\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{n} \mathop{+}_{i\in[n]} {f^i}_\star\phi \to \mathrm{p}\phi.$$

28.4.4 Teorema ergódico pontual

⊢ **Teorema 28.6** (Ergódico Pontual[2]). *Seja $(\boldsymbol{X}, f)$ um sistema dinâmico de medida (σ-finita). Para toda função $\phi \in L^1(X,\mathbb{R})$, existe $\widehat{\phi} \in \mathscr{I}^1(X,\mathbb{R})$ f-invariante tal que, para quase todo $x \in X$,*

$$\widehat{\phi}(x) = \lim_{n\to\infty} \frac{1}{n} \mathop{+}_{i\in[n]} \phi \circ f^i(x).$$

Se a medida é finita, então

$$\int \widehat{\phi} = \int \phi.$$

[1] Este teorema é conhecido como Teorema Ergódico de Von Neumann, em homenagem ao matemático húngaro *John von Neumann* (28/12/1903 – 08/02/1957).

[2] Este teorema é conhecido como Teorema Ergódico de Birkhoff, em homenagem ao matemático estadunidense *George David Birkhoff* (21/03/1884 – 12/11/1944).

$\vdash$ **Proposição 28.7.** *Seja $(\boldsymbol{X}, f)$ um sistema dinâmico de medida (probabilidade). São equivalentes*

1. *Para toda $\phi \in \mathscr{I}^1(X, \mathbb{R})$, $\widehat{\phi}$ é quase constante;*
2. *Para toda $\phi \in \mathscr{I}^1(X, \mathbb{R})$, $\widehat{\phi} = \int \phi$;*
3. *Para toda $\phi \in \mathscr{I}^1(X, \mathbb{R})$, se ϕ é quase f-invariante, então é quase constante.*
4. *Todo conjunto mensurável invariante ($f^{-1}(M) = M$) é quase vazio ou quase total ($m(M) = 0$ ou $m(M) = 1$);*
5. *Todo conjunto mensurável negativamente invariante ($f^{-1}(M) \subseteq M$) é quase vazio ou quase total;*
6. *Todo conjunto mensurável $M \subseteq X$ tal que $M \subseteq f^{-1}(M)$ (positivamente invariante?) é quase vazio ou quase total.*

Essas proposições são a definição de um sistema dinâmico ergódico, como veremos na seção seguinte.

28.5 Ergodicidade

Nesta seção, introduziremos o conceito de ergodicidade. A definição de transformação em sistemas ergódicos virá seguida de uma proposição que mostra várias equivalências de propriedades que caracterizam a ergodicidade. Essas equivalências são importantes não só como resultados teóricos interessantes e importantes ferramentas da teoria, mas também como embasamento intuitivo da ideia de ergodicidade. Relembramos que um conjunto é quase vazio quando tem medida 0 e quase total quando seu complementar tem medida 0.

$:\vdash$ **Definição 28.6.** Um *sistema dinâmico ergódico* é um sistema dinâmico de medida $\mathcal{S} = (\boldsymbol{X}, f)$ tal que todo conjunto mensurável invariante de $\mathcal{S}$ é quase vazio ou quase total: para todo $M \in \mathcal{M}$ tal que $f^{-1}(M) = M$, $M \mathrel{\overset{\circ}{=}} \emptyset$ ou $M \mathrel{\overset{\circ}{=}} X$. A dinâmica f é uma *dinâmica ergódica.*

$\vdash$ **Proposição 28.8.** *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico de medida. São equivalentes*

1. *f é ergódica;*
2. *Todo conjunto quase invariante de $\mathcal{S}$ é quase vazio ou quase total (para todo $M \in \mathcal{M}$ tal que $f^{-1}(M) \mathrel{\overset{\circ}{=}} M$, $M \mathrel{\overset{\circ}{=}} \emptyset$ ou $M \mathrel{\overset{\circ}{=}} X$);*
3. *A semiórbita positiva de quase todo ponto passa por qualquer conjunto de medida positiva: para todo $M \in \mathcal{M}$ tal que $M \mathrel{\overset{\circ}{\neq}} \emptyset$,*

$$\bigcup_{n \in \mathbb{Z}_{\geq 0}} f^{-n}(M) = \left\{ x \in X \;\middle|\; \mathcal{O}^+(x) \cap M \neq \emptyset \right\} \mathrel{\overset{\circ}{=}} X;$$

4. *Se a semiórbita positiva de quase nenhum ponto de um conjunto mensurável passa por um outro conjunto mensurável, algum dos conjuntos é quase vazio: para todos $N, M \in \mathcal{M}$ tais que*

$$N \cap \bigcup_{n \in \mathbb{Z}_{\geq 0}} f^{-n}(M) \mathrel{\mathring{=}} \emptyset,$$

$N \mathrel{\mathring{=}} \emptyset$ ou $M \mathrel{\mathring{=}} \emptyset$.

□ *Demonstração.* (1) $\Rightarrow$ (2) Pela proposição 28.1, existe um conjunto invariante $N \in \mathcal{M}$ tal que $M \mathrel{\mathring{=}} N$. Pela ergodicidade, segue que $N \mathrel{\mathring{=}} \emptyset$ ou $N \mathrel{\mathring{=}} X$, portanto segue que $M \mathrel{\mathring{=}} \emptyset$ ou $M \mathrel{\mathring{=}} X$.

(2) $\Rightarrow$ (3) Seja $M \in \mathcal{M}$ com $m(M) > 0$ e definamos

$$N := \bigcup_{n \in \mathbb{Z}_{\geq 0}} f^{-n}(M).$$

Segue que $f^{-1}(N) \subseteq N$. Isso implica que $f^{-1}(N) \cup N = N$ e que $f^{-1}(N) \cap N = f^{-1}(N)$. Portanto, como $m(f^{-1}(N)) = m(N)$, pois f preserva medida, segue que

$$\begin{aligned} m(N \triangle f^{-1}(N)) &= m(f^{-1}(N) \cup N) - m(f^{-1}(N) \cap N) \\ &= m(N) - m(f^{-1}(N)) = 0. \end{aligned}$$

Assim, temos que $f^{-1}(N) \mathrel{\mathring{=}} N$, e de (2) segue que $N \mathrel{\mathring{=}} \emptyset$ ou $N \mathrel{\mathring{=}} X$. Mas $N \not\mathrel{\mathring{=}} \emptyset$, pois $f^{-1}(M) \subseteq N$ e $m(f^{-1}(M)) = m(M) > 0$, logo temos que $N \mathrel{\mathring{=}} X$.

