

Relatividade Especial, Geral e Geometria Lorentziana

Rodney Josué Biezuner¹

Departamento de Matemática

Instituto de Ciências Exatas (ICEEx)

Universidade Federal de Minas Gerais (UFMG)

2 de dezembro de 2016

¹E-mail: biezunerufmg@gmail.com; homepage: <http://www.mat.ufmg.br/~rodney>.

Sumário

1 Teoria da Relatividade Especial	5
1.1 Breve Introdução	5
1.2 Dedução da existência de ondas eletromagnéticas através das equações de Maxwell	6
1.2.1 O Éter	9
1.3 Referenciais Inerciais	9
1.4 Princípio da Relatividade	10
1.5 Invariância do Intervalo Espaçotemporal	11
1.6 Relatividade da Simultaneidade	14
1.7 Relatividade de Intervalos Espaciais e de Intervalos Temporais	14
1.7.1 Contração do Espaço	14
1.7.2 Dilatação do Tempo	15
1.7.3 Invariância das Dimensões Transversais	17
2 O Espaçotempo de Minkowski	18
2.1 Produto Interno e Formas Quadráticas	18
2.2 Operadores Ortogonais e Isometrias	24
2.3 Métrica de Lorentz e Espaçotempo de Minkowski	28
2.3.1 Propriedades do Espaçotempo de Minkowski	29
2.4 Transformação de Lorentz do tipo Boost	33
2.4.1 Medida Quantitativa da Dilatação do Tempo	34
2.4.2 Matriz do Boost	35
2.4.3 Medida Quantitativa da Contração do Espaço	37
2.4.4 Resumo da Contração do Espaço e Dilatação do Tempo	37
2.4.5 Lei de Adição de Velocidades	39
2.4.6 Rotações Hiperbólicas	42
3 Paradoxos	44
3.1 Diagramas de Espaçotempo	44
3.2 Impossibilidade de se viajar acima da velocidade da luz	46
3.3 Geodésicas no Espaço de Minkowski	49
3.4 Paradoxo dos Gêmeos	50
4 Momento-Energia	54
4.1 4-Vetores	54
4.2 Momento-Energia	55
5 Aceleração no Espaçotempo de Minkowski	58
6 O Tensor do Campo Eletromagnético	59

7 O Princípio de Equivalência – Geodésicas no Espaçotempo Curvo	60
7.1 O Pensamento Mais Feliz da Vida de Einstein	60
7.1.1 Inércia	61
8 Topologia Diferencial	62
8.1 Variedades Diferenciáveis	62
8.2 Aplicações Diferenciáveis	63
8.2.1 Partições da Unidade	64
8.3 Vetores Tangentes	65
8.3.1 Vetores Tangentes a Curvas	65
8.3.2 Vetores Tangentes como Derivações	68
8.3.3 Diferencial de uma Aplicação Diferenciável	69
8.4 Coordenadas	70
8.4.1 Diferencial em Coordenadas	70
8.5 Fibrado Tangente	72
8.6 Fibrados Vetoriais	72
8.7 Campos Vetoriais	73
8.8 Colchete de Lie	75
8.9 Derivada de Lie	78
8.10 Campos Vetoriais que Comutam	81
8.11 Exercícios	84
9 Tensores	86
9.1 Vetores Contravariantes e Covariantes	86
9.1.1 Significado Real do Gradiente	88
9.2 Convenção da Soma de Einstein	89
9.3 Vetores e Covetores	89
9.3.1 Mudança de Coordenadas em Espaços Vetoriais	89
9.3.2 Covetores	90
9.3.3 O Espaço Bidual	92
9.4 Vetores e Covetores Tangentes	93
9.4.1 Mudança de Coordenadas no Espaço Tangente $T_p M$	93
9.4.2 Covetores Tangentes	94
9.5 Tensores	95
9.5.1 Definição	95
9.5.2 Produto Tensorial	96
9.5.3 Mudança de Base	98
9.5.4 Traço de Tensores	99
9.6 Fibrados Tensoriais	101
9.7 Campos Tensoriais	101
9.8 Exercícios	102
10 Variedades Métricas	103
10.1 Definição e Exemplos	103
10.2 Comprimentos e Volumes	111
10.2.1 Comprimentos de Curvas Parametrizadas	111
10.2.2 Volumes em Variedades Riemannianas Orientáveis	112
10.3 Grupos de Lie e Álgebras de Lie	113
10.4 Exercícios	117

11 Conexões	118
11.1 Conexões e Derivada Covariante	118
11.2 Derivada Covariante ao longo de Curvas	120
11.3 Transporte Paralelo	122
11.4 Conexão Compatível com a Métrica	123
11.5 O Tensor Torsão	125
11.5.1 Conexão Simétrica	125
11.6 Conexões Métricas	126
11.6.1 Símbolos de Christoffel da Conexão Métrica	129
11.6.2 Interpretação Geométrica da Derivada Covariante	131
12 Geodésicas	133
12.1 Definição – A Equação Geodésica	133
12.2 Exemplos de Geodésicas	135
12.3 Fluxo Geodésico	139
12.4 A Aplicação Exponencial	141
12.5 Lema de Gauss	143
12.6 Geodésicas em Variedades Riemannianas	147
12.6.1 Geodésicas em Variedades Riemannianas minimizam distâncias localmente	147
12.6.2 Vizinhanças Convexas	150
12.7 Geodésicas em Variedades de Lorentz	152
12.8 Função Distância em Variedades Riemannianas	155
12.9 Variedades Riemannianas Completas e Teorema de Hopf-Rinow para Variedades Riemannianas	156
12.10 Apêndice: Geodésicas através do Método Variacional	159
12.11 Exercícios	162
13 Curvatura	163
13.1 Mais sobre Tensores	163
13.2 O Tensor Curvatura	164
13.2.1 O Endomorfismo Curvatura	164
13.2.2 O Significado da Curvatura Nula	165
13.2.3 Operação de Subir ou Descer um Índice	169
13.2.4 O Tensor Curvatura	172
13.3 Curvatura de Ricci e Curvatura Escalar	180
13.4 Curvatura Seccional	183
13.5 Variedades Riemannianas de Curvatura Seccional Constante	188
13.6 Apêndice: Motivação para a definição do tensor curvatura	190
13.6.1 Efeitos de Maré em Mecânica Newtoniana	190
13.6.2 Efeitos de Maré em Relatividade Geral: Curvatura como uma medida da aceleração relativa de trajetórias geodésicas	191
13.6.3 Curvatura como uma medida do transporte paralelo em trajetórias fechadas	194
13.7 Exercícios	197
14 Derivada Covariante de Campos Tensoriais	198
14.1 Conexão nos Fibrados Tensoriais	198
14.2 Derivada Covariante Total	206
14.2.1 Divergência de Campos Tensoriais	213

15 O Tensor Momento-Energia	215
15.1 4-Vetores	215
15.2 Momento-Energia	216
15.3 Fluxos	218
15.3.1 Definição	218
15.4 Tensor Momento-Energia	219
15.4.1 Tensor Momento-Energia da Poeira	219
15.4.2 Simetria do Tensor Momento-Energia	221
15.4.3 Conservação do Momento-Energia	222
15.4.4 Tensor Momento-Energia de um Fluido Perfeito	222
15.5 Apêndice: Divergente e Fluxo	222
15.5.1 Significado Geométrico do Divergente Espacial	224
15.5.2 Significado Físico do Divergente	227
16 Equação de Einstein	229
16.1 Equação de Newton para o Potencial Gravitacional	229
16.2 Dedução da Equação de Einstein	231
16.2.1 O Tensor de Einstein tem Divergência Nula	233
16.2.2 O Limite Newtoniano	233
16.3 Postulado Geodésico como Consequência da Equação de Einstein	236
16.4 A Constante Cosmológica	238
17 A Solução de Schwarzschild	239
17.1 Métrica Plana em Coordenadas Esféricas	239
17.2 Métrica de Schwarzschild	240
17.3 Conservação de Energia e Momento Angular em Geodésicas do Espaçotempo de Schwarzschild	250
17.3.1 Equação Geodésica na Métrica de Schwarzschild	250
17.3.2 Energia e Momento Angular	252
17.4 Queda Livre	252
17.5 Cálculo da Deflexão da Luz	252
17.6 Precessão do Perihélio de Mercúrio	252
17.7 Buracos Negros de Schwarzschild	252
17.8 Produtos Dobrados	252
17.8.1 Espaçotempos Estáticos	252
18 Campos de Jacobi	254
18.1 A Equação de Jacobi	254
18.2 Campos de Jacobi	256
18.3 Campos de Jacobi em Variedades de Curvatura Seccional Constante	261
18.4 Velocidade de Afastamento das Geodésicas e Curvatura Seccional	263
18.5 Pontos Conjungados	267
19 Causalidade	269
19.1 Orientação Temporal	269
19.2 Teoria de Causalidade para Espaçotempos	270
19.2.1 Cronologia	270
19.3 Apêndice: O Recobrimento Duplo Temporalmemente Orientável	271

Capítulo 1

Teoria da Relatividade Especial

Principais personagens e porque estão listados aqui:

- *Aristóteles* (384 BCE–322 BCE): primeira teoria sistemática do movimento.
- *Galileo Galilei* (1564–1642): princípio da relatividade galileana e corpos caem com a mesma aceleração no vácuo.
- *Isaac Newton* (1642–1726): primeira e segunda leis da dinâmica, espaço e tempo absolutos, teoria da gravitação universal (com o conceito de *ação à distância* associado, modelado nos fenômenos elétrico e magnéticos).
- *Michael Faraday* (1791–1867): lei da indução de Faraday (inspiração direta para a teoria da relatividade especial) e a idéia de campos (eliminando a necessidade da idéia de *ação à distância*).
- *James Clerk Maxwell* (1831–1879): equações de Maxwell para o eletromagnetismo; luz é uma onda eletromagnética.
- *Ernst Mach* (1838–1916): princípio da inércia de Mach (inspiração indireta para a teoria da relatividade geral; depois refutado).
- *Albert Einstein* (1879–1955): teoria da relatividade especial e teoria da relatividade geral.
- *Hermann Minkowski* (1864–1909): conceito do espaço-tempo.

1.1 Breve Introdução

Aristóteles dividiu o movimento em dois tipos, *forçado* e *natural*. O melhor exemplo de um movimento natural é o de um corpo que cai. De acordo com Aristóteles, este é um movimento natural, não sendo causado por nada. Embora esta divisão pareça bastante simplista, veremos que algumas das idéias de Aristóteles em um certo sentido tiveram um renascimento na Teoria da Relatividade Geral (veja [Petkov], p. 16 e pp. 21–22).

O modelo para o Universo que Aristóteles adotou foi o *geocêntrico* (a Terra no centro do Universo e tudo o mais girando em torno dela). Galileu adotou o modelo *heliocêntrico* de Copérnico e para justificá-lo foi forçado a abandonar a física de Aristóteles. Ele realizou uma série de experimentos, reais e imaginários (*thought experiments*, uma expressão que ficou consagrada de Einstein em diante) para chegar a duas conclusões fundamentais: é impossível determinar quando se está em repouso ou em movimento e a aceleração com que um corpo cai no vácuo é constante, independentemente de seu peso. Para justificar o primeiro princípio, conhecido como *relatividade galileana*, ele usou a experiência de uma pessoa dentro da cabine fechada de um

navio: quer o navio encontre-se em movimento ou repouso, todos os objetos dentro do navio se comportarão da mesma forma. Assim, se a água cai de um balde, ela cai na vertical, mesmo quando o navio se move, e se houver borboletas voando, elas não vão se acumular na traseira da cabine (veja [Topper], Cap. 1, pp. 3–7 e [Petkov], Cap 2, pp. 23–28). Quanto ao segundo princípio, é um debate histórico se Galileo realmente realizou um experimento de cima da Torre de Pisa, já que ele não há nenhum documento dele afirmando isso. O que se conhece é o seu célebre experimento imaginário, em que ele imaginou três corpos com o mesmo peso caindo portanto com a mesma aceleração no vácuo; se dois deles são colados por uma cola sem peso, isso não deve afetar a sua aceleração, mas agora temos dois corpos, um com o dobro do peso do primeiro. (Veja [Topper], Cap. 11, pp 85–88.)

Newton aceitou os dois princípios de Galileo. O primeiro ele imortalizou na sua primeira lei do movimento, definindo mais precisamente o movimento no princípio de relatividade de Galileo como sendo movimento retilíneo uniforme, enquanto que o segundo é consequência da sua segunda lei do movimento, onde ele introduziu o conceito de *massa inercial* (uma medida de resistência ao movimento) e da Lei da Gravitação Universal, onde ele introduziu o conceito de *massa gravitacional* (uma medida da intensidade da força gravitacional); o fato da aceleração dos corpos que caem ser constante era consequência da equivalência das duas. Na Lei da Gravitação Universal, Newton introduziu também o conceito de *ação à distância* ou força oculta, algo que foi aceito com muita dificuldade na época (apesar de Newton invocar os fenômenos elétrico e magnético como exemplos de que isso ocorre na realidade). O conceito de ação à distância só foi aceito porque sua teoria funcionava. Eventualmente, quando Faraday introduziu os conceitos de campos elétrico e magnético, o conceito de campo gravitacional seguiu-se naturalmente. Também a questão de porque os corpos se atraem, bastante debatida na época depois da sua introdução por Newton, foi de forma semelhante ignorada por ele, como escreveu no final das últimas edições do seu *Principia Mathematica*: “É suficiente que a gravidade realmente existe e age de acordo com as leis que estabelecemos e é suficiente para explicar todos os movimentos dos corpos celestes e do nosso mar [isto é, as marés].”; ou seja, não pergunte mais, ou indague questões físicas, mais profundas, pois a matemática funciona e isto é suficiente (veja [Topper], Cap. 4, pp. 25–28). Foi o primeiro exemplo em Física da atitude “cale a boca e faça as contas” (veja [Petkov], Cap. 2, p. 2). A atitude de Newton foi aceita até que Einstein formulou a Teoria da Relatividade Geral.

1.2 Dedução da existência de ondas eletromagnéticas através das equações de Maxwell

Lembramos as equações de Maxwell no vácuo (1873):

$$\begin{aligned} \operatorname{div} \mathbf{E} &= \frac{\rho}{\varepsilon_0}, \\ \operatorname{rot} \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \operatorname{div} \mathbf{B} &= 0, \\ \operatorname{rot} \mathbf{B} &= \mu_0 \left(\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right), \end{aligned}$$

que também podem ser escritas na notação

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0}, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \times \mathbf{B} &= \mu_0 \left(\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right), \end{aligned}$$

onde \mathbf{E} é o campo elétrico, \mathbf{B} o campo magnético, ρ a densidade (volumétrica) de carga elétrica, ε_0 a constante de permissividade elétrica do vácuo, μ_0 a constante de permeabilidade magnética do vácuo e \mathbf{J} a densidade de corrente elétrica (veja [Fleisch1], para uma introdução curta mas muito boa sobre as equações de Maxwell).

1.1 Lema. *Seja V um campo vetorial. Então*

$$\begin{aligned}\operatorname{rot}(\operatorname{rot} \mathbf{V}) &= \operatorname{grad}(\operatorname{div} \mathbf{V}) - \operatorname{div}(\operatorname{grad} \mathbf{V}) \\ &= \nabla(\operatorname{div} \mathbf{V}) - \Delta \mathbf{V}.\end{aligned}$$

Em outra notação,

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{V}) &= \nabla(\nabla \cdot \mathbf{V}) - \nabla \cdot (\nabla \mathbf{V}) \\ &= \nabla(\nabla \cdot \mathbf{V}) - \nabla^2 \mathbf{V}.\end{aligned}$$

Prova: Escrevendo $\mathbf{V} = (V_1, V_2, V_3)$, temos

$$\nabla \times \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_1 & V_2 & V_3 \end{vmatrix} = \left(\frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z}, \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x}, \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \right),$$

donde

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{V}) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} & \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} & \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{vmatrix} \\ &= \left(\frac{\partial^2 V_2}{\partial y \partial x} - \frac{\partial^2 V_1}{\partial y^2} - \frac{\partial^2 V_1}{\partial z^2} + \frac{\partial^2 V_3}{\partial z \partial x}, \right. \\ &\quad - \frac{\partial^2 V_2}{\partial x^2} + \frac{\partial^2 V_1}{\partial x \partial y} + \frac{\partial^2 V_3}{\partial z \partial y} - \frac{\partial^2 V_2}{\partial z^2}, \\ &\quad \left. \frac{\partial^2 V_1}{\partial x \partial z} - \frac{\partial^2 V_3}{\partial x^2} - \frac{\partial^2 V_3}{\partial y^2} + \frac{\partial^2 V_2}{\partial y \partial z} \right) \\ &= \left(\frac{\partial^2 V_1}{\partial x^2} + \frac{\partial^2 V_2}{\partial y \partial x} + \frac{\partial^2 V_3}{\partial z \partial x} - \frac{\partial^2 V_1}{\partial x^2} - \frac{\partial^2 V_1}{\partial y^2} - \frac{\partial^2 V_1}{\partial z^2}, \right. \\ &\quad \frac{\partial^2 V_1}{\partial x \partial y} + \frac{\partial^2 V_2}{\partial y^2} + \frac{\partial^2 V_3}{\partial z \partial y} - \frac{\partial^2 V_2}{\partial x^2} - \frac{\partial^2 V_2}{\partial y^2} - \frac{\partial^2 V_2}{\partial z^2}, \\ &\quad \left. \frac{\partial^2 V_1}{\partial x \partial z} + \frac{\partial^2 V_2}{\partial y \partial z} + \frac{\partial^2 V_3}{\partial z^2} - \frac{\partial^2 V_3}{\partial x^2} - \frac{\partial^2 V_3}{\partial y^2} - \frac{\partial^2 V_3}{\partial z^2} \right) \\ &= \left(\frac{\partial}{\partial x} \left(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z} \right) - \Delta V_1, \right. \\ &\quad \frac{\partial}{\partial y} \left(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z} \right) - \Delta V_2, \\ &\quad \left. \frac{\partial}{\partial z} \left(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z} \right) - \Delta V_3 \right) \\ &= \nabla(\operatorname{div} \mathbf{V}) - \Delta \mathbf{V}.\end{aligned}$$