(3) $\Rightarrow$ (4) Vamos provar a recíproca. Sejam $M, N \in \mathcal{M}$ tais que $M \not\mathrel{\mathring{=}} \emptyset$ e $N \not\mathrel{\mathring{=}} \emptyset$. Por (3), sabemos que

$$m\left(\bigcup_{n \in \mathbb{Z}_{\geq 0}} f^{-n}(M)\right) = 1,$$

e disso segue que

$$0 < m(N) = m\left(\bigcup_{n \in \mathbb{Z}_{\geq 0}} N \cap f^{-n}(M)\right) \leq \mathop{+}_{n \in \mathbb{Z}_{\geq 0}} m(N \cap f^{-n}(M)).$$

Logo, existe $n \in \mathbb{Z}_{\geq 0}$ tal que $m(N \cap f^{-n}(M)) > 0$.

(4) $\Rightarrow$ (1) Seja $M \in \mathcal{M}$ tal que $f^{-1}(M) = M$. Então segue que $f^{-n}(M) = M$ para todo $n \in \mathbb{Z}_{\geq 0}$, logo

$$0 = m(M \cap M^{\complement}) = m(f^{-n}(M) \cap M^{\complement})$$

para todo $n \in \mathbb{Z}_{\geq 0}$, o que implica por (4) que $M \mathrel{\mathring{=}} \emptyset$ ou $M^{\complement} \mathrel{\mathring{=}} \emptyset$, logo $M \mathrel{\mathring{=}} \emptyset$ ou $M \mathrel{\mathring{=}} X$, e a ergodicidade está provada. ■

⊢ **Proposição 28.9.** *1. A rotação R_α no círculo é ergódica se, e somente se, α é irracional.*

2. O shift de Bernoulli é ergódico.

3. A expansão no círculo é ergódica.

□ *Demonstração.* 1. Se $\alpha \in \mathbb{Q}$, escrevemos $\alpha = \frac{p}{q}$ fração reduzida. Então, tomando um conjunto mensurável $A \subseteq \mathbb{T}$ tal que $) < m(A) < \frac{1}{q}$, segue que o conjunto

$$M := A \cup R_\alpha(A) \cup R_\alpha^2(A) \cap \cdots \cup R_\alpha^{q-1}(A)$$

é invariante sob R_α, pois $R_\alpha^q = id_{\mathbb{T}}$. No entanto, $0 < m(M) < 1$, o que mostra que $\mathbb{R}_\alpha$ não é ergódica.

Se $\alpha \notin \mathbb{Q}$, então a rotação é densa em $\mathbb{T}$. Sendo assim, seja $M \subseteq \mathbb{T}$ invariante sob R_α. Podemos escolher, para todo ε uma função f contínua no toro tal que $\|f - \mathbf{1}_M\|_1 < \varepsilon$ e da invariância de M segue que $\|f \circ R_\alpha^n - f\|_1 < 2\varepsilon$. Da continuidade de f, segue que, para todo $t \in \mathbb{R}$, $\|f \circ R_t - f\|_1 \leq 2\varepsilon$. Agora, como a medida é invariante por rotação, usando o teorema de Fubini e a desigualdade triangular para integrais temos que

$$\begin{aligned}\left\|f - \int f(t)dt\right\|_1 &= \int \left|\int (f(x) - f(x+t))dt\right| dx \\ &\leq \int\int |f(x) - f(x+t)|dxdt \\ &\leq 2\varepsilon.\end{aligned}$$

Mas então

$$\|\mathbf{1}_M - m(M)\|_1 \leq \|\mathbf{1}_m - f\|_1 + \left\|f - \int f(t)dt\right\|_1 + \left\|\int f(t)dt - m(M)\right\|_1 < 4\varepsilon.$$

Concluímos, portanto, que $\mathbf{1}_M$ é constante, pois isso vale para qualquer ε. Mas então $m(M)$ é 0 ou 1, logo R_α é ergódica.

2. Seja $N \in \mathcal{M}$ invariante. Para qualquer $\varepsilon > 0$, existe uma união finita de cilindros M tais que $m(M \triangle N) < \varepsilon$, o que implica $|m(M) - m(N)| < \varepsilon$. O conjunto M pode ser escrito como

$$M = \{x \in X : x_{-m} \cdots x_m \in F\}$$

para algum m e $F \subseteq \{0, 1, \cdots, n\}^{\{-m,\cdots,m\}}$. Assim, para todo $k > 2m$, temos que

$$\sigma^{-k}(M) = \{x \in X : x_{k-m} \cdots x_{k+m} \in F\}.$$

Como esse dois conjuntos são independentes, segue que

$$\begin{aligned}m\left(\sigma^{-k}(M) \setminus M\right) &= m\left(\sigma^{-k}(M) \cap M^{\complement}\right) \\ &= m\left(\sigma^{-k}(M)\right) m\left(M^{\complement}\right) \\ &= m(M)m\left(M^{\complement}\right).\end{aligned}$$

Como N é invariante, temos que $m(N \triangle \sigma^{-1}(M))$, e então

$$m(\sigma^{-k}(M) \triangle N) = m(\sigma^{-k}(M) \triangle \sigma^{-k}(N)) = m(M \triangle N) < \varepsilon$$

logo $m(\sigma^{-k}(M) \triangle M) < 2\varepsilon$ e

$$m(\sigma^{-k}(M) \smallsetminus M) + m(M \smallsetminus \sigma^{-k}(M)) = m(\sigma^{-k}(M) \triangle M) < 2\varepsilon.$$

Finalmente, obtemos que

$$\begin{aligned} m(N)m\left(N^{\complement}\right) &< (m(M)+\varepsilon)(m(M^{\complement})+\varepsilon) \\ &= m(M)m(M^{\complement}) + \varepsilon\mu(M) + \varepsilon\mu(M^{\complement}) + \varepsilon^2 \\ &< 2\varepsilon + \varepsilon + \varepsilon^2. \end{aligned}$$

Como isso vale para qualquer ε, segue que $m(N)m\left(N^{\complement}\right) = 0$; ou seja, que $m(N)$ é 0 ou 1, logo o sistema é ergódico.

3. Como os sistemas de Bernoulli Σ e a expansão E_2 são isomorfos por medida, segue que a expansão é ergódica.

■

28.6 Sistemas misturadores

Nesta seção consideraremos duas noções de um sistema misturador: a de *misturador* e de *fracamente misturador*. Um sistema misturador também é conhecido como fortemente misturador para distinguí-lo dos fracamente misturadores. Existem, de fato, outras noções de misturador, que generalizam a ideia de misturador, como *misturador de ordem k* (cf. [**ein**]), mas aqui consideraremos apenas essas duas. Antes de introduzir a definição, vale ressaltar uma outra definição equivalente de sistema ergódico, enunciada abaixo, cuja demonstração pode ser achada em [**ein**]. Essa classificação é uma consequência do teorema da média ergódica.