■

1.2 Teorema (Existência de Ondas Eletromagnéticas). *Sejam \mathbf{E} e \mathbf{B} os campos elétrico e magnético, respectivamente. Então, no vácuo, em um ambiente livre de cargas elétricas, os campos elétrico e magnético satisfazem a equação da onda homogênea:*

$$\begin{aligned}\frac{\partial^2 \mathbf{E}}{\partial t^2} &= \frac{1}{\mu_0 \varepsilon_0} \Delta \mathbf{E}, \\ \frac{\partial^2 \mathbf{B}}{\partial t^2} &= \frac{1}{\mu_0 \varepsilon_0} \Delta \mathbf{B}.\end{aligned}$$

Prova. Do lema anterior temos que

$$\begin{aligned}\operatorname{rot}(\operatorname{rot} \mathbf{E}) &= \nabla(\operatorname{div} \mathbf{E}) - \Delta \mathbf{E}, \\ \operatorname{rot}(\operatorname{rot} \mathbf{B}) &= \nabla(\operatorname{div} \mathbf{B}) - \Delta \mathbf{B}.\end{aligned}$$

No vácuo, em um ambiente livre de cargas elétricas, a primeira e terceira equações de Maxwell (leis de Gauss para o campo elétrico e para o campo magnético, respectivamente) se tornam

$$\begin{aligned}\operatorname{div} \mathbf{E} &= 0, \\ \operatorname{div} \mathbf{B} &= 0,\end{aligned}$$

logo

$$\begin{aligned}\operatorname{rot}(\operatorname{rot} \mathbf{E}) &= -\Delta \mathbf{E}, \\ \operatorname{rot}(\operatorname{rot} \mathbf{B}) &= -\Delta \mathbf{B}.\end{aligned}\tag{1.1}$$

Por outro lado, no vácuo, em um ambiente livre de cargas elétricas, a segunda e quarta equações de Maxwell (lei de indução de Faraday e lei de Ampère com a correção de Maxwell)

$$\begin{aligned}\operatorname{rot} \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \operatorname{rot} \mathbf{B} &= \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}.\end{aligned}$$

Portanto,

$$\begin{aligned}\operatorname{rot}(\operatorname{rot} \mathbf{E}) &= -\frac{\partial}{\partial t}(\operatorname{rot} \mathbf{B}) = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}, \\ \operatorname{rot}(\operatorname{rot} \mathbf{B}) &= \mu_0 \varepsilon_0 \frac{\partial}{\partial t}(\operatorname{rot} \mathbf{E}) = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}.\end{aligned}\tag{1.2}$$

Reunindo (1.1) e (1.2), obtemos o resultado desejado. ■

As soluções das equações da onda do Teorema 1.2 são chamadas **ondas eletromagnéticas**. Pode-se mostrar, usando as equações de Maxwell, que as ondas eletromagnéticas são *ondas transversais*, ou seja, as vibrações dos campos elétrico e magnético são na direção perpendicular à direção de propagação da onda (para uma dedução simples e curta veja, por exemplo, [Woodhouse], Seções 3.3 e 3.4, pp. 42–44; as ondas de som são um exemplo de *ondas longitudinais*, em que as vibrações são na mesma direção de propagação da onda).

1.3 Corolário. *A velocidade de uma onda eletromagnética no vácuo é*

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 2,99792458 \times 10^8 \text{ m/s.}$$

Prova. Pois

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = c^2 \Delta \mathbf{E},$$

$$\frac{\partial^2 \mathbf{B}}{\partial t^2} = c^2 \Delta \mathbf{B}.$$

■

Como este valor é igual ao da velocidade da luz, Maxwell concluiu que a luz é uma onda eletromagnética. Vários experimentos posteriores corroboraram esta idéia. As equações de Maxwell foram as primeiras leis da Física a conterem uma velocidade como uma constante fundamental.

1.2.1 O Éter

Na teoria ondulatória da luz de Fresnel (1821), a luz era uma onda mecânica que se propaga em um meio chamado o *éter luminífero*. Ou seja, as vibrações da luz correspondiam a vibrações das moléculas etéreas. A velocidade de propagação c da luz era uma propriedade do éter. Para conter ondas transversais, o éter tinha que ser um meio material elástico, embora fosse estranho que não existissem ondas longitudinais neste meio. Além disso, para produzir ondas com esta enorme velocidade de propagação (em um segundo, a luz dá 7 voltas ao redor da Terra), o éter tinha que ser extremamente rígido. O éter tinha que preencher todo o Universo, pois a luz se propaga em todo lugar. Era lógico considerar o éter como estando em repouso em relação ao espaço absoluto de Newton, funcionando como uma espécie de materialização deste.

No entanto, uma tal substância onipresente deveria produzir outros efeitos mecânicos, além dos fenômenos luminosos. Como podem os planetas se moverem através do éter sem perder energia? O éter penetraria através dos corpos móveis sem perturbá-los ou seria arrastado por eles? O éter parecia ser uma substância intangível sem qualquer outro efeito além de ser o lugar onde os fenômenos luminosos ocorriam ([Springer], Seção 1.3.3). Todo o conceito do éter soa bastante estranho para nós, até que lembramos que hoje em dia a teoria da *matéria escura*, formada de misteriosas partículas que interagem apenas gravitacionalmente, é bastante aceita.

Da mesma forma Maxwell viu suas ondas eletromagnéticas como um fenômeno mecânico em um meio propagador. Após reconhecer que a luz era uma onda eletromagnética, ele identificou este meio propagador com o éter de Fresnel. Portanto ele acreditava que suas equações eram válidas em um referencial fixado neste meio.

De maneira semelhante, em um experimento imaginário (*thought experiment*) célebre, que Einstein afirmou em sua autobiografia ter tido na idade de 16 anos, ele imaginou o que aconteceria se viajássemos em um raio de luz. Seguindo [Topper] (uma excelente história científica sobre a descoberta de Einstein da Teoria da Relatividade) vamos descrever o experimento imaginário de Einstein de uma forma diferente: imagine-se estando na cabine do piloto de uma nave espacial, tendo atrás da poltrona uma fonte de luz. À medida que a nave fosse acelerando, haveria uma hora em que ela ultrapassaria a velocidade da luz e então a cabine repentinamente ficaria às escuras. Desta forma seria possível determinar se a nave estava em repouso ou não. Mas isso significaria abandonar o princípio de relatividade de Galileu, também fundamental na Mecânica Newtoniana, que afirmava que não se podia distinguir quando se estava em repouso e quando se estava em movimento retilíneo uniforme. Para não violar este princípio, a luz teria que se mover com uma velocidade igual à soma da velocidade da nave e da velocidade da luz.

Vários experimentos no final do século XIX falharam em descobrir qualquer variação na velocidade da luz devido ao movimento da fonte, notavelmente o experimento de Michelson-Morley (1881; veja [Springer], Seção 1.3, para uma descrição breve deste e de outros experimentos e [Ferraro], Capítulo 2, para uma discussão bem mais detalhada pelo mesmo autor).

1.3 Referenciais Inerciais

1.4 Definição. Um referencial é um sistema de coordenadas no espaço e no tempo. □

Desta forma, associamos \mathbb{R}^4 ao espaço e ao tempo. Em geral, denotaremos as coordenadas de um referencial por (t, x, y, z) . Outras notações para as coordenadas serão introduzidas mais tarde.

Observe que com esta associação, por enquanto o espaço e o tempo formam juntos apenas um espaço vetorial de dimensão 4, mas não há uma unidade entre eles. O espaço é medido de uma forma diferente da forma em que o tempo é medido, as unidades de medida de espaço (metros) são diferentes das unidades de medida de tempo (segundos). Em outras palavras este espaço vetorial ainda não tem uma métrica associada.

No que se segue, o termo *partícula-teste* se refere a uma partícula com massa suficientemente pequena para não ter influências gravitacionais perceptíveis dentro de uma precisão previamente fixada. Dizemos que ela é *livre* se ela não está sujeita a nenhuma força.

1.5 Definição. Dizemos que um referencial para uma certa região de espaço e de tempo é um **referencial inercial** se naquela região de espaço e tempo, de acordo com uma precisão previamente especificada, toda partícula-teste livre inicialmente em repouso com relação àquele referencial permanece em repouso, e toda partícula-teste inicialmente em movimento com relação àquele referencial continua seu movimento ao longo de uma reta (isto é, sem mudança de velocidade em módulo ou direção). \square

Referenciais iniciais são eminentemente *locais*. Um elevador suficientemente pequeno em queda livre sob a ação de um campo gravitacional uniforme (por exemplo, o campo gravitacional de um planeta ou estrela, a uma distância razoável da superfície) durante um intervalo de tempo suficientemente pequeno é um referencial inercial, dentro de uma precisão fixada, porque todas as partículas-teste caem com a mesma aceleração. À medida que o tempo passa, efeitos de maré passam a predominar e partículas-teste inicialmente em repouso na direção do campo gravitacional começam a se distanciar (partículas-teste mais próximas da fonte do campo gravitacional sofrem aceleração maior), enquanto que partículas-teste inicialmente em repouso situadas em uma direção perpendicular à direção do campo gravitacional começam a se aproximar (já que cada partícula cai seguindo a sua linha do campo gravitacional). Outro referencial inercial é uma região do espaço suficientemente longe da fonte de um campo gravitacional, durante algum intervalo de tempo razoável. Para uma discussão mais detalhada sobre referenciais iniciais, veja as seções 2.1-2.5 do livro Spacetime Physics [Taylor-Wheeler], pp. 25–36.

1.6 Definição. Um ponto no espaço e no tempo \mathbb{R}^4 é chamado um **evento**.

Uma **observação** feita por um observador em um referencial inercial é simplesmente o ato de atribuir a qualquer evento as suas coordenadas (t, x, y, z) , ou seja, as coordenadas espaciais x, y, z da localização espacial da sua ocorrência e o instante de tempo t lido no relógio situado neste ponto em que o evento ocorre. \square

1.4 Princípio da Relatividade

1.7 Axioma (Princípio da Relatividade de Einstein). *Todas as leis da Física são as mesmas em qualquer referencial inercial.*

O Princípio da Relatividade diz que uma vez que as leis da Física foram estabelecidas em um referencial inercial, elas podem ser aplicadas sem modificação em qualquer outro referencial inercial. A forma matemática das leis da Física e os valores numéricos das constantes físicas básicas que estas leis contêm são as mesmas em qualquer referencial inercial. No que se refere às leis da Física, todos os referenciais iniciais são equivalentes. Consequentemente,

1.8 Corolário. *Nenhum teste das leis da Física produz uma maneira de distinguir um referencial inercial de outro referencial inercial.*

1.9 Corolário. *As equações de Maxwell no vácuo são as mesmas em qualquer referencial inercial.*

Daí o título do artigo original em que Einstein (1905) propôs a Teoria da Relatividade Especial: *On the Electrodynamics of Moving Bodies*.

1.10 Corolário. A velocidade da luz no vácuo é a mesma em qualquer referencial inercial.

Assim, quando se mede a velocidade da luz em um referencial inercial sempre se obtém o mesmo valor c , dado no Corolário 1.3. No artigo original de Einstein, este resultado é enunciado como um segundo postulado.

O nome *Teoria da Relatividade* a rigor não é apropriado. Na verdade o Princípio da Relatividade diz que as leis da Física são absolutas, no sentido de que não dependem do referencial inercial em que são testadas (por exemplo, a velocidade da luz tem o mesmo valor absoluto, que não é relativo ao referencial inercial em que é medido). Um melhor nome para a Teoria da Relatividade seria a *Teoria da Invariância*. No entanto, uma consequência do Princípio de Relatividade, como veremos, é a *relatividade do espaço e do tempo*, isto é, medidas do *deslocamento espacial* (separação entre dois pontos no espaço) e do *intervalo temporal* (separação entre dois momentos no tempo) dependem do referencial inercial em que elas são realizadas. Dada que esta consequência não intuitiva foi a maior surpresa da Teoria da Relatividade quando ela foi formulada e a maior dificuldade em aceitá-la, ela recebeu o nome deste fenômeno.

No que se segue, escolhemos unidades de espaço e tempo de tal forma que $c = 1$.

1.5 Invariância do Intervalo Espaçotemporal

1.11 Definição. O intervalo entre dois eventos quaisquer do espaço-tempo separados por incrementos $\Delta t, \Delta x, \Delta y, \Delta z$ é definido por

$$\Delta s^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2$$

□

Observe que podemos ter

$$\Delta s^2 > 0$$

ou

$$\Delta s^2 < 0$$

ou

$$\Delta s^2 = 0.$$

1.12 Lema. $\Delta s^2 = 0$ se e somente se os dois eventos ocorrem um após o outro na velocidade da luz.

Consequentemente, se \mathcal{O} e $\bar{\mathcal{O}}$ são dois referenciais inerciais, então

$$\Delta s^2 = 0$$

se e somente se

$$\Delta \bar{s}^2 = 0.$$

Prova. Pois, pelo Princípio de Relatividade, a velocidade da luz é constante igual a 1 em qualquer referencial inercial. ■

Eventos que ocorrem na velocidade da luz são, por exemplo, os eventos emissão de um fóton e absorção do mesmo.

No que se segue, usaremos a seguinte notação:

$$\begin{aligned}\Delta t &= \Delta x^0, \\ \Delta x &= \Delta x^1, \\ \Delta y &= \Delta x^2, \\ \Delta z &= \Delta x^3.\end{aligned}$$

1.13 Teorema (Invariância do Intervalo). Sejam \mathcal{O} e $\bar{\mathcal{O}}$ dois referenciais inerciais. Então

$$\Delta \bar{s}^2 = \Delta s^2.$$

Prova. [Shutz] Observe que a relação entre as coordenadas de \mathcal{O} e $\bar{\mathcal{O}}$ é *linear*, porque caso contrário, existiria um objeto em movimento retilíneo uniforme em um referencial (trajetória reta) que não estaria em movimento retilíneo uniforme no outro referencial (sua trajetória não seria uma reta), contrariando o princípio de relatividade (até mesmo galileano). Por simplicidade, assuma também que suas origens coincidem, isto é, os eventos $t = x = y = z = 0$ e $\bar{t} = \bar{x} = \bar{y} = \bar{z} = 0$ são os mesmos.

Da hipótese de linearidade, segue que $\Delta\bar{t}, \Delta\bar{x}, \Delta\bar{y}, \Delta\bar{z}$ são funções lineares de $\Delta t, \Delta x, \Delta y, \Delta z$, logo $\Delta\bar{s}^2$ é uma função quadrática de $\Delta t, \Delta x, \Delta y, \Delta z$, digamos

$$\Delta\bar{s}^2 = \sum_{i,j=0}^3 M_{ij} \Delta x^i \Delta x^j. \quad (1.3)$$

Observe que em princípio podemos ter $M_{ij} = M_{ij}(\mathbf{v})$, onde \mathbf{v} é a velocidade relativa entre os dois referenciais (independe dos eventos). Podemos assumir, sem perda de generalidade, que $M_{ij} = M_{ji}$ (trocando, se necessário, M_{ij} por $(M_{ij} + M_{ji})/2$).

Para dois eventos em que

$$\Delta\bar{s}^2 = \Delta s^2 = 0$$

temos

$$M_{00} (\Delta t)^2 + 2 \left(\sum_{i=1}^3 M_{0i} \Delta x^i \right) \Delta t + \sum_{i,j=1}^3 M_{ij} (\Delta x^i) (\Delta x^j) = 0. \quad (1.4)$$

Como $\Delta\bar{s}^2 = \Delta s^2 = 0$ continua valendo se substituirmos todos os Δx^i por $-\Delta x^i$, $i = 1, 2, 3$, obtemos

$$M_{00} (\Delta t)^2 - 2 \left(\sum_{i=1}^3 M_{0i} \Delta x^i \right) \Delta t + \sum_{i,j=1}^3 M_{ij} (\Delta x^i) (\Delta x^j) = 0.$$

Subtraindo uma equação da outra, segue que

$$\left(\sum_{i=1}^3 M_{0i} \Delta x^i \right) \Delta t = 0$$

Escolhendo

$$\begin{aligned} \Delta x^i &= \Delta t = 1, \\ \Delta x^j &= \Delta x^k = 0, \quad \text{se } j, k \neq i, \end{aligned}$$

concluímos que

$$M_{0i} = 0 \quad \text{para } i = 1, 2, 3 \quad (1.5)$$

e, portanto,

$$M_{00} (\Delta t)^2 + \sum_{i,j=1}^3 M_{ij} (\Delta x^i) (\Delta x^j) = 0.$$

Daí, com a mesma escolha acima, segue que

$$M_{ii} = -M_{00} \quad \text{para } i = 1, 2, 3.$$

Se $i \neq j$, escolhendo agora

$$\begin{aligned} \Delta x^i &= 1, \Delta x^j = 1, \Delta t = \sqrt{2}, \\ \Delta x^k &= 0, \quad \text{se } k \neq i, j, \end{aligned}$$

e

$$\begin{aligned}\Delta x^i &= 1, \Delta x^j = -1, \Delta t = \sqrt{2}, \\ \Delta x^k &= 0, \quad \text{se } k \neq i, j,\end{aligned}$$

obtemos, respectivamente,

$$\begin{aligned}2M_{00} + 2M_{00} + 2M_{ij} &= 0, \\ 2M_{00} + 2M_{00} - 2M_{ij} &= 0;\end{aligned}$$

subtraindo estas duas equações, concluímos que

$$M_{ij} = 0$$

se $i \neq j$.

Em outras palavras,

$$M_{ij} = -M_{00}\delta_{ij} \quad \text{para } i, j = 1, 2, 3, \quad (1.6)$$

e

$$\Delta \bar{s}^2 = M_{00} \left[(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \right] = -M_{00}(\mathbf{v}) \Delta s^2. \quad (1.7)$$

Denote

$$\phi(\mathbf{v}) = -M_{00}(\mathbf{v}). \quad (1.8)$$

Afirmamos que $\phi(\mathbf{v}) \equiv 1$.

Em primeiro lugar, observamos que ϕ é radial, de modo que

$$\phi(\mathbf{v}) = \phi(|\mathbf{v}|) = \phi(v). \quad (1.9)$$

De fato, escolhendo dois eventos que são simultâneos no referencial \mathcal{O} , isto é, $\Delta t = 0$ (por exemplo, o comprimento de uma barra) concluímos que

$$\begin{aligned}&(\text{distância entre dois eventos simultâneos em } \bar{\mathcal{O}}) \\ &= \phi(\mathbf{v}) (\text{distância entre dois eventos simultâneos em } \mathcal{O}).\end{aligned}$$

Como esta distância não pode depender da direção de \mathbf{v} (comprimentos não dependem da direção da velocidade, já que comprimentos perpendiculares à direção da velocidade não podem sofrer variações, conforme argumentaremos na Seção 1.7.3; veja também em maiores detalhes [Shutz], pp. 10–11), concluímos que $\phi(\mathbf{v}) = \phi(|\mathbf{v}|)$.

Agora considere três referenciais \mathcal{O} , $\bar{\mathcal{O}}$ e $\bar{\bar{\mathcal{O}}}$, sendo que $\bar{\mathcal{O}}$ move-se com velocidade \mathbf{v} com relação a \mathcal{O} , enquanto que $\bar{\bar{\mathcal{O}}}$ move-se com velocidade $-\mathbf{v}$ com relação a $\bar{\mathcal{O}}$, ou seja, os referenciais \mathcal{O} e $\bar{\bar{\mathcal{O}}}$ são idênticos. Segue que

$$\begin{aligned}\Delta \bar{s}^2 &= \phi(v) \Delta s^2, \\ \Delta \bar{\bar{s}}^2 &= \phi(v) \Delta \bar{s}^2,\end{aligned}$$

onde

$$\Delta \bar{\bar{s}}^2 = [\phi(v)]^2 \Delta s^2.$$

Mas os referenciais \mathcal{O} e $\bar{\bar{\mathcal{O}}}$ são idênticos, logo obtemos de fato que

$$\Delta s^2 = [\phi(v)]^2 \Delta s^2,$$

onde

$$\phi \equiv \pm 1.$$

Como o quadrado de comprimentos (isto é, $\Delta t = 0$) deve ser positivo em qualquer referencial inercial, segue que $\phi \equiv 1$. ■

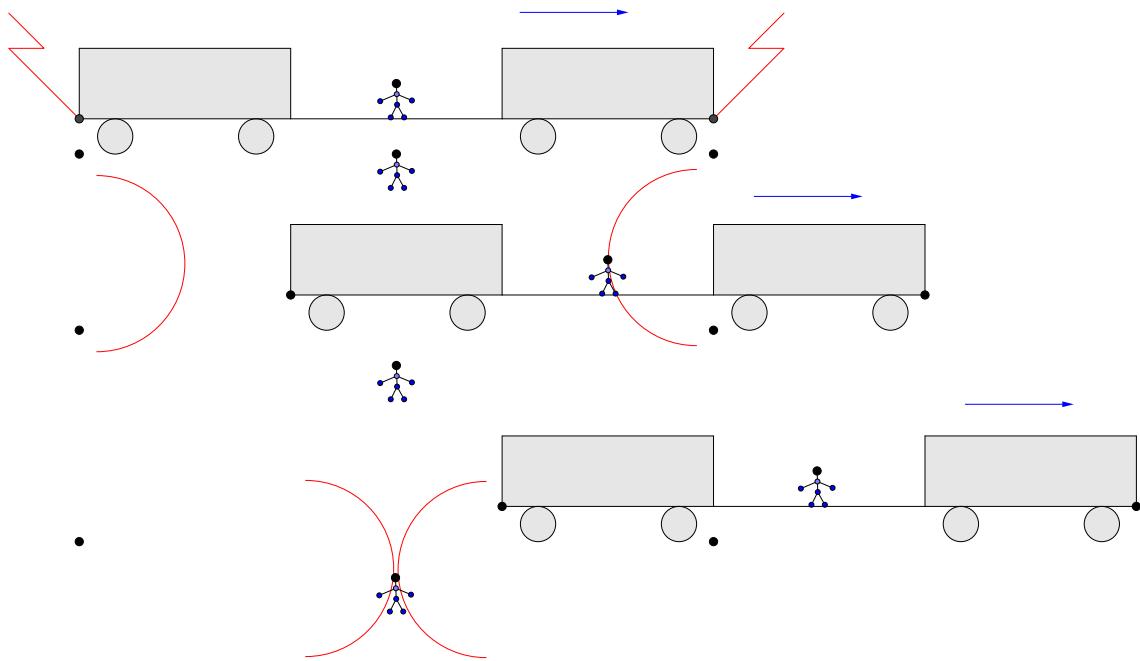
1.6 Relatividade da Simultaneidade

A maior parte dos paradoxos em Relatividade se originam de mal-entendidos com relação ao conceito de simultaneidade. A simultaneidade de dois eventos depende do referencial inercial em que eles são medidos (observados). Eventos que são simultâneos em um referencial inercial, não serão simultâneos em outro referencial inercial que se desloca em relação ao primeiro com velocidade v não perpendicular ao vetor que conecta as duas posições espaciais dos eventos, como veremos.

A relatividade da simultaneidade é ilustrada pelo *paradoxo do trem de Einstein*. Relâmpagos atingem ambas as extremidades de um trem que se move, deixando marcas enegrecidas no trem e no trilho. O brilho do impacto do relâmpago é emitido em todas as direções.

Um observador que se situa no chão fora do trem, no meio entre as duas marcas, conclui que os impactos foram simultâneos, já que os brilhos o atingem ao mesmo tempo.

Um observador que se situa no meio do trem conclui que os impactos não foram simultâneos, já que a velocidade da luz é a mesma em todos os referenciais e o brilho do impacto na frente do trem o atinge antes do brilho do impacto na parte traseira do trem.



1.7 Relatividade de Intervalos Espaciais e de Intervalos Temporais

1.7.1 Contração do Espaço

No paradoxo do trem, o observador no chão O mede o comprimento do trem como sendo o valor da medida entre as duas marcas enegrecidas no trilho, já que elas ocorreram simultaneamente. Já o observador dentro do trem O' não concorda com esta medição: dado que para ele a marca frontal foi feita antes da marca traseira, e o trem se moveu enquanto isso, para ele a medição feita pelo observador O deu uma medida menor para o trem.

É fácil entender isso usando a invariância do intervalo. Os dois eventos são: evento A , impacto do relâmpago com a parte dianteira do trem, e evento B , impacto do relâmpago com a parte traseira do trem. Digamos que estes dois eventos tem coordenadas (x_A, t_A) e (x_B, t_B) no referencial inercial do observador no chão O , e coordenadas (x'_A, t'_A) e (x'_B, t'_B) no referencial inercial do observador dentro do trem O' . Sejam

$$\begin{aligned}\Delta x &= x_B - x_A, \\ \Delta t &= t_B - t_A,\end{aligned}$$

e

$$\begin{aligned}\Delta x' &= x'_B - x'_A, \\ \Delta t' &= t'_B - t'_A,\end{aligned}$$

as separações espaciais e temporais destes dois eventos em cada um dos referenciais inerciais. Pela invariância do intervalo espaço-temporal,

$$(\Delta x')^2 - (\Delta t')^2 = (\Delta x)^2 - (\Delta t)^2.$$

Mas

$$\Delta t = 0,$$

já que os eventos são simultâneos no referencial inercial O , consequentemente

$$(\Delta x')^2 - (\Delta t')^2 = (\Delta x)^2$$

e

$$\Delta x' > \Delta x. \quad (1.10)$$

Assim, o comprimento de uma barra medido em um referencial em que ela está em repouso é sempre maior que o comprimento dela medido em um referencial em que ela se move. Este efeito é a chamada **contração de Lorenz**. O comprimento da barra medido em um referencial em que ela está em repouso é chamado o seu **comprimento próprio**. Portanto, o comprimento próprio é sempre **maior** que o comprimento medido em um referencial em que a barra está se movendo. Na Seção 2.4.3 obteremos uma expressão que nos permitirá calcular exatamente o quanto o comprimento é diminuído no referencial que se move; o fator de contração dependerá exclusivamente da velocidade deste.

1.7.2 Dilatação do Tempo

Imagine uma nave espacial viajando em velocidade uniforme ao sistema estelar de Canopus, situado 99 anos-luz de distância da Terra. Suponha que em relação ao referencial inercial Terra a nave espacial viaja à velocidade de 99/101 vezes a velocidade da luz, ou seja, 99 anos-luz de distância em 101 anos de tempo. Em outras palavras, no referencial Terra, a nave espacial leva 101 anos para chegar até Canopus. As coordenadas dos eventos A , saída da nave espacial da Terra, e B , chegada da nave a Canopus (ignorando o tempo gasto para acelerar a nave até a velocidade $(99/101)c$) no referencial inercial Terra são

$$\begin{aligned}(x_A, t_A) &= (0, 0), \\ (x_B, t_B) &= (99, 101).\end{aligned}$$

As coordenadas destes eventos no referencial inercial Nave são

$$\begin{aligned}(x'_A, t'_A) &= (0, 0), \\ (x'_B, t'_B) &= (0, T),\end{aligned}$$

pois no referencial inercial Nave os dois eventos, saída da Terra e chegada em Canopus, ocorrem na mesma posição, isto é, na própria nave. Para calcular o tempo T decorrido entre os dois eventos de acordo com o

referencial Nave, calculamos as separações espaciais e temporais nos dois referenciais inerciais e usamos a invariância do intervalo espaço-temporal. Temos

$$\begin{aligned}\Delta x &= x_B - x_A = 99, \\ \Delta t &= t_B - t_A = 101,\end{aligned}$$

e

$$\begin{aligned}\Delta x' &= x'_B - x'_A = 0, \\ \Delta t' &= t'_B - t'_A = T.\end{aligned}$$

Como

$$(\Delta t')^2 - (\Delta x')^2 = (\Delta t)^2 - (\Delta x)^2,$$

e

$$\Delta x' = 0$$

segue que

$$(\Delta t')^2 - (\Delta x')^2 = (\Delta t)^2 - (\Delta x)^2,$$

isto é,

$$(\Delta t')^2 = (\Delta t)^2 - (\Delta x)^2,$$

e consequentemente,

$$\Delta t' < \Delta t. \quad (1.11)$$

No caso, mais especificamente temos

$$\begin{aligned}T^2 &= 101^2 - 99^2 \\ &= 10.201 - 9.801 \\ &= 400\end{aligned}$$

e portanto

$$T = 20.$$

Enquanto na Terra se passaram 101 anos, dentro da nave espacial, decorreram apenas 20 anos. Assim, o intervalo de tempo entre dois eventos medido em um referencial em que eles ocorrem na mesma posição é sempre menor que o intervalo de tempo entre os dois eventos medido em um referencial que se move em relação ao primeiro. Este efeito é a chamada **dilatação do tempo**. O intervalo de tempo entre dois eventos medido em um referencial em que eles ocorrem na mesma posição é chamado o seu **tempo próprio**. Portanto, o tempo próprio é sempre **menor** que o tempo medido em um referencial que está se movendo em relação a este referencial. Na Seção 2.4.1 obteremos uma expressão que nos permitirá calcular exatamente o quanto o tempo é dilatado no referencial que se move; o fator de dilatação dependerá exclusivamente da velocidade deste.

Observe que quanto mais próxima da velocidade da luz é a velocidade da nave, menor o tempo decorrido na nave. No limite, se a nave se movesse à velocidade da luz (o que é impossível para partículas materiais, como veremos no final deste capítulo), o intervalo de tempo seria zero, pois neste caso teríamos

$$\Delta x = \Delta t = 99,$$

de modo que

$$(\Delta t')^2 = (\Delta t)^2 - (\Delta x)^2 = 0.$$

1.7.3 Invariância das Dimensões Transversais

As dimensões de objetos que se movem em direções transversais à direção de movimento relativo de dois referenciais iniciais tem a mesma medida nos dois referenciais.

Isso decorre de um argumento de simetria: não há nenhuma direção privilegiada segundo a qual pode haver uma contração ou dilatação (fisicamente, dizemos que o espaço é **isotrópico**, isto é, é o mesmo em todas as direções). Imagine dois cilindros ocos de mesmo raio e comprimento viajando em velocidade constante, um na direção do outro, ao longo do seu eixo comum de simetria. Imagine que um dos cilindros está em repouso (cilindro *A*) em relação a um certo referencial inercial. Se o outro cilindro (cilindro *B*) for medido como tendo o seu raio contraído [dilatando] em relação a este referencial, então ele passará por dentro [por fora] do cilindro *A*. Por outro lado, em relação a um segundo referencial em repouso ao cilindro *B* exatamente o contrário se verificará: o cilindro *A* passará por dentro [por fora] do cilindro *B*. Isto é intolerável do ponto de vista da lógica. Uma objeção mais fundamental é esta escolha arbitrária do eixo de contração [dilatação]. Se os dois cilindros estão viajando paralelamente, qual seria o eixo ao longo do qual se verificaria a contração [dilatação] em relação a um referencial inercial externo aos cilindros, em relação ao qual eles se movem? Os cilindros se contrairiam [dilatariam] cada um em relação ao seu eixo de movimento ou em relação à reta equidistante deles? Para o observador inercial externo, qualquer eixo paralelo ao movimento dos cilindros é equivalente. Veja [[Taylor-Wheeler](#)], Seção 3.6, págs. 65–67, para uma discussão detalhada.

Da mesma forma, dois eventos com separações espaciais transversais à direção de movimento relativo de dois referenciais iniciais são simultâneos em um dos referenciais se e somente se eles são simultâneos no outro referencial.

Isso decorre também de um argumento de simetria. Se, em um cilindro viajando com velocidade uniforme em relação a um referencial inercial externo, dois eventos ocorrem simultaneamente na periferia do cilindro de acordo com o referencial inercial do próprio cilindro, como decidir qual evento ocorreria antes ou depois no referencial inercial externo? Também no referencial externo os dois eventos devem ser simultâneos.

Capítulo 2

O Espaçotempo de Minkowski

2.1 Produto Interno e Formas Quadráticas

2.1 Definição. Seja \mathcal{V} um espaço vetorial real de dimensão finita.

Uma **forma bilinear** em \mathcal{V} é uma aplicação $f : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ que é linear em cada variável, isto é,

$$\begin{aligned} f(\alpha v + \beta w, u) &= \alpha f(v, u) + \beta f(w, u), \\ f(u, \alpha v + \beta w) &= \alpha f(u, v) + \beta f(u, w), \end{aligned}$$

para todos $u, v, w \in \mathcal{V}$ e para todos $\alpha, \beta \in \mathbb{R}$. Uma forma bilinear f é **simétrica** se

$$f(v, w) = f(w, v)$$

para todos $v, w \in \mathcal{V}$. Uma forma bilinear B é **não-degenerada** se

$$f(v, u) = 0$$

para todo $u \in \mathcal{V}$ implicar $v = 0$ e se

$$f(u, w) = 0$$

para todo $u \in \mathcal{V}$ implicar $w = 0$.

Um **produto interno** (ou **métrica**) em \mathcal{V} é uma forma bilinear simétrica não-degenerada, que geralmente denotaremos por g ou por $\langle \cdot, \cdot \rangle$.

Um produto interno $\langle \cdot, \cdot \rangle$ é **positivo definido** se

$$\langle v, v \rangle > 0,$$

para todo $v \in \mathcal{V}$ e **negativo definido** se

$$\langle v, v \rangle < 0,$$

para todo $v \in \mathcal{V}$.

Dizemos que dois vetores v, w são **ortogonais** se $\langle v, w \rangle = 0$. \square

2.2 Exemplo. Seja \mathcal{V} um espaço vetorial real de dimensão finita e

$$\mathcal{B} = \{e_1, \dots, e_n\}$$

uma base para \mathcal{V} . Para $0 \leq k \leq n$ e $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, $\alpha_i > 0$ para todo i , definimos o produto interno

$$\langle v, w \rangle = - \sum_{i=1}^k \alpha_i v^i w^i + \sum_{i=k+1}^n \alpha_i v^i w^i,$$

onde $v = (v^1, \dots, v^n)$ e $w = (w^1, \dots, w^n)$ são as coordenadas de v e w em relação à \mathcal{B} .

De fato, esta é claramente uma forma bilinear simétrica. Se

$$\langle v, w \rangle = 0$$

para todo $w \in \mathcal{V}$, isso vale em particular para os vetores da base canônica e_1, \dots, e_n . Mas $\langle v, e_i \rangle = \pm \alpha_i v^i$, logo $v^i = 0$ para todo i e portanto ela é não-degenerada. Se $k = 0$ este produto interno é positivo definido, se $k = n$ este produto interno é negativo definido e se $0 < k < n$ este produto interno não é nem positivo definido nem negativo definido, pois vetores da forma

$$v = (v^1, \dots, v^k, 0, \dots, 0)$$

satisfazem $\langle v, v \rangle < 0$, enquanto que vetores v da forma

$$v = (0, \dots, 0, v^{k+1}, \dots, v^n)$$

satisfazem $\langle v, v \rangle > 0$.