⊢ **Proposição 28.10.** *Um sistema $(X, \mathcal{M}, m, T)$ é ergódico se, e somente se, para todos $M, N \in \mathcal{M}$,*

$$\frac{1}{n} \mathop{+}_{i=0}^{n-1} m(M \cap T^{-i}(N)) \to m(M)m(N)$$

quando $n \to \infty$.

A definição a seguir, de um sistema misturador, destaca a seguinte ideia. Em alguns sistema, se tomamos dois subconjuntos mensuráveis quaisquer dos espaços, notamos que sempre ocorre que um deles é independente da n-ésima imagem inversa do outro quando n é grande. O sentido de independência empregado é que, em

espaços de probabilidade, eventos são pensados como conjuntos mensuráveis e dois eventos E, F são independentes quando $m(E \cap F) = m(E)m(F)$; ou seja, a probabilidade dos dois ocorrerem é o produto das probabilidades de cada um deles ocorrer separadamente. Pensando em conjuntos, e agora considerando uma ideia geométrica que motiva o nome *misturador*, essa ideia é equivalente a pensar que os conjuntos, quando iterados, estão muito bem separados no sistema, misturados, de modo que se fixamos um deles, o primeiro, a medida das imagens inversas do segundo que estão nesses primeiro conjunto dividida pela medida desse primeiro é igual à medida do segundo; ou seja, a quantidade do segundo que está no primeiro é proporcionalmente igual à quantidade desse segundo no espaço todo. A definição a seguir formaliza o que foi aqui explicado.

:⊢ **Definição 28.7.** Um sistema *misturador* é um sistema $(X, \mathcal{M}, m, T)$ que preserva medida tal que, para todos $M, N \in \mathcal{M}$,

$$m(M \cap T^{-n}N) \to m(M)m(N)$$

quando $n \to \infty$

⊢ **Proposição 28.11.** *A rotação R_α no círculo não é misturadora.*

□ *Demonstração.* Tomando uma sequência $n_i \to \infty$ tal que $n_i\alpha \mod 1 \to 0$. No caso de α irracional isso é possível pois a rotação é densa, e no caso de racional, basta tomar uma sequência constante igual a zero. Tomando os conjuntos $M = N = [0, \frac{1}{2}]$, segue que, quando $i \to \infty$,

$$m(M \cap R_\alpha^{-n_i}(N)) = m(M \cap R_{-n_i\alpha}(M)) \to m(M) = \frac{1}{2} \neq \left(\frac{1}{2}\right)^2 = m(M)m(N).$$

■

Embora o conceito de sistema misturador como definido acima é mais natural, a noção levemente modificada abaixo é, muitas vezes, mais útil e simples de se aplicar. Esse é o conceito de fracamente misturador.

:⊢ **Definição 28.8.** Um sistema *fracamente misturador* é um sistema $(X, \mathcal{M}, m, T)$ tal que, para todos $M, N \in \mathcal{M}$,

$$\frac{1}{n} \mathop{+}_{i=0}^{n-1} \left| m(M \cap T^{-i}(N)) - m(M)m(N) \right| \to 0$$

quando $n \to \infty$

É interessante comentar algo sobre os diferentes tipos de convergência aqui considerados. Além da convergência tradicional, $a_n \to a$, existem outras noções de convergência de uma sequência. Uma delas é convergência de Cesàro, que considera a convergência da sequência $\frac{1}{n} +_{i=0}^{n-1} a_n$ das n-ésimas médias da sequência original. Pelas definições e proposições anteriores, considerando a sequência $a_i := m(M \cap T^{-i}(N)) - m(M)m(N)$, temos que um sistema é *ergódico* se, e somente se, a_n converge a 0 pela convergência de Cesàro, é *misturador* se, e somente se, a_n converge a 0 no sentido tradicional, e é *fracamente misturador* se $|a_n|$ converge a 0 pela convergência de Cesàro. Assim, como, para qualquer sequência a_n que converge a 0, vale que $|a_n|$ converge a 0 por Cesàro e, então, a_n converge a 0 por Cesàro, segue a proposição abaixo.

$\vdash$ **Proposição 28.12.** *Seja $(X, \mathcal{M}, m, T)$ um sistema que preserva medida. Então*

1. *Se o sistema é misturador, é fracamente misturador.*
2. *Se o sistema é fracamente misturador, é ergódico.*

$\square$ *Demonstração.* Basta considerar a sequência

$$a_i := m(M \cap T^{-i}(N)) - m(M)m(N)$$

e os comentários do parágrafo anterior. ■

$:\vdash$ **Definição 28.9.** A *densidade* de um conjunto $C \subseteq \mathbb{N}$ é o número

$$d(C) := \lim_{n\to\infty} \frac{1}{n} |\{k \in C : 0 \leq k < n\}| = \frac{1}{n} |(C \cap [0, n)|$$

quando o limite existir.

Antes de demonstrar um teorema importante a seguir com várias equivalências envolvendo a propriedade de ergodicidade e mistura fraca, demonstraremos um lema sobre convergência de sequências.

$\vdash$ **Lema 28.13.** *Seja $(a_n)_{n\in\mathbb{N}}$ uma sequência limitada de números reais positivos. São equivalentes as seguintes noções de convergência.*

1. $\lim_{n\to\infty} \frac{1}{n} \mathop{+}_{i=0}^{n-1} a_i = 0$.
2. *Existe um conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $a_n \to 0$ para $n \to \infty$, $n \notin C$.*
3. $\lim_{n\to\infty} \frac{1}{n} \mathop{+}_{i=0}^{n-1} a_i^2 = 0$.

□ *Demonstração.* Vamos mostrar que (1) implica (2). Definamos $C_k := \{j \in \mathbb{N} : a_j > \frac{1}{k}\}$. Então temos que

$$C_1 \subseteq C_2 \subseteq C_3 \subseteq \cdots, \tag{28.1}$$

pois, se $x \in C_k$, então $x > \frac{1}{k} > \frac{1}{k+1}$, logo $x \in C_{k+1}$. Então temos que, para todo $k \geq 1$,

$$\frac{1}{k}|C_k \cap [0,n)| = \mathop{+}_{i \in C_k \cap [0,n)} \frac{1}{k} < \mathop{+}_{i \in C_k \cap [0,n)} a_i \leq \mathop{+}_{i=0}^{n} a_i$$

e então segue que

$$\frac{1}{n}|C_k \cap [0,n)| \leq k\frac{1}{n} \mathop{+}_{i=0}^{n} a_i \to 0$$

quando $n \to \infty$. Logo, para todo $k \geq 1$, C_k tem densidade 0. Como cada C_k tem densidade 0, podemos escolher inteiros $0 < l_1 < l_2 < \cdots$ tais que, para todo $k \geq 1$ e para todo $n \geq l_k$,