Em geral, fixada uma base em um espaço vetorial \mathcal{V} , se os seus vetores são expressos em coordenadas em relação a esta base, a fórmula acima define um produto interno em \mathcal{V} . \square

2.3 Definição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$.

A **forma quadrática** induzida por $\langle \cdot, \cdot \rangle$ (ou associada a $\langle \cdot, \cdot \rangle$) é a aplicação $q : \mathcal{V} \rightarrow \mathbb{R}$ definida por

$$q(v) = \langle v, v \rangle.$$

Dizemos que v é um vetor do tipo **luz** se $q(v) = 0$.

Dizemos que v é um vetor **unitário** se $q(v) = \pm 1$.

Dado um espaço vetorial com produto interno \mathcal{V} , dizemos que uma base $\mathcal{B} = \{e_1, \dots, e_n\}$ para \mathcal{V} é **ortonormal** se os seus vetores são dois a dois ortogonais e unitários. \square

2.4 Exemplo. Para a forma quadrática q induzida pelo produto interno g em \mathcal{V} do Exemplo 2.2 com $\alpha_i = 1$ para todo i , isto é,

$$\langle v, w \rangle = - \sum_{i=1}^k v^i w^i + \sum_{i=k+1}^n v^i w^i,$$

onde $v = (v^1, \dots, v^n)$ e $w = (w^1, \dots, w^n)$ são as coordenadas de v e w em relação a uma base $\mathcal{B} = \{e_1, \dots, e_n\}$ fixada, de modo que

$$q(v) = - \sum_{i=1}^k (v^i)^2 + \sum_{i=k+1}^n (v^i)^2,$$

a própria base \mathcal{B} é ortonormal, com

$$q(e_i) = \begin{cases} -1 & \text{se } i = 1, \dots, k, \\ 1 & \text{se } i = k+1, \dots, n. \end{cases}$$

\square

2.5 Exemplo. Se q é uma forma quadrática induzida por um produto interno positivo definido ou por um produto interno negativo definido, então v é um vetor do tipo luz se e somente se $v = 0$. Se q é induzida por um produto interno que não é definido, então mesmo que $v \neq 0$ pode acontecer que v seja um vetor do tipo luz. Por exemplo, para a forma quadrática do Exemplo 2.4, se $k \neq 0$ temos

$$q(e_i + e_j) = 0$$

para todos $1 \leq i \leq k$ e $k+1 \leq j \leq n$. De fato, é possível obter uma base para \mathcal{V} composta inteiramente de vetores do tipo luz. No caso $k=1$, basta tomar

$$\mathcal{B}' = \{e_1 + e_2, e_1 - e_2, e_1 + e_3, \dots, e_1 + e_n\}.$$

Para ver que \mathcal{B}' é uma base, basta verificar que os n vetores que a formam são LI. Se

$$\alpha_1(e_1 + e_2) + \alpha_2(e_1 - e_2) + \alpha_3(e_1 + e_3) + \dots + \alpha_n(e_1 + e_n) = 0,$$

então

$$\left(\sum_{i=1}^n \alpha_i \right) e_1 + (\alpha_1 - \alpha_2) e_2 + \sum_{i=3}^n \alpha_i e_i = 0.$$

Como \mathcal{B} é uma base, segue que

$$\begin{aligned} \sum_{i=1}^n \alpha_i &= 0, \\ \alpha_1 - \alpha_2 &= 0, \\ \alpha_3 &= \dots = \alpha_n = 0, \end{aligned}$$

onde $\alpha_i = 0$ para todo i . \square

Note que

$$q(v) = q(-v),$$

para todo $v \in \mathcal{V}$, pois

$$q(-v) = \langle -v, -v \rangle = \langle v, v \rangle = q(v).$$

De modo geral,

$$q(\alpha v) = \alpha^2 q(v).$$

2.6 Proposição (Identidade Polar). *Um produto interno é completamente determinado por sua forma quadrática associada. Mais especificamente, se q é a forma quadrática induzida pelo produto interno $\langle \cdot, \cdot \rangle$, então*

$$\langle v, w \rangle = \frac{1}{2} [q(v+w) - q(v) - q(w)].$$

e

$$\langle v, w \rangle = \frac{1}{4} [q(v+w) - q(v-w)]$$

Em particular, toda forma bilinear simétrica induz uma única forma quadrática.

Prova. Pois,

$$\begin{aligned} q(v+w) - q(v) - q(w) &= \langle v+w, v+w \rangle - \langle v, v \rangle - \langle w, w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle - \langle v, v \rangle - \langle w, w \rangle \\ &= 2\langle v, w \rangle, \end{aligned}$$

e

$$\begin{aligned} q(v+w) - q(v-w) &= \langle v+w, v+w \rangle - \langle v-w, v-w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle - [\langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle] \\ &= 4\langle v, w \rangle. \end{aligned}$$



2.7 Definição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$ e q sua forma quadrática associada. A **norma** de um vetor $v \in \mathcal{V}$ é definida por

$$\|v\| = \sqrt{|\langle v, v \rangle|} = \sqrt{|q(v)|}.$$

□

Observe que a norma de um vetor é sempre um número positivo ou zero. Vetores do tipo luz tem norma zero.

2.8 Teorema (Teorema de Sylvester). *Sejam \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$ e q sua forma quadrática associada. Então \mathcal{V} possui uma base ortonormal.*

Além disso, o número de vetores e_i de qualquer base ortonormal $\mathcal{B} = \{e_1, \dots, e_n\}$ para \mathcal{V} tais que

$$q(e_i) = -1$$

é o mesmo.

Prova. Para mostrar a existência de uma base ortonormal faremos indução sobre a dimensão de \mathcal{V} . Seja $n = \dim \mathcal{V}$ e verifiquemos primeiro o caso $n = 1$. Afirmamos que existe $u \in \mathcal{V}$ tal que

$$q(u) \neq 0.$$

De fato, como g é não-degenerada, existem vetores $v, w \in \mathcal{V}$ tais que

$$\langle v, w \rangle \neq 0.$$

Se $q(v) \neq 0$ ou $q(w) \neq 0$, a afirmação está provada. Caso contrário, se $q(v) = q(w) = 0$, escolhemos $u = v + w$, pois

$$\begin{aligned} q(v + w) &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &= q(u) + 2\langle v, w \rangle + q(w) \\ &= 2\langle v, w \rangle \\ &\neq 0. \end{aligned}$$

Basta então tomar

$$e_1 = \frac{u}{\|u\|}.$$

Assuma agora $n > 1$. Como no passo anterior, seja $u \in \mathcal{V}$ tal que

$$q(u) \neq 0.$$

Seja $U = \langle u \rangle$ o subespaço gerado por u e

$$U^\perp = \{v \in \mathcal{V} : \langle u, v \rangle = 0\}$$

o subespaço ortogonal a u . Afirmamos que

$$\mathcal{V} = U \oplus U^\perp.$$

De fato, os subespaços U e U^\perp são linearmente independentes, pois um vetor típico de U é da forma αu para algum escalar $\alpha \in \mathbb{R}$ e se $\alpha u \in U^\perp$ também então

$$\langle u, \alpha u \rangle = 0$$

por definição, mas

$$\langle u, \alpha u \rangle = \alpha \langle u, u \rangle = \alpha q(u)$$

e como $q(u) \neq 0$, segue que $\alpha = 0$. Além disso, $\mathcal{V} = U + U^\perp$: dado $v \in \mathcal{V}$, definindo

$$w = v - \frac{\langle u, v \rangle}{\langle u, u \rangle} u,$$

segue que $w \in U^\perp$, pois

$$\begin{aligned} \langle u, w \rangle &= g \left(u, v - \frac{\langle u, v \rangle}{\langle u, u \rangle} u \right) \\ &= \langle u, v \rangle - \frac{\langle u, v \rangle}{\langle u, u \rangle} \langle u, u \rangle \\ &= 0; \end{aligned}$$

portanto

$$v = \frac{\langle u, v \rangle}{\langle u, u \rangle} u + w$$

com o primeiro vetor da soma em U e o segundo em U^\perp . Agora, pela hipótese de indução existe uma base ortonormal

$$\mathcal{B}' = \{e_1, \dots, e_{n-1}\}$$

para U^\perp . Tomando $e_n = u$, obtemos uma base ortonormal

$$\mathcal{B} = \{e_1, \dots, e_{n-1}, e_n\}$$

para \mathcal{V} .

Para provar a última parte do teorema (segundo [Lang]), sejam

$$\begin{aligned} \mathcal{B} &= \{e_1, \dots, e_r, e_{r+1}, \dots, e_n\}, \\ \mathcal{B}' &= \{f_1, \dots, f_s, f_{s+1}, \dots, f_n\} \end{aligned}$$

duas bases ortonormais para \mathcal{V} ordenadas de tal forma que

$$\begin{aligned} q(e_i) &= \begin{cases} -1 & \text{se } i = 1, \dots, r, \\ 1 & \text{se } i = r+1, \dots, n, \end{cases} \\ q(f_i) &= \begin{cases} -1 & \text{se } i = 1, \dots, s, \\ 1 & \text{se } i = s+1, \dots, n. \end{cases} \end{aligned}$$

Afirmamos que

$$e_1, \dots, e_r, f_{s+1}, \dots, f_n$$

são linearmente independentes. De fato, se

$$\alpha_1 e_1 + \dots + \alpha_r e_r + \beta_{s+1} f_{s+1} + \dots + \beta_n f_n = 0,$$

escrevemos

$$\alpha_1 e_1 + \dots + \alpha_r e_r = -\beta_{s+1} f_{s+1} - \dots - \beta_n f_n,$$

e tomamos o produto interno desta equação consigo mesma, obtendo

$$-\alpha_1^2 - \dots - \alpha_r^2 = \beta_{s+1}^2 + \dots + \beta_n^2,$$

já que $g(e_i, e_i) = q(e_i) = -1$ para $i = 1, \dots, r$ e $g(f_i, f_i) = q(f_i) = 1$ para $i = s+1, \dots, n$. Como o lado esquerdo é não positivo e o lado direito é não negativo, segue que ambos devem ser do tipo luz e portanto

$$\alpha_1 = \dots = \alpha_r = \beta_{s+1} = \dots = \beta_n = 0,$$

provando a afirmação. Como $\dim \mathcal{V} = n$, temos que

$$r + (n - s) \leq n,$$

onde

$$r \leq s.$$

Por simetria do argumento, segue também que $s \leq r$ e portanto $r = s$. ■

Segue que dado um espaço vetorial \mathcal{V} com produto interno g e forma quadrática associada q , \mathcal{V} possui uma base ortonormal $\mathcal{B} = \{e_1, \dots, e_n\}$ tal que em relação a esta base o produto interno e a forma quadrática se escrevem na forma

$$\begin{aligned} \langle v, w \rangle &= \sum_{i=1}^k v^i w^i - \sum_{i=k+1}^n v^i w^i, \\ q(v) &= \sum_{i=1}^k (v^i)^2 - \sum_{i=k+1}^n (v^i)^2, \end{aligned}$$

para todos

$$v = v^i e_i, w = w^i e_i \in \mathcal{V}.$$

É óbvio que a mesma conclusão do Teorema de Sylvester vale para os vetores e_j tais que $q(e_j) = +1$. A razão do resultado ter sido enunciado em termos da contagem de vetores unitários negativos decorre da seguinte convenção:

2.9 Definição. Dado um espaço vetorial \mathcal{V} com produto interno g e sua forma quadrática associada q , o índice de g (e de q) é o número de vetores de uma base ortonormal de \mathcal{V} tais que $q(e_i) = -1$. □

2.10 Notação. Definimos

$$\eta_{ij}^k = \begin{cases} -1 & \text{se } i = j = 1, \dots, k, \\ 1 & \text{se } i = j = k+1, \dots, n, \\ 0 & \text{se } i \neq j. \end{cases}$$

O símbolo η_{ij}^k é uma generalização do delta de Kronecker δ_{ij} , isto é,

$$\eta_{ij}^0 = \delta_{ij}.$$

Ele é às vezes chamado o **eta de Kronecker**. Quando $k = 1$, omitimos o índice superescrito e escrevemos simplesmente η_{ij} , isto é,

$$\eta_{ij} = \eta_{ij}^1$$

Na representação matricial, temos

$$\eta^k = \begin{bmatrix} -\text{id}_k & 0 \\ 0 & \text{id}_{n-k} \end{bmatrix},$$

onde id_m é a matriz identidade $m \times m$. □

2.11 Exemplo. Em \mathbb{R}^n , se e_1, \dots, e_n são os vetores da base canônica, definimos uma métrica com índice k por $g_{ij}^k = \langle e_i, e_j \rangle = \eta_{ij}^k$, □

2.2 Operadores Ortogonais e Isometrias

2.12 Definição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$. Dizemos que um operador linear $L : \mathcal{V} \rightarrow \mathcal{V}$ é um **operador ortogonal** se ele *preserva* o produto interno, isto é, se

$$\langle Lv, Lw \rangle = \langle v, w \rangle$$

para todos $v, w \in \mathcal{V}$.

Uma **matriz ortogonal** é uma matriz que representa um operador ortogonal em relação a qualquer base. ■

Lembramos que se A é uma matriz que representa o operador linear L em relação à base $\mathcal{B} = \{e_1, \dots, e_n\}$ de \mathcal{V} , então a matriz A tem a forma em colunas

$$A = [\begin{array}{ccc} Le_1 & \dots & Le_n \end{array}],$$

onde cada coluna é

$$Le_j = \sum_{i=1}^n A_j^i e_i = \begin{bmatrix} A_j^1 \\ \vdots \\ A_j^n \end{bmatrix}.$$

Assim, A é ortogonal se

$$\langle Av, Aw \rangle = \langle v, w \rangle.$$

Se g é uma métrica, denotaremos também por g a matriz

$$g = (g_{ij}).$$

2.13 Proposição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$ e forma quadrática associada q de índice k . Então P é uma matriz ortogonal se e somente se

$$P^T \eta^k P = \eta^k.$$

Prova. Seja

$$\mathcal{B} = \{e_1, \dots, e_n\}$$

uma base ortonormal para \mathcal{V} satisfazendo

$$q(e_i) = \begin{cases} -1 & \text{se } i = 1, \dots, k, \\ 1 & \text{se } i = k+1, \dots, n. \end{cases}$$

Se

$$v = \sum_{i=1}^n v^i e_i,$$

denote

$$v^{\#k} = - \sum_{i=1}^k v^i e_i + \sum_{i=k+1}^n v^i e_i.$$

Note que em notação matricial (vetores como colunas)

$$v^{\#k} = \eta^k v$$

e

$$\langle v, w \rangle = v^T w^{\#k}$$

Seja P uma matriz ortogonal representando o operador ortogonal L em relação a \mathcal{B} . Então

$$\langle Le_i, Le_j \rangle = \langle e_i, e_j \rangle = \eta_{ij}^k$$

e

$$\begin{aligned} P\eta^k P &= \begin{bmatrix} Le_1 \\ \dots \\ Le_n \end{bmatrix} \begin{bmatrix} -\text{id}_k & 0 \\ 0 & \text{id}_{n-k} \end{bmatrix} \begin{bmatrix} Le_1 & \dots & Le_n \end{bmatrix} \\ &= \begin{bmatrix} Le_1 \\ \dots \\ Le_n \end{bmatrix} \begin{bmatrix} (Le_1)^{\#k} & \dots & (Le_n)^{\#k} \end{bmatrix} \\ &= [\langle Le_i, Le_j \rangle] \\ &= \eta^k. \end{aligned}$$

Reciprocamente, suponha que $P^T \eta^k P = \eta^k$. Então, se $v, w \in \mathcal{V}$ são vetores arbitrários, multiplicando esta equação à esquerda por v^T e à direita por w , temos

$$v^T (P^T \eta^k P) w = v^T \eta^k w. \quad (2.1)$$

Observando, como visto acima, que

$$v^T \eta^k w = v^T w^{\#k} = \langle v, w \rangle,$$

segue que o lado esquerdo de (2.1) é

$$\begin{aligned} v^T (P^T \eta^k P) w &= (Pv)^T \eta^k Pw \\ &= \langle Pv, Pw \rangle, \end{aligned}$$

enquanto que o seu lado direito é $\langle v, w \rangle$, donde

$$\langle Pv, Pw \rangle = \langle v, w \rangle.$$

■

2.14 Proposição. *O conjunto \mathcal{G} dos operadores ortogonais é um grupo sob a operação de composição de operadores.*

Prova. Claramente, o operador identidade está em \mathcal{G} . Se $L, M \in \mathcal{G}$, então

$$\begin{aligned} \langle (L \circ M) v, (L \circ M) w \rangle &= \langle L(Mv), L(Mw) \rangle \\ &= \langle Mv, Mw \rangle \\ &= \langle v, w \rangle, \end{aligned}$$

logo o produto $L \circ M \in \mathcal{G}$. Finalmente, se $L \in \mathcal{G}$, usaremos a não-degeneracidade do produto interno para mostrar que L é invertível e que $L^{-1} \in \mathcal{G}$. Se $v \in \ker L$, então para todo $w \in \mathcal{V}$ temos

$$\langle v, w \rangle = \langle Lv, Lw \rangle = \langle 0, Lw \rangle = 0,$$

e a não-degeneracidade do produto interno implica que $v = 0$, e portanto L é invertível. Daí,

$$\langle v, w \rangle = \langle L(L^{-1}v), L(L^{-1}w) \rangle = \langle L^{-1}v, L^{-1}w \rangle,$$

logo $L^{-1} \in \mathcal{G}$. ■

\mathcal{G} é chamado o **grupo ortogonal**.

2.15 Definição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno e forma quadrática associada q . Uma aplicação $F : \mathcal{V} \rightarrow \mathcal{V}$ tal que

$$q(F(v) - F(w)) = q(v - w)$$

para todos $v, w \in \mathcal{V}$ é chamada uma **isometria** de \mathcal{V} . ■

2.16 Proposição. O conjunto $\text{Isom}(\mathcal{V})$ das isometrias de um espaço vetorial com produto interno \mathcal{V} é um grupo sob a operação de composição de operadores.

Prova. Claramente, o operador identidade é uma isometria. Se $F, G \in \text{Isom}(\mathcal{V})$, então

$$\begin{aligned} q[(F \circ G)(v) - (F \circ G)(w)] &= q[F(G(v)) - F(G(w))] \\ &= q(G(v) - Gw) \\ &= q(v - w), \end{aligned}$$

logo o produto $F \circ G \in \text{Isom}(\mathcal{V})$.

Se $F \in \text{Isom}(\mathcal{V})$, para mostrar que F é invertível, seja T uma translação que leva $F(0)$ em 0. Então $T \circ F$ é uma isometria tal que $(T \circ F)(0) = 0$ e, pela Proposição 2.17 a seguir, é um operador ortogonal, em particular bijetiva; logo $F = T^{-1} \circ (T \circ F)$ é uma composição de bijeções, logo é uma bijeção e portanto invertível. Como

$$q(v - w) = q(F(F^{-1}(v)) - F(F^{-1}(w))) = q(F^{-1}(v) - F^{-1}(w)),$$

segue que $F^{-1} \in \text{Isom}(\mathcal{V})$. ■

Todo operador ortogonal é uma isometria, pois

$$\begin{aligned} q(L(v) - L(w)) &= q(L(v - w)) \\ &= \langle L(v - w), L(v - w) \rangle \\ &= \langle v - w, v - w \rangle \\ &= q(v - w), \end{aligned}$$

logo o grupo ortogonal é um subgrupo do grupo de isometrias.

2.17 Proposição. Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno. Isometrias $F : \mathcal{V} \rightarrow \mathcal{V}$ tais que $F(0) = 0$ são operadores ortogonais.

Prova: Passo 1. $q(F(v)) = q(v)$ para todo $v \in \mathcal{V}$.

Pois

$$q(F(v)) = q(F(v) - F(0)) = q(v - 0) = q(v).$$

Passo 2. F preserva o produto interno.

Pela identidade polar (Proposição 2.6), para todos $v, w \in \mathcal{V}$ vale

$$\begin{aligned} \langle F(v), F(w) \rangle &= \frac{1}{2} [q(F(v) - F(w)) - q(F(v)) - q(F(w))] \\ &= \frac{1}{2} (q(v - w) - q(v) - q(w)) \\ &= \langle v, w \rangle. \end{aligned}$$

Passo 3. F é linear.

Seja

$$\mathcal{B} = \{e_1, \dots, e_n\}$$

uma base ortonormal para \mathcal{V} . Suponha que

$$F(e_i) = e_i \quad (2.2)$$

para todo i . Para todo $v \in \mathcal{V}$ temos

$$\langle v, e_i \rangle = \langle F(v), F(e_i) \rangle = \langle F(v), e_i \rangle.$$

Por linearidade do produto interno segue que

$$\langle v, w \rangle = \langle F(v), w \rangle$$

para todo $w \in \mathcal{V}$, de modo que pela não-degeneracidade do produto interno temos

$$F(v) = v,$$

isto é,

$$F = \text{id},$$

que é um operador ortogonal. Se (2.2) não se cumpre, pelo Passo 2 F leva bases ortonormais em bases ortonormais, de modo que se

$$F(e_i) = f_i,$$

então

$$\mathcal{B}' = \{f_1, \dots, f_n\}$$

é uma base ortonormal. Se L é o operador ortogonal que leva \mathcal{B}' em \mathcal{B} , segue que $G = L \circ F$ é uma isometria que satisfaz

$$G(e_i) = e_i,$$

logo pelo argumento anterior segue que

$$G = \text{id}$$

e portanto

$$F = L^{-1}$$

é um operador ortogonal. ■

2.18 Proposição. *Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno. Se $F : \mathcal{V} \rightarrow \mathcal{V}$ é uma isometria, então existe um único operador ortogonal L e uma única translação T tais que*

$$F = T \circ L.$$

Prova: (*Existência*) Seja $T : \mathcal{V} \rightarrow \mathcal{V}$ a translação

$$T(p) = p + F(0).$$

Então $T^{-1} \circ F : \mathcal{V} \rightarrow \mathcal{V}$ é uma isometria tal que

$$(T^{-1} \circ F)(0) = T^{-1}(F(0)) = F(0) - F(0) = 0,$$

de modo que pelo resultado anterior

$$L := T^{-1} \circ F$$

é um operador ortogonal.

(*Unicidade*) Se L_1, L_2 são operadores ortogonais e T_1, T_2 são translações tais que

$$F = T_1 \circ L_1 = T_2 \circ L_2,$$

então

$$L_1 \circ L_2^{-1} = T_1^{-1} \circ T_2.$$

Em particular, $T_1^{-1} \circ T_2$ é um operador linear e como operadores lineares deixam o vetor nulo fixo, segue que $T_1^{-1} \circ T_2$ é a translação nula, isto é, a identidade. Logo,

$$\begin{aligned} T_1^{-1} \circ T_2 &= \text{id} \implies T_1 = T_2, \\ L_1 \circ L_2^{-1} &= \text{id} \implies L_1 = L_2. \end{aligned}$$

■

2.3 Métrica de Lorentz e Espaçotempo de Minkowski

2.19 Definição. Uma **métrica de Lorentz** em um espaço vetorial de dimensão 4 é um produto interno de índice 1.

Um **espaçotempo de Minkowski** é \mathbb{R}^4 munido de uma métrica de Lorentz. Ele será denotado por \mathbb{M} . Os pontos de \mathbb{M} serão chamados **eventos**. □

Em particular, segue da observação logo após o Teorema de Sylvester, Teorema 2.8, que todo espaço-tempo de Minkowski \mathbb{M} possui uma base $\mathcal{B} = \{e_0, e_1, e_2, e_3\}$ com a propriedade de que seu produto interno $\langle \cdot, \cdot \rangle$ e sua forma quadrática associada q em relação a esta base são dados por

$$\begin{aligned} \langle v, w \rangle &= -v^0 w^0 + v^1 w^1 + v^2 w^2 + v^3 w^3, \\ q(v) &= -(v^0)^2 + (v^1)^2 + (v^2)^2 + (v^3)^2 \end{aligned}$$

para todos $v = v^i e_i, w = w^i e_i \in \mathbb{M}$. Usando notação matricial, se

$$\eta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

a métrica de Lorentz pode ser escrita na forma

$$\langle v, w \rangle = \eta_{ij} v^i w^j.$$

Segue do Princípio da Relatividade (Teorema 1.13 da Invariância do Intervalo) que o espaço-tempo físico na Teoria da Relatividade Especial é um espaço-tempo de Minkowski, com o intervalo entre dois eventos v, w sendo exatamente $q(v - w)$. Mais tarde, quando estudarmos a Teoria da Relatividade Geral, veremos que isso é verdade apenas localmente.

Vale a pena ressaltar que foi Minkowski quem traduziu a Teoria da Relatividade Especial para a linguagem do *espaçotempo* (em 1907), isto é, foi ele quem reconheceu que a consequência principal da teoria foi a unificação do espaço e do tempo em uma só unidade. Em suas palavras: “*De agora em diante, espaço por si só e tempo por si estão condenados a desaparecer em meras sombras, e somente uma união dos dois preservará uma realidade independente.*” ([Taylor-Wheeler], Seção 1.5, p. 15.) Inicialmente, Einstein não aceitou bem a idéia de Minkowski. Em sua opinião, estender o espaço para a quarta dimensão mais mistificava do que esclarecia a teoria. Na verdade, em resposta à formulação de Minkowski, Einstein é quotado como tendo dito: “*Desde que os matemáticos invadiram a teoria da relatividade, eu próprio não a comprehendo mais.*” Pouco tempo depois, no entanto, Einstein mudou radicalmente de opinião e reconheceu a importância do conceito de espaço-tempo, que foi fundamental para a elaboração da Teoria da Relatividade Geral; Minkowski foi a primeira entre as duas únicas pessoas que ele reconheceu no artigo de publicação da teoria como sendo instrumentais na sua formulação (veja Seção 7.1).

Na realidade, o *espaçotempo* é muito mais que uma linguagem ou uma formalização diferente da teoria. Ele dá sentido físico à Teoria da Relatividade, como [Petkov] argumenta muito persuasivamente no Capítulo 5 (especialmente pp. 75–76). De fato, o próprio Minkowski (quotado lá, p. 78) argumentou convincentemente que o Postulado da Relatividade deveria ser trocado pelo *Postulado do Mundo Absoluto*. Ou seja, a principal idéia da Teoria da Relatividade Especial, que resume tudo e de onde todas as consequências seguem, é trocar o espaço e o tempo absoluto de Newton pelo **espaçotempo absoluto** com a métrica de Lorentz (o espaço-tempo de Minkowski). [Petkov], Capítulo 6, argumenta que a relutância inicial de Einstein em aceitar o espaço-tempo de Minkowski desempenhou um papel importante na relativa demora em desenvolver a Teoria da Relatividade Geral, mesmo após ele ter tido o insight do princípio da equivalência (seu pensamento mais feliz da vida; veja Seção 7.1).

2.20 Definição. O grupo das isometrias do espaço-tempo de Minkowski $\text{Isom}(\mathbb{M})$ é chamado o **grupo de Poincaré**.

Um operador ortogonal do espaço-tempo de Minkowski é chamado uma **transformação de Lorentz**. O grupo das transformações de Lorentz é chamado o **grupo de Lorentz**. ■

Assim, transformações de Lorentz são operadores lineares no espaço-tempo de Minkowski que preservam a métrica de Lorentz. Consequentemente, o grupo de Lorentz é um subgrupo do grupo de Poincaré.

Pode-se provar que o grupo de Poincaré é um grupo a 10 parâmetros. Um subgrupo particularmente fisicamente interessante é o **subgrupo de Poincaré ortocrono próprio**, obtido quando se exclui as reflexões espaciais e temporais (estas últimas são chamadas *reversões temporais*; elas se referem a transformações de coordenadas entre dois referenciais inerciais em que o relógio de um está andando na direção contrária à do relógio do outro), que ainda é um grupo a 10 parâmetros gerado pelas 4 translações (nas direções t, x, y, z) e por 6 transformações de Lorentz, 3 das quais são as rotações espaciais da geometria euclidiana em relação aos eixos x, y, z e as outras 3 são os boosts nas direções x, y, z , a serem explicados na próxima seção. Para maiores detalhes sobre o subgrupo de Poincaré ortocrono próprio, veja [Rowe], Seção 1.9, pp. 31–34 e [Naber], Seção 1.3, pp. 15–42. Para maiores detalhes sobre o grupo de Poincaré geral, veja [Tsamparlis], Seção 1.6, pp. 18–26, para uma dedução com insight geométrico, e Seção 1.7, pp. 26–40, para uma dedução puramente algébrica.

2.3.1 Propriedades do Espaçotempo de Minkowski

2.21 Definição. Vetores $v \in \mathbb{M}$ tais que

- $q(v) = 0$ são chamados do tipo **luz**,
- $q(v) > 0$ são chamados do tipo **espaço**,
- $q(v) < 0$ são chamados do tipo **tempo**.

□

Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno. Dizemos que um subespaço $W \subset \mathcal{V}$ é **não degenerado** se o produto interno de \mathcal{V} restrito a W é não degenerado; caso contrário, dizemos que W é **degenerado**. Por exemplo, no espaço de Minkowski \mathbb{M} , o subespaço gerado por e_1, e_2, e_3 é não degenerado, enquanto que o subespaço gerado por e_0 é degenerado.

2.22 Lema. *Seja \mathcal{V} um espaço vetorial com produto interno. Um subespaço $W \subset \mathcal{V}$ é não degenerado se e somente se*

$$\mathcal{V} = W \oplus W^\perp.$$

Prova: Lembremos o seguinte resultado de Álgebra Linear sobre dimensões de subespaços:

$$\dim(W + W^\perp) + \dim(W \cap W^\perp) = \dim W + \dim W^\perp. \quad (2.3)$$

Afirmamos que para qualquer subespaço $W \subset \mathcal{V}$ vale

$$\dim W + \dim W^\perp = n. \quad (2.4)$$

De fato, seja $\{e_1, \dots, e_m\}$ uma base para W e complete esta base até uma base $\{e_1, \dots, e_m, e_{m+1}, \dots, e_n\}$ para \mathcal{V} . Um vetor $v \in \mathcal{V}$ pertence a W^\perp se e somente se $\langle v, e_i \rangle = 0$ para $i = 1, \dots, m$, ou seja, se e somente se $v = v^j e_j$ satisfaz o sistema

$$g_{ij}v^j = 0 \quad \text{para } i = 1, \dots, m.$$

Como a matriz (g_{ij}) é invertível, o subespaço solução deste sistema tem dimensão exatamente igual a $n - m$.

De (2.3) e (2.15) segue que $\mathcal{V} = W + W^\perp$ se e somente se $W \cap W^\perp = 0$. Assim, qualquer uma destas condições é equivalente a $\mathcal{V} = W \oplus W^\perp$. Mas

$$W \cap W^\perp = \{w \in W : w \perp W\},$$

logo $W \cap W^\perp = 0$ se e somente se W é não degenerado. ■

2.23 Proposição. *Se $v \in \mathbb{M}$ é do tipo tempo e $w \neq 0$ é ortogonal a v , então w é um vetor do tipo espaço.*

Em particular, $\langle v \rangle^\perp$ é um subespaço vetorial formado apenas por vetores do tipo espaço, e portanto onde o produto interno é positivo definido.

Prova: Pelo lema, podemos decompor

$$\mathbb{M} = \langle v \rangle \oplus \langle v \rangle^\perp.$$

Como os índices de \mathbb{M} e $\langle v \rangle$ são iguais a 1, pelo Teorema de Sylvester o índice de $\langle v \rangle^\perp$ deve ser zero, caso contrário obteríamos uma base para \mathbb{M} com mais de um vetor com valor q igual a -1 . Portanto, todo vetor ortogonal a v não nulo é do tipo espaço. ■

2.24 Definição. Denote por \mathbb{M}^- o conjunto dos vetores do tipo tempo. Para $v \in \mathbb{M}^-$, definimos o **cone temporal** de v como sendo o conjunto

$$\mathcal{C}(v) = \{w \in \mathbb{M}^- : \langle v, w \rangle < 0\}.$$

Seu **cone temporal oposto** é o conjunto

$$\mathcal{C}(-v) = -\mathcal{C}(v) = \{w \in \mathbb{M}^- : \langle v, w \rangle > 0\}.$$

□

2.25 Proposição. *Dois vetores v, w do tipo tempo estão no mesmo cone temporal se e somente se*

$$\langle v, w \rangle < 0.$$

Prova: Seja $v \in \mathcal{C}(u)$, para $u \in \mathbb{M}^-$ que podemos tomar unitário, de modo que $\langle u, u \rangle = -1$. Mostraremos que $w \in \mathbb{M}^-$ pertence também a $\mathcal{C}(u)$ se e somente se $\langle v, w \rangle < 0$. Escreva

$$\begin{aligned} v &= au + x, \\ w &= bu + y, \end{aligned}$$

para $x, y \in \langle u \rangle^\perp$. Temos

$$\langle v, w \rangle = \langle au + x, bu + y \rangle = -ab + \langle x, y \rangle. \quad (2.5)$$

Como

$$\begin{aligned} 0 > \langle v, v \rangle &= \langle au + x, au + x \rangle = -a^2 + \|x\|^2, \\ 0 > \langle w, w \rangle &= \langle bu + y, bu + y \rangle = -b^2 + \|y\|^2, \end{aligned}$$

segue que

$$\begin{aligned}|a| &> \|x\|, \\ |b| &> \|y\|,\end{aligned}$$

e pela desigualdade de Cauchy-Schwartz, que pela Proposição 2.23 é valida em $\langle u \rangle^\perp$, obtemos

$$|\langle x, y \rangle| \leq \|x\| \|y\| < |ab|.$$

Segue de (2.5) que

$$\operatorname{sign} \langle v, w \rangle = -\operatorname{sign}(ab).$$

Como $v \in \mathcal{C}(u)$,

$$0 > \langle u, v \rangle = \langle u, au + x \rangle = -a,$$

isto é,

$$a > 0,$$

onde

$$\operatorname{sign} \langle v, w \rangle = -\operatorname{sign} b,$$

Mas $w \in \mathcal{C}(u)$ se e somente se $b > 0$, pois

$$\langle u, w \rangle = \langle u, bu + y \rangle = -b,$$

logo segue o resultado. ■

Existem apenas dois cones temporais, pois pela Proposição 2.23 o produto escalar de dois vetores do tipo tempo não pode ser nulo, logo ou é negativo e eles estão no mesmo cone temporal pela Proposição 2.25, ou é positivo e eles estão em cones temporais opostos. Portanto, temos a união disjunta

$$\mathbb{M}^- = \mathcal{C}(v) \cup \mathcal{C}(-v)$$

para qualquer vetor do tipo tempo v .

Da mesma forma que o conjunto dos vetores do tipo tempo \mathbb{M}^- , cones temporais não são subespaços vetoriais de \mathbb{M} . Mas cada cone temporal é um conjunto convexo: se $a, b > 0$ e $v, w \in \mathcal{C}$, então $av + bw \in \mathcal{C}$.