$$\frac{1}{n}|C_k \cap [0,n)| \leq \frac{1}{k}. \tag{28.2}$$

Basta escolher l_1 e tomar um l_2 suficientemente grande e assim por diante. Definimos então o conjunto

$$C := \bigcup_{k \in \mathbb{N}} (C_k \cap [l_k, l_{k+1})). \tag{28.3}$$

Vamos mostrar que a_i converge a 0 para $n \notin C$ e que C tem densidade 0. Primeiro, notamos que $C_k \cap [l_k, \infty) \subseteq C$, pois, se $x \in C_k$, então está em todo C_p para $p \geq k$, e, então, se estiver em $[l_p, l_{p+1})$, estará em $C_p \cap [l_p, l_{p+1})$, logo em C. Assim, para todo $n \geq l_k$, então se $n \notin C$, temos $n \notin C_k$ e, portanto, $a_n \leq \frac{1}{k}$, logo $a_n \to 0$. Agora, para mostrar que C tem densidade zero, notemos que, para todo $n \in [l+k, l_{k+1})$ temos que $C \cap [0,n) \subseteq C_k \cap [0,n)$, pois se $x \in C \cap [0,n)$, está em algum $[l_p, l_{p+1})$ e, por consequência, em $C_k \cap [0,n)$, já que todo C_p está contido no C_k para $p \leq k$. Isso implica que

$$\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{n}|C_k \cap [0,n)| \leq \frac{1}{k} \tag{28.4}$$

e, portanto, C tem densidade zero.

Agora, vamos mostrar que (2) implica (1). Como a sequência é limitada, seja $L > 0$ tal que $a_n \leq L$ para todo $n \geq 1$. Como a_n converge para zero fora de C e o conjunto C tem densidade zero, podemos escolher, para cada $k \geq 1$, $n_k \in \mathbb{N}$ tal que, para todo $n \notin C$, $n \geq n_k$, temos que $a_n < \frac{1}{k}$ e, para todo $n \geq n_k$, $\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{k}$.

Sendo assim, para todo $n \geq kn_k$, temos que

$$\begin{aligned}\frac{1}{n}\mathop{+}_{i=0}^{n-1} a_i &= \frac{1}{n}\left(\mathop{+}_{i=0}^{n_k-1} a_i + \mathop{+}_{i\in C\cap[n_k,n)} a_i + \mathop{+}_{i\in C^{\complement}\cap[n_k,n)} a_i\right)\\ &< \frac{1}{n}\left(Ln_k + L|C\cap[0,n)| + n\frac{1}{k}\right)\\ &\leq \frac{2L+1}{k},\end{aligned}$$

o que mostra que a média parcial converge para 0.

Para mostrar a equivalência de (3), basta notar que pela equivalência de (1) e (2), e como $a_n \to 0$ se, e somente se, $a_n^2 \to 0$, pois a sequência é positiva, temos a equivalência de (1) e (3). ■

⊢ **Teorema 28.14.** *Seja $(X, \mathcal{M}, m, T)$ um sistema que preserva medida. São equivalentes*

1. *T é fracamente misturadora.*
2. *$T \times T$ é ergódica com respeito a $m \times m$.*
3. *$T \times T$ é fracamente misturadora com respeito a $m \times m$.*
4. *Para todo sistema ergódico $(Y, \mathcal{Y}, \nu, S)$ que preserva medida, o sistema $(X \times Y, \mathcal{M} \otimes \mathcal{Y}, m \times \nu, T \times S)$ é ergódico.*
5. *Para todos $M, N \in \mathcal{M}$, existe conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $m(M \cap T^{-n}(N)) \to m(M)m(N)$ quando $n \to \infty$ e $n \notin C$.*
6. *Para todos $M, N \in \mathcal{M}$,*

$$\frac{1}{n}\mathop{+}_{i=0}^{n-1}\left|m(M\cap T^{-i}(N)) - m(M)m(N)\right|^2 \to 0$$

quando $n \to \infty$.

□ *Demonstração.* A equivalência de (1),(5) e (6) é consequência do lema anterior tomando como a_i a sequência $|m(M \cap T^{-i}(N)) - m(M)m(N)|$. A demonstração das demais equivalências seguirá o seguinte esquema: (5) ⇒ (3) ⇒ (1) ⇒ (4) ⇒ (2) ⇒ (6) ⇒ (5).

(5) ⇒ (3). Sejam $M_1, M_2, N_1, N_2 \in \mathcal{M}$, existem conjuntos $C_1, C_2 \subseteq \mathbb{N}$ com densidade 0 tais que

$$\begin{aligned}&m(M_1 \cap T^{-n}(N_1)) \to m(M_1)m(N_1) \quad \text{para } n \notin C_1\\ &m(M_2 \cap T^{-n}(N_2)) \to m(M_2)m(N_2) \quad \text{para } n \notin C_2.\end{aligned}$$

Seja $C := C_1 \cup C_2$. Então C tem densidade zero, pois $C \cap [0,n) = (C_2 \cap [0,n)) \cap (C_2 \cap [0,n))$, o que implica

$$\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{n}(|C_1 \cap [0,n)| + |C_2 \cap [0,n)|) \to 0.$$

Assim, temos que, para todo $n \notin C$,

$$\begin{aligned}
&\Big((m \times m)(M_1 \times M_2) \cap (T \times T)^{-n}(N_1 \times N_2) \\
&\qquad - (m \times m)(M_1 \times M_2)(m \times m)(N_1 \times N_2)\Big) \\
&= \Big(m(M_1 \cap T^{-n}(N_1))m(M_2 \cap T^{-n}(N_2)) - m(M_1)m(M_2)m(N_1)m(N_2)\Big) \\
&\to 0.
\end{aligned}$$

Isso mostra que $T \times T$ é fracamente misturadora, pois os retângulos geram $\mathcal{M} \otimes \mathcal{M}$.

(3) $\Rightarrow$ (1). Sejam $M, N \in \mathcal{M}$. Então, para $M \times X, N \times X \in \mathcal{M} \otimes \mathcal{M}$ e um conjunto com densidade zero $C \subseteq \mathbb{N}$, e para todo $n \notin C$,

$$\begin{aligned}
0 &= \lim_{n \notin C} \Big((m \times m)(M \times X) \cap (T \times T)^{-n}(N \times X) \\
&\qquad - (m \times m)(M \times X)(m \times m)(N \times X)\Big) \\
&= \lim_{n \notin C} \Big(m(M \cap T^{-n}(N))m(X \cap T^{-n}(X)) - m(M)m(X)m(N)m(X)\Big) \\
&= \lim_{n \notin C} \Big(m(M \cap T^{-n}(N)) - m(M)m(N)\Big),
\end{aligned}$$

pois $m(X) = 1$ e T é invariante, o que implica que T é fracamente misturadora.

(1) $\Rightarrow$ (4). Consideremos um sistema ergódico $(Y, \mathcal{Y}, \nu, S)$ que preserva medida. Sejam $M_1, N_1 \in \mathcal{M}$ e $M_2, N_2 \in \mathcal{Y}$. Então

$$\begin{aligned}
&\lim_{n \to \infty} \frac{1}{n} \mathop{+}_{i=0}^{n-1} \Big(m \times \nu)(M_1 \times M_2 \cap (T \times S)^{-i}(N_1 \times N_2)\Big) \\
&= \lim_{n \to \infty} \frac{1}{n} \mathop{+}_{i=0}^{n-1} m(M_1 \cap T^{-i}(N_1))\nu(M_2 \cap S^{-i}(N_2)) \\
&= \lim_{n \to \infty} \frac{1}{n} \mathop{+}_{i=0}^{n-1} m(M_1)m(M_2)\nu(M_2 \cap S^{-i}(N_2)) \\
&\qquad + \frac{1}{n} \mathop{+}_{i=0}^{n-1} \Big(m(M_1 \cap T^{-i}(N_1)) - m(M_1)m(M_2)\Big)\nu(M_2 \cap S^{-i}(N_2)).
\end{aligned}$$

Mas a primeira parcela dessa soma final é $m(M_1)m(N_1)\nu(M_1)\nu(N_2)$, pois, pela ergodicidade de ν, temos que $+_{i=0}^{n-1} \nu(M_2 \cap S^{-i}(N_2)) \to \nu(M_2)\nu(N_2)$, e a segunda

parcela da soma é zero, pois $\nu(M_2 \cap S^{-i})(N_2) \leq 1$ é limitado e T é fracamente misturadora. Como $m(M_1)m(N_1)\nu(M_1)\nu(N_2) = (m\times\nu)(M_1\times M_2)(m\times\nu)(N_1\times N_2)$, temos o resultado procurado.

(4) $\Rightarrow$ (2). Consideremos o sistema $(Y, \mathcal{Y}, \nu, S)$, com $Y = \{y\}$ e $S = id_Y$, que é ergódico e preserva medida. Nesse caso, $T \times S$ é isomórfico a T e, então, por (4), T é ergódico. Usando (4) novamente, e tomando $Y = X$, temos que $T \times T$ é ergódico.

(2) $\Rightarrow$ (6). Sejam $M, N \in \mathcal{M}$. Pela ergodicidade de $T \times T$, temos que

$$\begin{aligned}
&\frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} m(M \cap T^{-i}(N)) \\
&= \frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} \Big((m \times m)(M \times X \cap (T \times T)^{-i}(N \times X))\Big) \\
&\to (m \times m)(M \times X)(m \times m)(N \times X) = m(M)m(N)
\end{aligned}$$

e, analogamente,

$$\frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} m(M \times T^{-i}(N))^2 \to m(M)^2\mu(N)^2.$$

Disso, segue que

$$\begin{aligned}
&\frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} \Big(m(M \cap T^{-i}(N)) - m(M)m(N)\Big)^2 \\
&= \frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} m(M \cap T^{-i}(N))^2 \\
&\qquad - 2m(M)m(N)\frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} m(M \cap T^{-i}(N)) + m(M)^m(N)^2\frac{1}{n}\mathop{\mathop{+}}_{i=0}^{n-1} 1 \\
&\to m(M)^2\mu(N)^2 - 2\mu(M)m(N)m(M)m(N) + m(M)^2\mu(N)^2 = 0.
\end{aligned}$$

Isso completa a demostração do teorema, pois já sabemos que (6) $\Rightarrow$ (5). ■

28.7 Entropia de medida

Para as próximas definições, denotamos a função *logaritmo* por log.

:⊢ **Definição 28.10.** Sejam $\boldsymbol{X}$ um espaço de medida e $\mathcal{P} = (P_i)_{i=0}^n$ uma partição finita de X. A *entropia* de $\mathcal{P}$ é o número

$$E(\mathcal{P}) := E(m(P_0), \ldots, m(P_n)) = -\mathop{\mathop{+}}_{i=0}^{n} m(P_i)\log(m(P_i)).$$

:⊢ **Definição 28.11.** Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida irrevertível e $\mathcal{P}$ uma partição finita de X. A *entropia* de f com respeito a $\mathcal{P}$ é o número

$$E(f, \mathcal{P}) := \lim_{n\to\infty} \frac{1}{n} E\left(\bigvee_{t=0}^{n} f^{-t}(\mathcal{P})\right).$$

Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida revertível e $\mathcal{P}$ uma partição finita de X. A *entropia* de f com respeito a $\mathcal{P}$ é o número

$$E(f, \mathcal{P}) := \lim_{n\to\infty} \frac{1}{n} E\left(\bigvee_{t=-n}^{n} f^{-t}(\mathcal{P})\right).$$

Em ambos os casos, a *entropia* de f é o número

$$E(f) := \sup\left\{E(f, \mathcal{P}) \mid \mathcal{P} \text{ é quase partição de } \boldsymbol{X}\right\}.$$

28.8 Medida invariante para contrações e expansões

⊢ **Proposição 28.15.** *Seja $(\boldsymbol{M}, f)$ sistema dinâmico métrico em que $\boldsymbol{M}$ é um espaço métrico completo e $f\colon M \to M$ uma contração cujo ponto fixo é p. A única probabilidade sobre (a σ-álgebra topológica de) M preservada por f é a medida atômica em p.*

□ *Demonstração.* Sejam m uma probabilidade sobre M preservada por f e $c \in [0, 1[$ uma constante de contração de f. Primeiro, notemos que, para todo $r \in]0, \infty[$ e todo $n \in \mathbb{N}$,

$$\mathrm{B}\,(p; r) \subseteq f^{-n}(\mathrm{B}\,(p; c^n r)). \tag{28.5}$$

Para mostrar isso, seja $x \in \mathrm{B}\,(p; r)$. Então $|p, x| \leq r$. Como p é ponto fixo, $f^n(p) = p$, portanto

$$|p, f^n(x)| = |f^n(p), f^n(x)| \leq c^n\,|p, x| \leq c^n r,$$

o que mostra que $f^n(x) \in \mathrm{B}\,(p; c^n r)$. Como f preserva m, segue de (28.5) que

$$m(\mathrm{B}\,(p; r)) \leq m(f^{-n}(\mathrm{B}\,(p; c^n r))) = m(\mathrm{B}\,(p; c^n r)). \tag{28.6}$$

Agora, notemos que

$$\lim_{n\to\infty} \Theta(\mathrm{B}\,(p; c^n r)) = \lim_{n\to\infty} 2c^n r = 0.$$