2.26 Definição. Fixada uma base ortonormal

$$\mathcal{B} = \{e_0, e_1, e_2, e_3\}$$

para o espaço de Minkowski \mathbb{M} com $q(e_0) = -1$, o **cone temporal futuro** de \mathbb{M} é o conjunto

$$\mathcal{C}^+ = \mathcal{C}(e_0) = \{v \in \mathbb{M}^- : \langle v, e_0 \rangle < 0\},$$

enquanto que o **cone temporal passado** de \mathbb{M} é o conjunto

$$\mathcal{C}^- = \mathcal{C}(-e_0) = \{v \in \mathbb{M}^- : \langle v, e_0 \rangle > 0\}.$$

Assim,

$$\mathbb{M}^- = \mathcal{C}^+ \cup \mathcal{C}^-.$$

A escolha de qual cone temporal é designado como o cone temporal futuro e qual cone temporal é designado como o cone temporal passado depende portanto da escolha da base. Isso determina uma orientação no espaço de Minkowski.

Como observado na demonstração da Proposição 2.25, existem vetores do tipo espaço que satisfazem a desigualdade de Cauchy-Schwartz. Isso não é verdade para todos os vetores do tipo espaço. Por exemplo,

$$\begin{aligned} v &= (-1, 1, 1, 1), \\ w &= (1, 1, 1, 1), \end{aligned}$$

são vetores do tipo espaço, pois

$$\langle v, v \rangle = \langle w, w \rangle = 2,$$

mas

$$\langle v, w \rangle = 4,$$

enquanto que

$$\|v\| \|w\| = \sqrt{2}\sqrt{2} = 2,$$

portanto

$$\langle v, w \rangle > \|v\| \|w\|.$$

Por outro lado, vetores do tipo tempo *sempre* satisfazem a *desigualdade de Cauchy-Schwartz reversa*:

2.27 Proposição (Desigualdade de Cauchy-Schwartz Reversa). *Se $v, w \in \mathbb{M}$ são do tipo tempo, então*

$$|\langle v, w \rangle| \geq \|v\| \|w\|$$

com a igualdade valendo se e somente se v e w são múltiplos escalares um do outro.

Prova: Escreva

$$w = av + z$$

com $z \in \langle v \rangle^\perp$. Então

$$\langle w, w \rangle = a^2 \langle v, v \rangle + \langle z, z \rangle,$$

e daí

$$\begin{aligned} \langle v, w \rangle^2 &= a^2 \langle v, v \rangle^2 = (a^2 \langle v, v \rangle) \langle v, v \rangle \\ &= (\langle w, w \rangle - \langle z, z \rangle) \langle v, v \rangle \\ &\geq \langle w, w \rangle \langle v, v \rangle \\ &= \|v\| \|w\|, \end{aligned}$$

pois $\langle z, z \rangle \geq 0$ pela Proposição 2.23 e $\langle w, w \rangle, \langle v, v \rangle < 0$. A igualdade vale se e somente se $z = 0$. ■

2.28 Proposição (Desigualdade Triangular Reversa). *Se $v, w \in \mathbb{M}$ são do tipo tempo pertencentes ao mesmo cone temporal, então*

$$\|v\| + \|w\| \leq \|v + w\|$$

com a igualdade valendo se e somente se v e w são múltiplos escalares um do outro.

Prova: Como v, w estão no mesmo cone temporal e cones temporais são convexos, segue que $v + w$ também está no mesmo cone temporal e em particular também é um vetor do tipo tempo, ou seja,

$$\langle v + w, v + w \rangle < 0.$$

Pela desigualdade de Cauchy-Schwartz reversa,

$$\|v\| \|w\| \leq -\langle v, w \rangle.$$

Logo,

$$\begin{aligned}
 (\|v\| + \|w\|)^2 &= \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
 &\leq \|v\|^2 - 2\langle v, w \rangle + \|w\|^2 \\
 &= -\langle v, v \rangle - 2\langle v, w \rangle - \langle w, w \rangle \\
 &= -\langle v + w, v + w \rangle \\
 &= \|v + w\|^2
 \end{aligned}$$

A igualdade só valerá se e somente se a igualdade valer na desigualdade de Cauchy-Schwartz reversa. ■
Assim, a reta que liga dois pontos no espaço de Minkowski não é o caminho mais curto entre dois pontos;
na verdade, é o caminho *mais longo*.

2.29 Proposição (Interpretação Geométrica do Produto Interno). *Se $v, w \in \mathbb{M}$ são vetores do tipo tempo pertencentes ao mesmo cone temporal, então existe um único real $\theta \geq 0$ tal que*

$$\langle v, w \rangle = -\|v\|\|w\|\cosh\theta.$$

θ é chamado o **ângulo hiperbólico** entre v e w .

Prova: Pela desigualdade de Cauchy-Schwartz reversa,

$$\frac{|\langle v, w \rangle|}{\|v\|\|w\|} \geq 1,$$

logo existe um único número não negativo θ tal que

$$\cosh\theta = \frac{|\langle v, w \rangle|}{\|v\|\|w\|}.$$

Pela Proposição 2.25, $\langle v, w \rangle < 0$, logo

$$|\langle v, w \rangle| = -\langle v, w \rangle,$$

onde

$$\cosh\theta = -\frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

■

2.30 Definição. Qualquer vetor do tipo tempo ou do tipo luz é chamado um vetor do tipo **causal**.

O fecho de um cone temporal é chamado um **cone causal**. □

Qualquer vetor que não é do tipo espaço é um vetor do tipo causal. A fronteira de um cone temporal consiste exatamente de vetores do tipo luz. A interseção dos dois cones causais é exatamente o vetor nulo. O motivo da terminologia “causal” será vista no próximo capítulo.

2.4 Transformação de Lorentz do tipo Boost

Nesta seção vamos obter a transformação de Lorentz entre dois sistemas de coordenadas para o espaço-tempo de Minkowski cujas origens coincidem, e tais que, *espacialmente*, eles tem eixos paralelos e um se move em relação ao outro ao longo de um dos eixos x, y ou z . Este tipo de transformação de Lorentz é chamada um **boost** (impulso) na direção x, y ou z .

Vamos obter o boost na direção x . Considere dois referenciais iniciais em movimento relativo. Digamos que temos um referencial inercial O , cujas coordenadas denotaremos por (t, x, y, z) , e um referencial inercial O' , cujas coordenadas denotaremos por (t', x', y', z') , que se move com velocidade v , $|v| < 1$, na direção

positiva do eixo x do referencial inercial O , com os eixos y, z, y', z' transversais à direção do movimento relativo e a origem dos dois referenciais coincidindo no tempo $t = t' = 0$. Portanto,

$$\begin{aligned} y &= y', \\ z &= z'. \end{aligned}$$

2.4.1 Medida Quantitativa da Dilatação do Tempo

Para fixar idéias, começemos com dois eventos A e B , ambos consistindo na emissão de uma faísca, o primeiro evento ocorrendo na origem comum dos dois sistemas de coordenadas, o segundo ocorrendo algum tempo depois na origem espacial do sistema O' . As coordenadas destes eventos (ignorando as coordenadas y, z, y', z') nestes dois referenciais iniciais são portanto

$$\begin{aligned} (x_A, t_A) &= (0, 0), \\ (x_B, t_B) &= (x, t). \end{aligned}$$

e

$$\begin{aligned} (x'_A, t'_A) &= (0, 0), \\ (x'_B, t'_B) &= (0, t'). \end{aligned}$$

de modo que

$$\begin{aligned} \Delta x &= x, \\ \Delta t &= t, \\ \Delta x' &= 0, \\ \Delta t' &= t'. \end{aligned}$$

Temos

$$x = vt$$

e pela invariância do intervalo spaçotemporal

$$(t')^2 = t^2 - x^2 = (1 - v^2) t^2,$$

de modo que

$$t = \frac{1}{\sqrt{1 - v^2}} t'.$$

2.31 Definição. O número

$$\gamma = \frac{1}{\sqrt{1 - v^2}} \quad (2.6)$$

é chamado o **fator de dilatação do tempo**. \square

Observe que t' é o tempo próprio no referencial O' e como

$$\gamma \geq 1,$$

isso mostra o que já vimos antes, que o tempo próprio é sempre menor que o tempo medido em um outro referencial que está se movendo em relação ao referencial onde o tempo próprio é medido. Agora temos uma expressão que nos dá exatamente o quanto o tempo é dilatado.

Concluímos que para dois eventos cujas coordenadas $x' = 0$ a transformação de Lorentz é

$$t = \gamma t', \quad (2.7)$$

$$x = v\gamma t'. \quad (2.8)$$

Observe que

$$v = \pm \frac{\sqrt{\gamma^2 - 1}}{\gamma},$$

pois a velocidade relativa entre os referenciais pode ser positiva ou negativa, dependendo da direção em que um se move em relação ao outro, portanto opta-se por escrever v explicitamente, e não em função do parâmetro γ . Se $v > 0$ (isto é, o referencial O' se move para a direita do referencial O), então

$$\begin{aligned} t &= \gamma t', \\ x &= \sqrt{\gamma^2 - 1} t', \end{aligned}$$

enquanto que se $v < 0$ (isto é, o referencial O' se move para a esquerda do referencial O), então

$$\begin{aligned} t &= \gamma t', \\ x &= -\sqrt{\gamma^2 - 1} t', \end{aligned}$$

2.4.2 Matriz do Boost

No caso geral,

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix},$$

ou seja,

$$\begin{aligned} t &= a_{11}t' + a_{12}x', \\ x &= a_{21}t' + a_{22}x'. \end{aligned}$$

Fazendo $x' = 0$, caímos no caso anterior e obtemos

$$\begin{aligned} a_{11} &= \gamma, \\ a_{21} &= v\gamma. \end{aligned}$$

Portanto, simplificando a notação,

$$\begin{aligned} t &= \gamma t' + ax', \\ x &= v\gamma t' + bx'. \end{aligned}$$

Para obter as constantes a e b usamos novamente a invariância do intervalo espaço-temporal. Escolhendo agora eventos mais gerais

$$\begin{aligned} (x_A, t_A) &= (0, 0), \\ (x_B, t_B) &= (x, t). \end{aligned}$$

e

$$\begin{aligned} (x'_A, t'_A) &= (0, 0), \\ (x'_B, t'_B) &= (x', t'). \end{aligned}$$

de modo que

$$\begin{aligned} \Delta x &= x, \\ \Delta t &= t, \\ \Delta x' &= x', \\ \Delta t' &= t', \end{aligned}$$

segue que

$$(t')^2 - (x')^2 = t^2 - x^2,$$

Substituindo, temos

$$\begin{aligned} (t')^2 - (x')^2 &= (\gamma t' + ax')^2 - (v\gamma t' + bx')^2 \\ &= (\gamma^2 - v^2\gamma^2)(t')^2 + (2\gamma a - 2v\gamma b)t'x' \\ &\quad + (a^2 - b^2)(x')^2 \\ &= (t')^2 + 2\gamma(a - vb)t'x' + (a^2 - b^2)(x')^2 \end{aligned}$$

Como t', x' são arbitrários, obtemos

$$\begin{aligned} a - vb &= 0, \\ a^2 - b^2 &= -1. \end{aligned}$$

Da primeira equação segue que

$$a = vb,$$

que substituímos na segunda equação para obter

$$(v^2 - 1)b^2 = -1$$

donde

$$b = \frac{1}{\sqrt{1-v^2}} = \gamma.$$

Logo,

$$a = v\gamma.$$

Assim,

$$t = \gamma t' + v\gamma x', \tag{2.9}$$

$$x = v\gamma t' + \gamma x'. \tag{2.10}$$

e

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix}. \tag{2.11}$$

Observe que a matriz da transformação de Lorentz boost é simétrica. Em termos apenas de v podemos escrever

$$t = \frac{t' + vx'}{\sqrt{1-v^2}}, \tag{2.12}$$

$$x = \frac{x' + vt'}{\sqrt{1-v^2}}. \tag{2.13}$$

A inversa da transformação de Lorentz é dada por

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}. \tag{2.14}$$

Ela é facilmente obtida uma vez que notamos que o referencial O se move com velocidade $-v$ na direção positiva do eixo x do referencial inercial O' , logo basta substituir v por $-v$ em (2.11).

As matrizes das transformações de Lorentz boosts nas direções y e z são, respectivamente,

$$\begin{bmatrix} \gamma & 0 & v\gamma & 0 \\ 0 & 1 & 0 & 0 \\ v\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} \gamma & 0 & 0 & v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ v\gamma & 0 & 0 & \gamma \end{bmatrix}.$$

A composta de boosts em direções diferentes não é um boost e portanto o conjuntos dos boosts não forma um subgrupo do grupo de Poincaré (é uma rotação, chamada **rotação de Thomas** ou **precessão de Thomas**; veja [Rowe], Seção 1.9.3, pp. 33–34 e [Rahaman], Seção 3.4, p. 31). Este fato só foi descoberto nos anos 1920s e diz-se que surpreendeu Einstein. É claro que se considerarmos apenas boosts em uma única direção eles formam um subgrupo a 1 parâmetro, isomorfo a \mathbb{R} , como veremos na Seção 2.4.5.

2.4.3 Medida Quantitativa da Contração do Espaço

Usando o fator de dilatação do tempo podemos também encontrar o fator de contração do espaço, isto é, a contração medida do comprimento de um objeto ao longo de sua direção de movimento em relação à medida do comprimento deste objeto em relação a um referencial em que ele se encontra em repouso.

Imagine que o referencial móvel O' é uma nave espacial viajando da Terra para um sistema estelar a uma distância L quando medida no referencial O , que é exatamente a Terra; em relação ao referencial Terra, é como se tivéssemos uma barra enorme em repouso ligando a Terra ao sistema estelar de comprimento L . O referencial Nave deve medir uma distância menor L' , conforme vimos. De acordo com o referencial Terra O , a viagem leva o tempo

$$t = \frac{L}{v},$$

enquanto que no referencial Nave Espacial O' , o tempo próprio é

$$t' = \frac{t}{\gamma} = \frac{L}{v\gamma} = \frac{\sqrt{1-v^2}}{v} L.$$

A medida da distância L' entre a Terra e o sistema estelar de acordo com o referencial inercial Nave O' é dada por

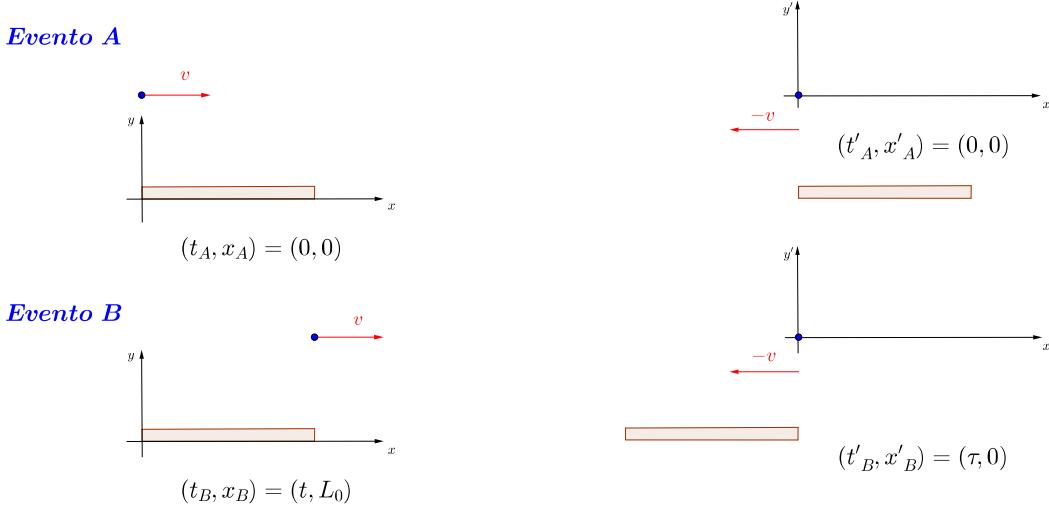
$$L' = vt',$$

pois v é a velocidade da nave quando ela deixou a primeira extremidade desta longuíssima barra (que é o planeta Terra) e t' é o tempo que ela levou para chegar à segunda extremidade da barra (o sistema estelar). Portanto

$$L' = \sqrt{1-v^2} L. \tag{2.15}$$

2.4.4 Resumo da Contração do Espaço e Dilatação do Tempo

Segundo [Ferraro], resumimos agora a contração do espaço e dilatação do tempo em uma única situação. Considere uma partícula que se move ao longo de uma barra de comprimento L_0 com velocidade v .



O evento A é a partícula passando pela extremidade esquerda da barra, enquanto que o evento B é a partícula passando pela extremidade direita da barra. No referencial (t, x) da barra (isto é, no referencial em que a barra encontra-se em repouso), os eventos A e B tem coordenadas

$$(t_A, x_A) = (0, 0), \\ (t_B, x_B) = (t, L_0),$$

de modo que o comprimento próprio L_0 da barra é

$$L_0 = vt,$$

enquanto que no referencial (t', x') da partícula (isto é, no referencial em que a partícula está em repouso) os eventos A e B tem coordenadas

$$(t'_A, x'_A) = (0, 0), \\ (t'_B, x'_B) = (\tau, 0),$$

de modo que o comprimento L da barra medido neste referencial é

$$L = v\tau.$$

Em particular,

$$\frac{t}{\tau} = \frac{L_0}{L},$$

o que mostra que não é possível que os comprimentos (distâncias espaciais) sejam relativos, isto é, variem com a velocidade do referencial, sem que os tempos (intervalos temporais) também sejam e vice-versa.

Pela invariância do intervalo,

$$\tau^2 = t^2 - L_0^2.$$

Substituindo $L_0 = vt$, obtemos

$$\tau = \sqrt{1 - v^2}t,$$

de modo que o tempo próprio, medido no referencial em que os eventos ocorrem na mesma posição, é sempre menor que o tempo medido em um referencial em que os eventos ocorrem em diferentes posições. Dividindo ambos os lados por τ e substituindo t/τ por L_0/L , obtemos

$$L = \sqrt{1 - v^2} L_0,$$

de modo que o comprimento próprio, isto é, o comprimento de um objeto em um referencial em que ele está em repouso, é sempre maior que o comprimento do objeto medido em um referencial em relação ao qual ele se move (mover-se em relação a um objeto “causa” a contração do mesmo na dimensão ao longo da direção do movimento).

2.4.5 Lei de Adição de Velocidades

Considere três referenciais inerciais em movimento relativo. Temos como antes o referencial inercial $O(t, x, y, z)$ e o referencial inercial $O'(t', x', y', z')$ que se move com velocidade v' na direção positiva do eixo x do referencial inercial O . Temos agora um terceiro referencial inercial $O''(t'', x'', y'', z'')$ que se move com velocidade v'' na direção positiva do eixo x' do referencial inercial O' . Assumimos os eixos y, z, y', z', y'', z'' transversais à direção do movimento relativo e a origem dos três referenciais coincidindo no tempo $t = t' = t'' = 0$. Qual é a velocidade v do referencial inercial $O''(t'', x'', y'', z'')$ em relação ao referencial inercial O ?

Para obter a resposta, basta usar a propriedade de grupo das transformações de Lorentz. A transformação de Lorentz das coordenadas do referencial O'' para as coordenadas do referencial O é o produto da transformação de Lorentz das coordenadas do referencial O'' para as coordenadas do referencial O' pela transformação de Lorentz das coordenadas do referencial O' para as coordenadas do referencial O . Assim, temos ao mesmo tempo

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t'' \\ x'' \\ y'' \\ z'' \end{bmatrix}$$

e

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \gamma' & v'\gamma' & 0 & 0 \\ v'\gamma' & \gamma' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma'' & v''\gamma'' & 0 & 0 \\ v''\gamma'' & \gamma'' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t'' \\ x'' \\ y'' \\ z'' \end{bmatrix},$$

isto é,

$$\begin{aligned} \begin{bmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} \gamma' & v'\gamma' & 0 & 0 \\ v'\gamma' & \gamma' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma'' & v''\gamma'' & 0 & 0 \\ v''\gamma'' & \gamma'' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \gamma'\gamma''(1 + v'v'') & \gamma'\gamma''(v' + v'') & 0 & 0 \\ \gamma'\gamma''(v' + v'') & \gamma'\gamma''(1 + v'v'') & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Portanto,

$$\begin{aligned} \gamma &= \gamma'\gamma''(1 + v'v''), \\ v\gamma &= \gamma'\gamma''(v' + v''). \end{aligned}$$

Logo, a lei de adição de velocidades na teoria da relatividade é

$$v = \frac{v' + v''}{1 + v'v''},$$

o que é bastante diferente da lei de adição de velocidades na mecânica Newtoniana (lá, $v = v' + v''$). Em particular, se os referenciais viajam com velocidade menor ou igual à velocidade da luz, um em relação ao outro, a velocidade do terceiro referencial em relação ao primeiro também é menor que a velocidade da luz, como provamos a seguir.

2.32 Proposição. *Se*

$$|v'|, |v''| \leq 1$$

e

$$v = \frac{v' + v''}{1 + v'v''},$$

então

$$|v| \leq 1.$$

Prova:

Demonstração 1 (mais elementar). Se $|a| < 1$, a função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$f(x) = \frac{x+a}{1+ax}$$

é crescente, pois

$$f'(x) = \frac{1-a^2}{(1+ax)^2} > 0.$$

Como $f(-1) = -1$ e $f(1) = 1$, segue que $f((-1, 1)) = (-1, 1)$, donde o resultado.

Demonstração 2 (menos elementar, mas oferece um importante insight). A função tangente hiperbólica, definida da maneira esperada por

$$\tanh x = \frac{\sinh x}{\cosh x}, \quad (2.16)$$

é uma função estritamente crescente, pois sua derivada é

$$\begin{aligned} \frac{d}{dx} \tanh x &= \frac{\cosh x \cosh x - \sinh x \sinh x}{\cosh^2 x} = \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} \\ &= \frac{1}{\cosh^2 x} > 0, \end{aligned}$$

ímpar (pois é o produto de uma função par e uma função ímpar) e satisfaz

$$|\tanh x| < 1 \quad \text{para todo } x \in \mathbb{R} \quad (2.17)$$

pois

$$|\sinh x| < \cosh x$$

para todo $x \in \mathbb{R}$, já que sempre temos

$$\frac{|e^x - e^{-x}|}{2} < \frac{e^x + e^{-x}}{2}$$

e

$$\tanh(\mathbb{R}) = (-1, 1), \quad (2.18)$$

porque, pela regra de L'Hôpital,

$$\begin{aligned} \lim_{x \rightarrow +\infty} \tanh x &= \lim_{x \rightarrow +\infty} \frac{\sinh x}{\cosh x} = \lim_{x \rightarrow +\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ &= \lim_{x \rightarrow +\infty} \frac{e^{2x} - 1}{e^{2x} + 1} = \lim_{x \rightarrow +\infty} \frac{2e^{2x}}{2e^{2x}} = 1. \end{aligned}$$

Vale a identidade

$$\tanh(a+b) = \frac{\tanh a + \tanh b}{1 + \tanh a \tanh b}. \quad (2.19)$$

De fato, temos

$$\begin{aligned} \operatorname{senh}(a+b) &= \operatorname{senh} a \cosh b + \cosh a \operatorname{senh} b, \\ \cosh(a+b) &= \cosh a \cosh b + \operatorname{senh} a \operatorname{senh} b. \end{aligned}$$

Isso é facilmente verificado:

$$\begin{aligned} \operatorname{senh} a \cosh b + \cosh a \operatorname{senh} b &= \frac{e^a - e^{-a}}{2} \frac{e^b + e^{-b}}{2} + \frac{e^a + e^{-a}}{2} \frac{e^b - e^{-b}}{2} \\ &= \frac{e^{a+b} + e^{a-b} - e^{-a+b} - e^{-a-b}}{4} + \frac{e^{a+b} - e^{a-b} + e^{-a+b} - e^{-a-b}}{4} \\ &= \frac{e^{a+b} - e^{-a-b}}{2} \\ &= \operatorname{senh}(a+b), \\ \\ \cosh a \cosh b + \operatorname{senh} a \operatorname{senh} b &= \frac{e^a + e^{-a}}{2} \frac{e^b + e^{-b}}{2} + \frac{e^a - e^{-a}}{2} \frac{e^b - e^{-b}}{2} \\ &= \frac{e^{a+b} + e^{a-b} + e^{-a+b} + e^{-a-b}}{4} + \frac{e^{a+b} - e^{a-b} - e^{-a+b} + e^{-a-b}}{4} \\ &= \frac{e^{a+b} + e^{-a-b}}{2} \\ &= \cosh(a+b). \end{aligned}$$

Daí,

$$\begin{aligned} \tanh(a+b) &= \frac{\operatorname{senh}(a+b)}{\cosh(a+b)} = \frac{\operatorname{senh} a \cosh b + \cosh a \operatorname{senh} b}{\cosh a \cosh b + \operatorname{senh} a \operatorname{senh} b} \\ &= \frac{\operatorname{senh} a \cosh b + \cosh a \operatorname{senh} b}{\cosh a \cosh b + \operatorname{senh} a \operatorname{senh} b} \\ &= \frac{\cosh a \cosh b}{\cosh a \cosh b + \operatorname{senh} a \operatorname{senh} b} \\ &= \frac{\operatorname{senh} a}{\cosh a} + \frac{\operatorname{senh} b}{\cosh b} \\ &= \frac{1}{1 + \frac{\operatorname{senh} a \operatorname{senh} b}{\cosh a \cosh b}} \\ &= \frac{\tanh a + \tanh b}{1 + \tanh a \tanh b}. \end{aligned}$$

Segue de (2.18) e da injetividade da função tangente hiperbólica que se

$$|v'|, |v''| < 1,$$

então existem únicos $a, b \in \mathbb{R}$ tais que

$$\begin{aligned} v' &= \tanh a, \\ v'' &= \tanh b. \end{aligned}$$

Portanto, segue de (2.19) que

$$v = \frac{\tanh a + \tanh b}{1 + \tanh a \tanh b} = \tanh(a+b),$$

onde

$$|v| < 1.$$

Se $v' = 1$, então

$$v = \frac{1+v''}{1-v''} = 1$$

e se $v' = -1$, então

$$v = \frac{-1+v''}{1-v''} = -1.$$

As mesmas conclusões valem para $v'' = \pm 1$. ■

2.4.6 Rotações Hiperbólicas

Considere a transformação de Lorentz bidimensional

$$L = \begin{bmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

com

$$\gamma = \frac{1}{\sqrt{1-v^2}}.$$

Escrevendo

$$v = \tanh \theta,$$

para um único $\theta \in \mathbb{R}$, como vimos na demonstração da Proposição 2.32, como

$$\frac{1}{\sqrt{1-\tanh^2 \theta}} = \frac{1}{\sqrt{1-\frac{\operatorname{senh}^2 \theta}{\cosh^2 \theta}}} = \frac{\cosh \theta}{\sqrt{\cosh^2 \theta - \operatorname{senh}^2 \theta}} = \frac{\cosh \theta}{1},$$

segue que

$$\gamma = \cosh \theta. \tag{2.20}$$

Daí

$$v\gamma = \tanh \theta \cosh \theta = \operatorname{senh} \theta.$$

Podemos portanto escrever

$$L = \begin{bmatrix} \cosh \theta & \operatorname{senh} \theta & 0 & 0 \\ \operatorname{senh} \theta & \cosh \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

Nas duas primeiras coordenadas, isso é simplesmente

$$\begin{bmatrix} t \\ x \end{bmatrix} = \begin{bmatrix} \cosh \theta & \operatorname{senh} \theta \\ \operatorname{senh} \theta & \cosh \theta \end{bmatrix} \begin{bmatrix} t' \\ x' \end{bmatrix}.$$

2.33 Definição. A transformação

$$R_\theta = \begin{bmatrix} \cosh \theta & \operatorname{senh} \theta \\ \operatorname{senh} \theta & \cosh \theta \end{bmatrix} \tag{2.21}$$

é chamada uma **rotação hiperbólica de rapidez** θ . □

Em termos da rapidez, se denotarmos por $L_{O \rightarrow O'}$ a transformação de Lorentz das coordenadas do referencial O' para o referencial O , a lei de adição de velocidades da teoria da relatividade pode ser escrita na forma

$$\theta(L_{O \rightarrow O''}) = \theta(L_{O \rightarrow O'}) + \theta(L_{O' \rightarrow O''}). \quad (2.22)$$

Portanto, o subgrupo dos boosts em relação a uma direção fixada é isomorfo ao grupo a um parâmetro de adições \mathbb{R} . Dado que boosts são rotações hiperbólicas e que as rotações em \mathbb{R}^3 formam um grupo (ou seja, a composta de duas rotações no espaço em relação ao mesmo eixo ou em relação a eixos diferentes é uma rotação no espaço em relação ao mesmo eixo ou em relação a um eixo diferente dos outros dois, respectivamente) podemos entender a surpresa de Einstein ao ser informado que os boosts não formam um grupo, ou seja, que a composta de rotações hiperbólicas em \mathbb{R}^4 não é uma rotação hiperbólica (em \mathbb{R}^2 evidentemente é).

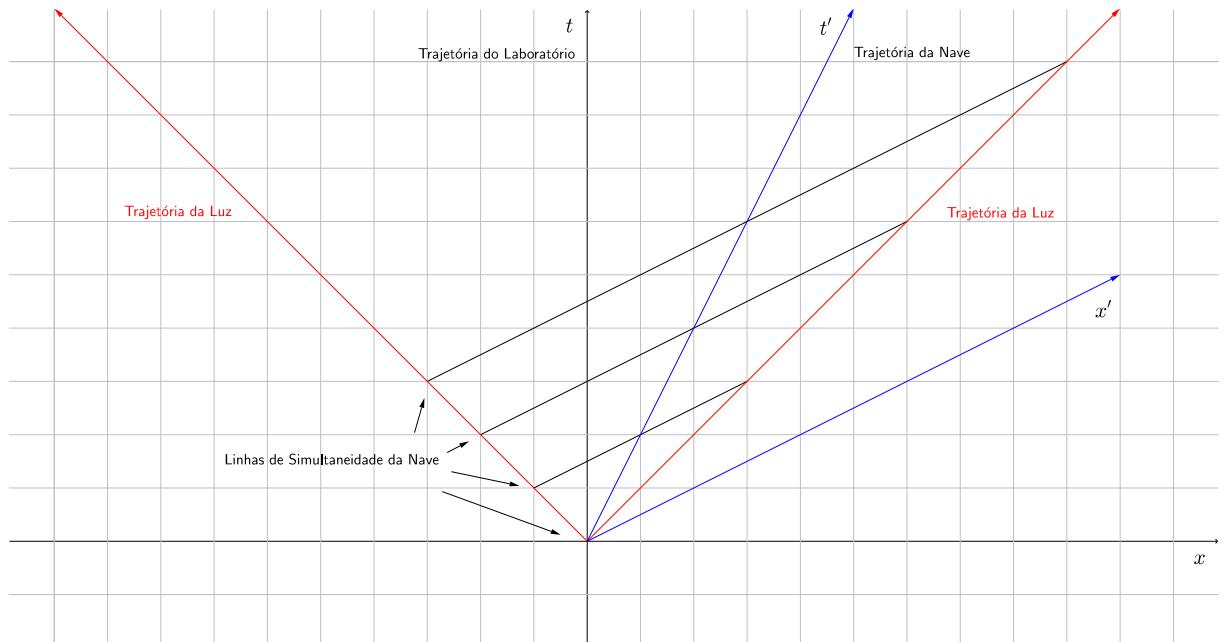
Capítulo 3

Paradoxos

3.1 Diagramas de Espaçotempo

Uma técnica eficiente, e com a vantagem adicional de ser visual e portanto fácil de entender, para resolver problemas de relatividade especial é a análise de **diagramas espaçotempo**. Nesta seção examinaremos alguns exemplos.

Em primeiro lugar, vamos entender como representar as coordenadas espaçotempo de dois referenciais inerciais no mesmo diagrama espaçotempo. Vamos nos restringir a uma dimensão temporal.



Escolha um dos referenciais inerciais. Escolhemos o referencial inercial *Laboratório*. Os eixos t, x neste referencial inercial são representados perpendicularmente, como usual. Em seguida representamos a trajetória no espaçotempo do referencial inercial *Nave* nas coordenadas do referencial Laboratório. Esta trajetória corresponde ao eixo t' do referencial inercial Nave, pois, ao longo desta trajetória, a posição da nave não

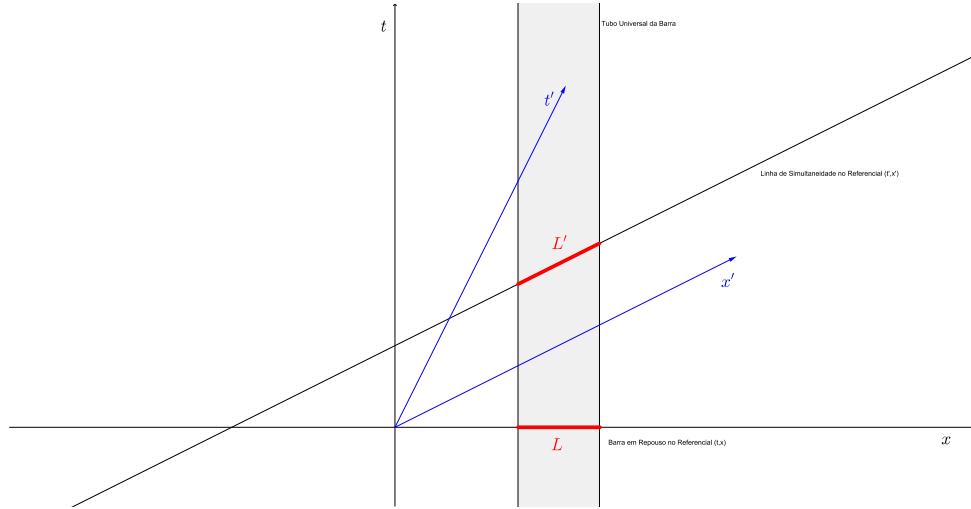
muda. Para obter o eixo x' do referencial inercial Nave, precisamos obter as linhas de simultaneidade em relação a este referencial, pois o eixo x' é exatamente o conjunto dos pontos que tem coordenada temporal $t' = 0$. Para obter as linhas de simultaneidade do referencial Nave, trace primeiro as duas trajetórias de raios de luz no referencial Laboratório; se os eixos tem a mesma escala (tempo e espaço medidos em metros), estas trajetórias fazem um ângulos de 45° com os eixos. As linhas de simultaneidade do referencial Nave são então as linhas paralelas tais que os pontos das interseções destas linhas com as trajetórias da luz são equidistantes da trajetória da nave (porque eles ocorrem ao mesmo tempo no referencial inercial Nave, já que a velocidade da luz é a mesma em todos os referenciais inerciais). Uma destas linhas de simultaneidade será o eixo x' ; qual delas dependerá da escolha da origem no referencial inercial Nave (na figura, as origens dos dois referenciais foi escolhida como sendo a mesma).