Como, para todo $n \in \mathbb{N}$, $p \in \mathrm{B}\,(p; c^n r)$, as bolas $\mathrm{B}\,(p; c^n r)$ são fechadas e encaixadas e M é completo, segue do teorema dos conjuntos fechados encaixados que

$$\bigcap_{n\in\mathbb{N}} \mathrm{B}\,(p; c^n r) = \{p\}.$$

Como $m(\mathrm{B}\,(p; r)) < \infty$, segue que

$$m(\{p\}) = m\left(\bigcap_{n\in\mathbb{N}} \mathrm{B}\,(p; c^n r)\right) = \lim_{n\to\infty} m(\mathrm{B}\,(p; c^n r)).$$

De (28.6) segue que

$$m(\{p\}) = \lim_{n\to\infty} m(\mathrm{B}\,(p; c^n r)) \geq m(\mathrm{B}\,(p; r)).$$

Como isso vale para todo $r \in \,]0, \infty[$, segue que, para todo $n \in \mathbb{N}$,

$$m(\{p\}) \geq m(\mathrm{B}\,(p; n)),$$

o que implica que

$$m(\{p\}) \geq \lim_{n\to\infty} m(\mathrm{B}\,(p; n)) = m\left(\bigcup_{n\in\mathbb{N}} \mathrm{B}\,(p; n)\right) = m(M) = 1.$$

Já que m é probabilidade, $m(\{p\}) \leq 1$, o que mostra que $m(\{p\}) = 1$. ■

⊢ **Proposição 28.16.** *Seja $(\boldsymbol{M}, f)$ sistema dinâmico métrico em que $\boldsymbol{M}$ é um espaço métrico completo e $f\colon M \to M$ uma expansão contínua com ponto fixo p. A única probabilidade sobre (a σ-álgebra topológica de) M preservada por f é a medida atômica em p.*

□ *Demonstração.* Sejam m uma probabilidade sobre M preservada por f e $c \in \,]1, \infty[$ uma constante de expansão de f. Primeiro, notemos que, para todo $r \in \,]0, \infty[$ e todo $n \in \mathbb{N}$,

$$\mathrm{B}\left(p; c^{-n} r\right) \supseteq f^{-n}(\mathrm{B}\,(p; r)). \tag{28.7}$$

Para mostrar isso, seja $x \in f^{-n}(\mathrm{B}\,(p; r))$. Então $f^n(x) \in \mathrm{B}\,(p; r)$, o que implica $|p, f^n(x)| \leq r$. Como p é ponto fixo, $f^n(p) = p$, portanto

$$|p, x| \leq c^{-n}\,|f^n(p), f^n(x)| = c^{-n}\,|p, f^n(x)| \leq c^{-n} r,$$

o que mostra que $x \in \mathrm{B}\,(p; c^{-n} r)$. Como f preserva m, segue de (28.7) que

$$m(\mathrm{B}\left(p; c^{-n} r\right)) \geq m(f^{-n}(\mathrm{B}\,(p; r))) = m(\mathrm{B}\,(p; r)). \tag{28.8}$$

Agora, notemos que

$$\lim_{n\to\infty} \Theta(\mathrm{B}\left(p; c^{-n}r\right)) = \lim_{n\to\infty} 2c^{-n}r = 0.$$

Como, para todo $n \in \mathbb{N}$, $p \in \mathrm{B}\,(p; c^{-n}r)$, as bolas $\mathrm{B}\,(p; c^{-n}r)$ são fechadas e encaixadas e M é completo, segue do teorema dos conjuntos fechados encaixados que

$$\bigcap_{n\in\mathbb{N}} \mathrm{B}\left(p; c^{-n}r\right) = \{p\}.$$

Como $m(\mathrm{B}\,(p; r)) < \infty$, segue que

$$m(\{p\}) = m\left(\bigcap_{n\in\mathbb{N}} \mathrm{B}\left(p; c^{-n}r\right)\right) = \lim_{n\to\infty} m(\mathrm{B}\left(p; c^{-n}r\right)).$$

De (28.8) segue que

$$m(\{p\}) = \lim_{n\to\infty} m(\mathrm{B}\left(p; c^{-n}r\right)) \geq m(\mathrm{B}\,(p; r)).$$

Como isso vale para todo $r \in]0, \infty[$, segue que, para todo $n \in \mathbb{N}$,

$$m(\{p\}) \geq m(\mathrm{B}\,(p; n)),$$

o que implica que

$$m(\{p\}) \geq \lim_{n\to\infty} m(\mathrm{B}\,(p; n)) = m\left(\bigcup_{n\in\mathbb{N}} \mathrm{B}\,(p; n)\right) = m(M) = 1.$$

Já que m é probabilidade, $m(\{p\}) \leq 1$, o que mostra que $m(\{p\}) = 1$. ■

28.9 Exemplos

28.9.1 Medida dos deslocamentos

Definimos nos conjuntos $N^{\mathbb{Z}_{\geq 0}}$ e $N^{\mathbb{Z}}$ uma topologia, a topologia produto, de modo que os sistemas $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ e $(N^{\mathbb{Z}}, \sigma)$ se tornaram sistemas dinâmicos topológicos. Temos agora dois modos de definir uma σ-álgebra nesses conjuntos. A primeira é definirmos uma σ-álgebra de conjuntos mensuráveis em $N^{\mathbb{Z}_{\geq 0}}$ ou $N^{\mathbb{Z}}$ da mesma forma que definimos sua topologia, usando a σ-álgebra discreta em N, que é gerada pelos conjuntos da forma $\{k\}$ com $k \in N$, e tomando a σ-álgebra produto. A segunda maneira é consideramos a σ-álgebra gerada pela topologia de $N^{\mathbb{Z}_{\geq 0}}$ ou $N^{\mathbb{Z}}$, a σ-álgebra topológica ou σ-álgebra de Borel. Esse dois metodos, a princípio, não

são equivalentes, pois gerado pelos cilindros sub-básicos é ser uma união qualquer de interseções finitas de cilindros sub-básicos, enquanto um conjunto mesurável gerados pelos cilindros sub-básicos não são uniões quaisquer e sua forma é mais delicada de se descrever. Não discutiremos por ora mais detalhes sobre essa σ-álgebras, mas o que podemos dizer é que a σ-álegbra produto está contida na σ-álgebra topológica. Adotaremos, no momento, a primeira construção, pois estudaremos somente a estrutura mensurável desses sistemas a princípio. Como anteriormente, a seguir consideraremos o conjunto $N^{\mathbb{Z}}$, mas a construção é análoga para $N^{\mathbb{Z}_{\geq 0}}$.

Para definirmos uma probabilidade no conjunto $N^{\mathbb{Z}}$ com a σ-álgebra produto, consideramos antes um vetor de probabilidade $p = (p_0, \cdots, p_{N-1}) \in \Delta^{N-1}$, e definimos a medida m_p na σ-álgebra de N como $m(\{k\}) := p_k$ para todo $k \in N$. Essa medida é uma probabilidade em N. A partir dessa probabilidade, definimos uma probabilidade em $N^{\mathbb{Z}}$ como a medida produto, que é dada da seguinte forma. Para cada cilindro $C_{t_0,\ldots,t_n}[k_0, \ldots, k_n]$, definimos

$$m(C_{t_0,\ldots,t_n}[k_0, \ldots, k_n]) := m_p(\{k_0\}) \cdots m_p(\{k_n\}) = p_{k_0} \cdots p_{k_n}.$$

A partir dessa definição, a medida produto está definida para todo conjunto mensurável do jeito canônico.

Vamos mostrar que a função σ preserva medida. Primeiro, notemos que ela é mensurável. Seja $C_t[k]$ um cilindro gerador. Então

$$\sigma^{-1}(C_t[k]) = \sigma^{-1}({\pi_t}^{-1}(\{k\})) = (\pi_t \circ \sigma)^{-1}(\{k\}) = {\pi_{t-1}}^{-1}(\{k\}) = C_{t-1}[k],$$

que é mensurável pois $C_{t-1}[k]$ é um cilindro gerador. Isso mostra que σ é mensurável. Agora, notemos que m preserva medida. Seja $C_{t_0,\ldots,t_n}[k_0, \ldots, k_n]$ um cilindro. Então

$$m(\sigma^{-1}(C_{t_0,\ldots,t_n}[k_0, \ldots, k_n]))) = m(C_{t_0-1,\ldots,t_n-1}[k_0, \ldots, k_n]) = \bigtimes_{i=0}^{n} p_i,$$

portanto m preserva medida. Assim, provamos a seguinte proposição.

$\vdash$ **Proposição 28.17.** *Os sistemas* $(N^{\mathbb{Z}_{\geq 0}}, \sigma)$ *e* $(N^{\mathbb{Z}}, \sigma)$ *são sistemas dinâmicos de medida.*

28.9.2 Itinerários de medida

Lembremos que se $\mathcal{P} = (P_i)_{i \in N}$ é uma partição finita de X, então o itinerário de um ponto $x \in X$ com respeito a $\mathcal{P}$ é a sequência $(x_t)_{t \in T} \in N^T$ tal que $x_t = k$ se, e somente se, $f^t(x) \in P_k$. Denotamos a função que leva x em seu itinerário por $\phi : X \to N^T$. Queremos agora investigar as condições em $\mathcal{P}$ para que ϕ seja uma quase função que preserva medida. Para isso, teremos que garantir que ela seja

mensurável e garantir que ela preserve medida. Como trabalhamos com espaços de medida, a noção de partição agora pode ser enfraquecida para uma partição de medida, que não exige que as partes da partição sejam disjuntas, mas somente quase disjuntas.

Notemos que $x_t = \pi_t(\phi(x))$, portanto podemos escrever que $\pi_t(\phi(x)) = k$ se, e somente se, $f^t(x) \in P_k$. Isso é o mesmo que dizer que

$$(\pi_t \circ \phi)^{-1}(k) = f^{-t}(P_k).$$

Essa identidade é a chave para a demonstração da proposição a seguir.

⊢ **Proposição 28.18.** *Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico topológico discreto, $\mathcal{P} = (P_i)_{i \in N}$ uma partição finita por medida de $\boldsymbol{X}$ tal que $m(P_k) = p_k$, e $\phi : X \to N^T$ a função que leva $x \in X$ em seu itinerário com respeito a $\mathcal{P}$. Então ϕ é uma quase função que preserva medida (e é semiconjugação).*

□ *Demonstração.* A relação ϕ é uma quase função porque, como $\mathcal{P}$ não é uma partição, mas uma partição de medida, podem existir pontos em duas partes distintas de $\mathcal{P}$. No entanto, esses pontos são um conjunto quase vazio, o que mostra que a menos de um conjunto quase vazio a função está definida. Pela mesma demostração para o caso sem medida, essa quase função é uma semiconjugação.

Mostremos agora que ϕ preserva medida. Para isso, seja $C_t[k]$ um cilindro gerador de $N^{\mathbb{T}}$. Nesse caso, segue que

$$\phi^{-1}(C_t[k]) = \phi^{-1}({\pi_t}^{-1}(k)) = (\pi_t \circ \phi)^{-1}(k) = f^{-t}(P_k),$$

que é mensurável pois f é mensurável e P_k é mensurável. Agora, notemos que

$$m(\phi^{-1}(C_t[k])) = m(f^{-t}(P_k)) = m(P_k) = p_k = m(C_t[k])),$$

pois f preserva medida, e isso mostra que ϕ preserva medida. ■

28.9.3 Rotação no círculo

⊢ **Proposição 28.19.** *1. A rotação R_α em $\mathbb{T}^1$ preserva a medida topológica completa (de Lebesgue).*

2. A expansão E_m em $\mathbb{T}^1$ preserva a medida topológica completa (de Lebesgue).

□ *Demonstração.* 1. É suficiente mostrar que a medida é preservada para intervalos. Mas isso é simples, pois dado $[a, b] \in \mathbb{T}$, como $R_\alpha^{-1}([a, b]) = [a - \alpha, b - \alpha]$, segue que

$$m(R_\alpha^{-1}([a, b])) = (b - \alpha) - (a - \alpha) = b - a = m([a, b]).$$

2. Como no caso anterior, é suficiente mostrar para intervalos. Dado $[a,b] \in \mathbb{T}$, como $E_m^{-1}([a,b]) = \bigsqcup_{i=0}^{m-1} \left[\frac{a+i}{m}, \frac{b+i}{m}\right]$, segue que

$$m(E_m^{-1}([a,b])) = \mathop{+}_{i=0}^{m-1} \frac{b+i}{m} - \frac{a+i}{m} = m\frac{b-a}{m} = m([a,b]).$$

■

⊢ **Proposição 28.20.** *Seja o sistema de Bernoulli* $(\Sigma, \mathcal{B}, m, \sigma)$*, em que* $\Sigma := \{0,1\}^{\mathbb{N}}$*,* $\mathcal{B}$ *é a* σ*-álgebra de Borel de* Σ*,* σ *é o shift de Bernoulli e* m *é a medida produto* $m := \bigsqcap_{\mathbb{N}} \nu$*, em que* ν *é a medida em* $\{0,1\}$ *definida por* $\nu(\{0\}) = \nu(\{1\}) = \frac{1}{2}$*. Então o mapa*

$$\begin{aligned} \phi\colon \Sigma &\longrightarrow \mathbb{T} \\ (x_n)_{n\in\mathbb{N}} &\longmapsto \mathop{+}_{n\in\mathbb{N}} \frac{x_n}{2^{n+1}} \end{aligned}$$

é um isomorfismo entre o sistema de Bernoulli e a expansão E_2 *no círculo.*

□ *Demonstração.* Primeiro notemos que o sistema de Bernoulli deve preservar medida pelo comentário anterior. Assim, notemos que, para $(x_n)_{n\in N}$,

$$\phi \circ \sigma((x_n)) = \phi((x_{n+1})) = \mathop{+}_{n\in\mathbb{N}} \frac{x_{n+1}}{2^{n+1}} = 2 \mathop{+}_{n\in\mathbb{N}} \frac{x_{n+1}}{2^{n+2}} = E_2 \circ \phi(x_n).$$

A inversa de ϕ é definida a menos de o conjunto de racionais diádicos, que tem medida nula, e então segue o resultado. ■

Capítulo 29

Sistemas dinâmicos diferenciais

:⊢ **Definição 29.1.** Um *sistema dinâmico diferencial* é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é uma variedade diferencial, (X, f) é um sistema dinâmico e f é uma ação diferencial em $\boldsymbol{X}$.

29.1 Estrutura hiperbólica

:⊢ **Definição 29.2.** Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico diferencial discreto e revertível em que $\boldsymbol{X}$ é uma variedade diferencial compacta, conexa e métrica[1]. Um conjunto *(uniformemente) hiperbólico* de $\mathcal{S}$ é um conjunto invariante $H \subseteq X$ para o qual existem constantes reais $c \in]0, +\infty[$, $\lambda \in]0, 1[$ e uma decomposição de $\mathrm{T}X|_H$ em subfibrados vetoriais E^+ e E^-, isto é,

$$\mathrm{T}X|_H = E^+ + E^-,$$

invariantes por $\mathrm{D}f$, isto é, para todo $p \in X$,

$$\mathrm{D}f|_p(E^{+,-}|_p) = E^{+,-}|_{f(p)},$$

tais que, para todo $p \in H$, todo $t \in \mathbb{I}^+$ e todos $v^+ \in E^+|_p$, $v^- \in E^-|_p$,

$$\left\|\mathrm{D}f^t|_p v^+\right\| \leq c\lambda^t \left\|v^+\right\| \qquad \text{e} \qquad \left\|\mathrm{D}f^{-t}|_p v^-\right\| \leq c\lambda^t \left\|v^-\right\|.$$

[2]

Os fibrados vetoriais E^+ e E^- são os fibrados tangentes *estável* e *instável*[3] de $\boldsymbol{X}$, respectivamente, e, para todo $p \in H$, os espaços $E^+|_p$ e $E^-|_p$ são os espaços *estável* e *instável* tangentes a p, respectivamente.

[1] Tem métrica riemanniana.

[2] A notação $\mathrm{D}f^t$ não é ambígua, pois $\mathrm{D}(f^t) = (\mathrm{D}f)^t$.

[3] Os superíndices podem novamente gerar confusão. A mesma definição para os conjuntos estável e instável é usada.

► **Exemplo 29.1** (Automorfismo Hiperbólico (mapa do gato)). *Consideremos o difeomorfismo*

$$\begin{aligned}[A]\colon \mathbb{T}^2 &\longrightarrow \mathbb{T}^2\\ x &\longmapsto \begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix} x.\end{aligned}$$

⊢ **Proposição 29.1.** *Os conjuntos estável e instável de $W^+(x)$ e $W^-(x)$ para $x \in H$ um conjunto hiperbólico, são subvariedade (imersas, não mergulhadas).*

⊢ **Proposição 29.2** (Teorema das variedades estável e instável). *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico diferencial discreto e revertível em que $\boldsymbol{X}$ é uma variedade diferencial compacta, conexa e métrica e $H \subseteq X$ um conjunto hiperbólico. Para todo $x \in H$,*

$$\mathrm{T}W^+(x)|_x = E^+|_x \qquad \text{e} \qquad \mathrm{T}W^-(x)|_x = E^-|_x$$

29.2 Axioma A

:⊢ **Definição 29.3.** Um sistema dinâmico *Axioma A* é um sistema dinâmico diferencial discreto e revertível $\mathcal{S} = (\boldsymbol{X}, f)$ tal que

1. $\overline{\Pi(f)} = \Omega(f)$; isto é, o conjunto dos pontos periódicos de f é denso no conjunto dos pontos não-errantes;
2. $\Omega(f)$ é um conjunto hiperbólico.

⊢ **Teorema 29.3** (Decomposição Espectral de Smale). *Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico Axioma A. O conjunto não-errante Ω pode ser particionado em uma quantidade finita n de conjuntos compactos $(\Omega_i)_{i\in N}$ topologicamente transitivos:*

$$\Omega = \bigcup_{i\in[n]} \Omega_i.$$

Para cada $i \in [n]$, Ω_i pode ser particionado em m_i conjuntos compactos $(\Omega_{i,j})_{j\in[m_i]}$ tais que $f(\Omega_{i,j}) = \Omega_{i,j+1}$ (com $j+1$ interpretado $\mod m_i$) e $\Omega_{i,j}$ topologicamente misturador com respeito a f^{m_i}.

Os conjuntos Ω_i são chamados de *peças básicas*.

29.3 Anosov

:⊢ **Definição 29.4.** Um sistema dinâmico *de Anosov* é um sistema dinâmico diferencial discreto e revertível $\mathcal{S} = (\boldsymbol{X}, f)$ em que X é hiperbólico.

As variedades $\{W^+(x)\}_{x\in X}$ e $\{W^-(x)\}_{x\in X}$ são folheações, as folheações estável e instável de X.

Os fibrados estável e instável são integráveis porque existe folheação cujo espaço tangente são esses fibrados.

29.4 Partições de Markov

Bibliografia

[Car71] Henri Cartan. *Differential Calculus*. Méthodes. Hermann, 1971 (ver pp. 508, 524, 549).

[CC00] Alberto Candel e Lawrence Conlon. *Foliations I*. Graduate Studies in Mathematics. American Mathematical Society, 2000. URL: https://bookstore.ams.org/gsm-23/ (ver p. 684).

[Col12] Rodney Coleman. *Calculus on Normed Vector Spaces*. Universitext. Springer New York, 2012. ISBN: 9781461438946. URL: https://www.springer.com/gp/book/9781461438939 (ver pp. 507, 535, 547, 549).

[Lee12] John M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer-Verlag New York, 2012. URL: https://www.springer.com/gp/book/9781441999818 (ver p. 667).

[Lim14] Elon Lages Lima. *Curso de Análise*. Vol. 2. Projeto Euclides. IMPA, 2014. ISBN: 978-85-244-0372-9 (ver pp. 506, 552).