Usando diagramas espaço-tempo podemos por exemplo entender melhor a contração de Lorentz e a dilatação do tempo. No diagrama a seguir, a barra em repouso no referencial laboratório (t, x) tem comprimento próprio L . No referencial nave (t', x') , o comprimento L' da barra é medido quando os seus extremos são simultâneos, isto é, ao longo de uma linha de simultaneidade deste referencial. Portanto, o segmento correspondente a L' no diagrama é a hipotenusa de um triângulo retângulo cujo cateto correspondente a L é paralelo ao eixo x ; o outro cateto é a diferença temporal Δt entre os extremos do segmento L' (vistos como eventos no espaço-tempo) quando eles são medidos no referencial (t, x) , onde eles não são simultâneos. Como

$$(L')^2 = L^2 - (\Delta t)^2,$$

segue que

$$L' < L.$$

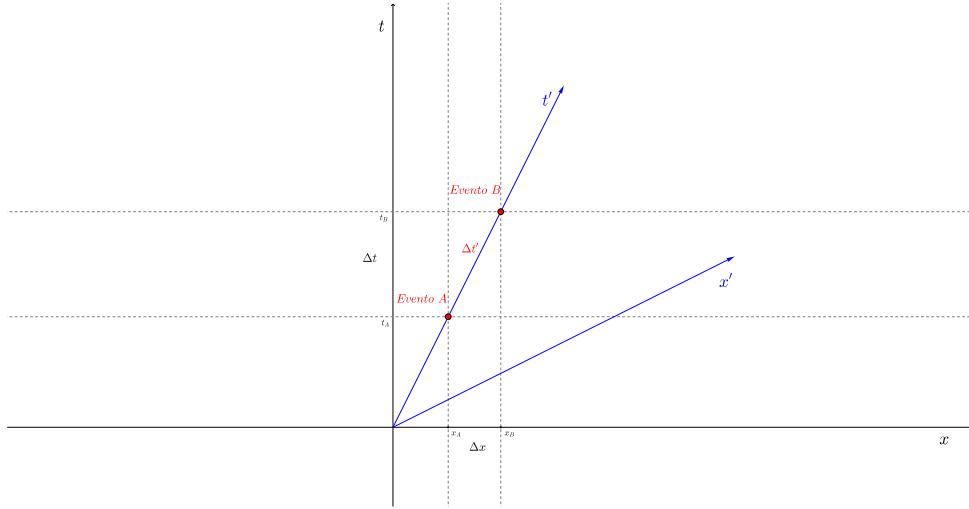


No próximo diagrama, os eventos A e B ocorrem na mesma posição $x' = 0$ no referencial nave (t', x') . O intervalo de tempo $\Delta t'$ decorrido entre eles é portanto o tempo próprio entre os dois eventos. Novamente, o segmento correspondente que tem estes eventos como extremos é a hipotenusa de um triângulo retângulo cujo cateto paralelo ao eixo t tem comprimento Δt , que é o intervalo de tempo entre os eventos A e B conforme eles são medidos no referencial laboratório (t, x) ; o outro cateto tem comprimento Δx e é o intervalo de espaço entre os eventos A e B conforme eles são medidos no referencial laboratório (t, x) . Como

$$(\Delta t')^2 = (\Delta t)^2 - (\Delta x)^2,$$

segue que

$$\Delta t > \Delta t'.$$



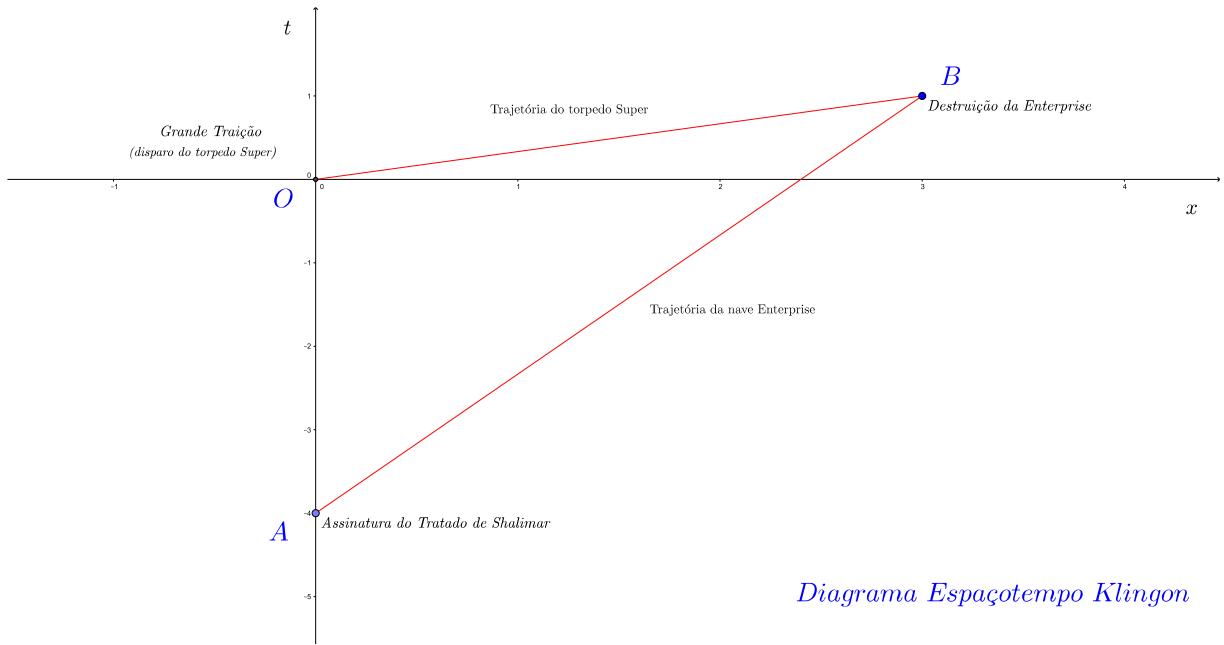
3.2 Impossibilidade de se viajar acima da velocidade da luz

Nesta seção quotamos diretamente de [Taylor-Wheeler], Box L-1, pp. 108–109, com pequenas modificações, para um exemplo dramático de como a possibilidade de um objeto material viajando à velocidade da luz viola a ordem natural de causa e efeito.

O *Tratado de Paz de Shalimar* entre a *Federação Unida dos Planetas* e o *Império Klingon* foi assinado 4 anos antes da *Grande Traição*. Sendo um evento tão crucial na história do *Quadrante Alfa* de nossa galáxia, ele foi tomado como o zero nos sistemas de coordenadas do espaço-tempo.

Pelo Tratado de Shalimar, os Klingons concordaram em parar de atacar os postos avançados da Federação em retorno por acesso ao *Banco de Dados Técnico da Federação*. Os negociadores da Federação deixaram o planeta Shalimar imediatamente após assinar o tratado, na nave estelar *Enterprise* viajando a 0,6 da velocidade da luz (por algum motivo até hoje inexplicado, o *warpdrive* não estava funcionando; os historiadores debatem se houve ou não sabotagem). Em 4 anos, os Klingons usaram o seu acesso garantido ao Banco de Dados Técnico da Federação para desenvolver um torpedo fotônico superavançado, de elevado alcance e poder de destruição, e com a *capacidade de viajar acima da velocidade da luz*, que eles chamaram de *Super*. Naquele dia negro da Grande Traição (evento de referência 0), os Klingons testaram o Super, disparando ele a partir do planeta Shalimar exatamente na *Enterprise* contendo os negociadores, a 3 vezes a velocidade da luz. O torpedo atingiu a nave destruindo-a completamente. Não houve sobreviventes. Um período ainda mais longo de guerras e batalhas estelares do que antes da assinatura do tratado teve sequência.

Os eventos estão descritos no diagrama de espaço-tempo abaixo, do ponto de vista dos Klingons, isto é, no referencial inercial do planeta Shalimar (coordenadas t, x).



Com obter uma visão equilibrada da História, vamos investigar os acontecimentos do ponto de vista da Federação, mais especificamente no referencial inercial da nave *Enterprise* (coordenadas t' , x'). Especificamente, quando ocorreram neste sistema de coordenadas a assinatura do Tratado de Shalimar (evento A) e o impacto do Super com a nave, que causou a sua completa destruição (evento B)?

Temos

$$(t_A, x_A) = (-4, 0), \\ (t_B, x_B) = (1, 3).$$

Como a velocidade da *Enterprise* em relação ao planeta Shalimar é 0,6, segue que o fator de dilatação do tempo é

$$\gamma = \frac{1}{\sqrt{1-v^2}} = \frac{1}{\sqrt{1-0,6^2}} = 1,25.$$

de modo que a transformação de Lorentz do referencial inercial de Shalimar para o referencial inercial *Enterprise* é

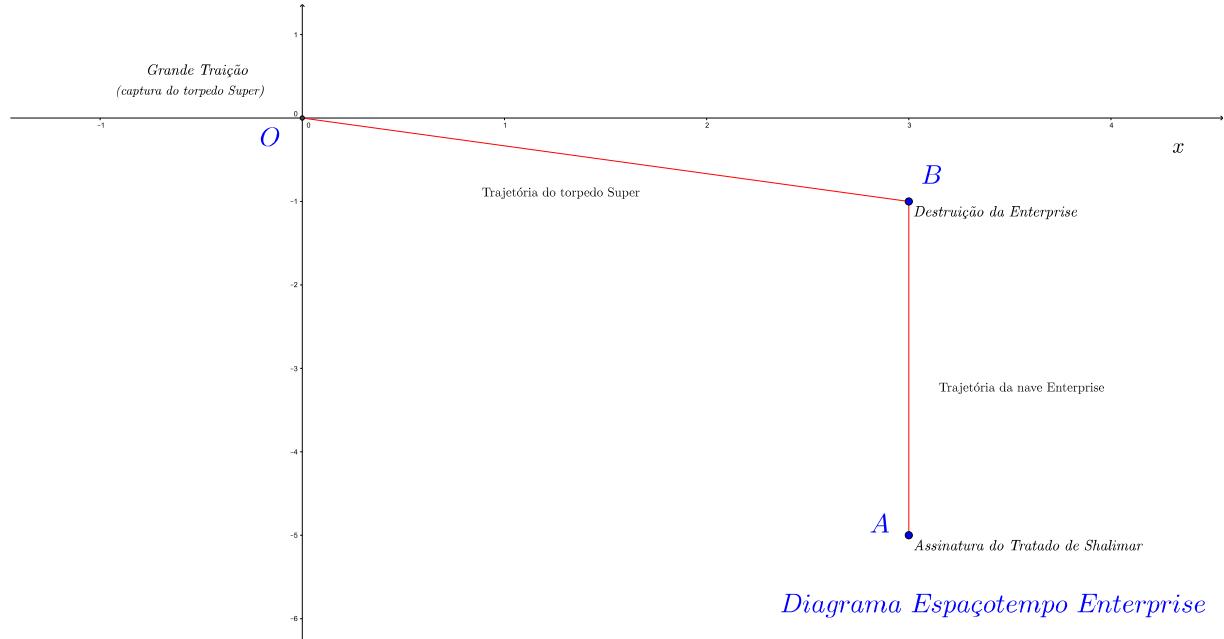
$$\begin{bmatrix} \gamma & -v\gamma \\ -v\gamma & \gamma \end{bmatrix} = \begin{bmatrix} 1,25 & -0,75 \\ -0,75 & 1,25 \end{bmatrix}.$$

Logo,

$$\begin{bmatrix} t'_A \\ x'_A \end{bmatrix} = \begin{bmatrix} 1,25 & -0,75 \\ -0,75 & 1,25 \end{bmatrix} \begin{bmatrix} t_A \\ x_A \end{bmatrix} = \begin{bmatrix} 1,25 & -0,75 \\ -0,75 & 1,25 \end{bmatrix} \begin{bmatrix} -4 \\ 0 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

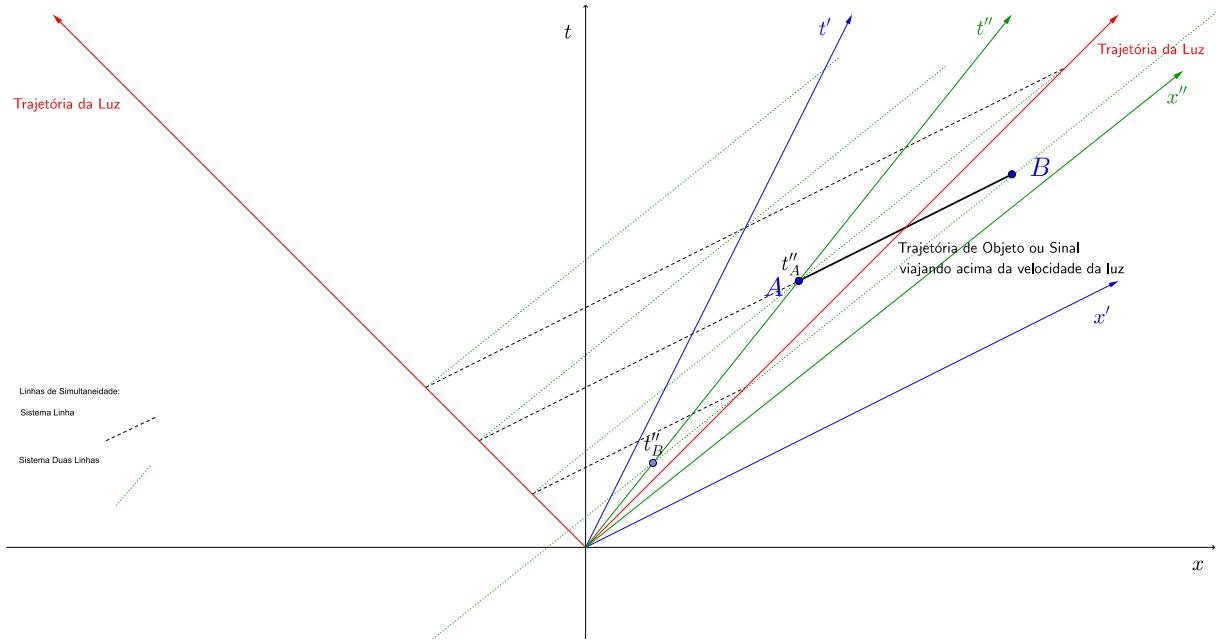
e

$$\begin{bmatrix} t'_B \\ x'_B \end{bmatrix} = \begin{bmatrix} 1,25 & -0,75 \\ -0,75 & 1,25 \end{bmatrix} \begin{bmatrix} t_B \\ x_B \end{bmatrix} = \begin{bmatrix} 1,25 & -0,75 \\ -0,75 & 1,25 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$



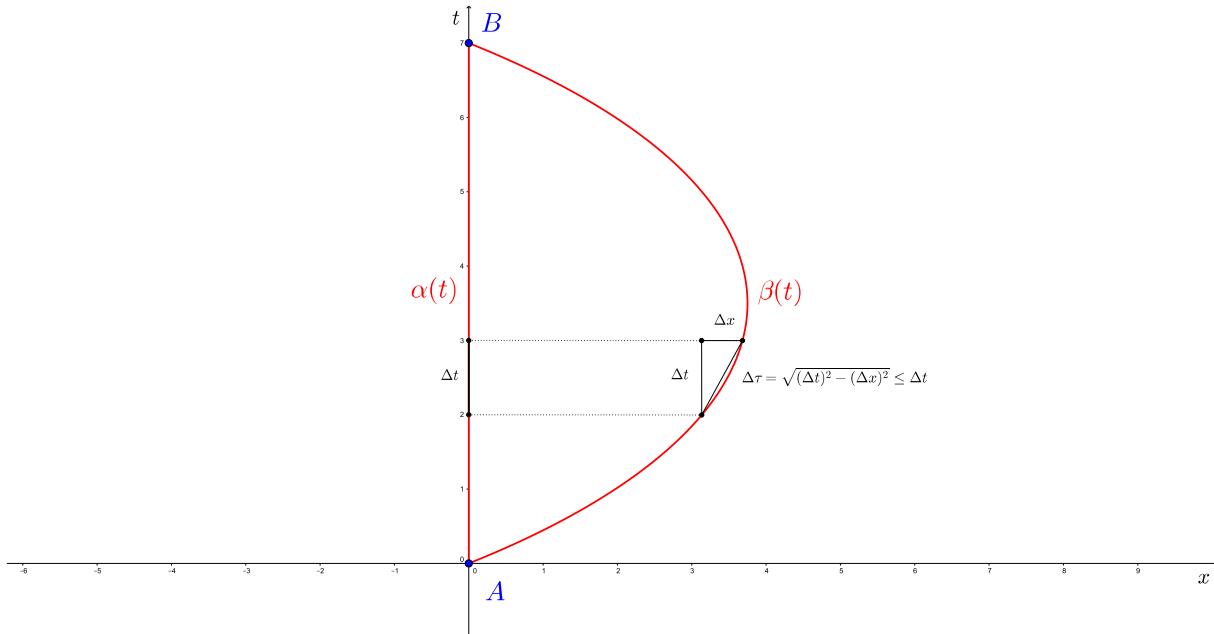
Portanto, no referencial inercial da *Enterprise*, sua explosão ocorre antes que os Klingons disparem o torpedo. Na verdade, se formos interpretar os eventos ocorridos no diagrama de espaçotempo da *Enterprise*, ela misteriosamente explode ao mesmo tempo que dispara um novo torpedo fotônico em orientação reversa (propulsor superluz na frente da ogiva) na direção do planeta Shalimar, torpedo este que miraculosamente entra no lançador dos Klingons sem causar-lhes quaisquer danos. O Império Klingon usa esta interpretação dos dados para acusar a Federação de duplicitade e de ela ter causado o reinício das hostilidades, disparando um torpedo defeituoso a partir da nave dos negociadores que causou a destruição da própria nave estelar agressora, mas que por sorte foi capturado pelos Klingons antes de causar danos.

Em geral, se algum objeto material ou sinal puder exceder a velocidade da luz, a ordem cronológica de eventos será alterada em algum referencial inercial. De fato, se isso ocorre, a trajetória no espaçotempo do objeto material ou sinal terá uma inclinação menor que 45° em relação ao eixo x . Logo, existirá um referencial inercial em que estes eventos ocorrem simultaneamente. E quando observados a partir de qualquer referencial inercial movendo-se com uma velocidade maior que este segundo referencial, a ordem cronológica dos eventos será alterada. Por exemplo, no diagrama spaçotempo abaixo, o evento A é a emissão de um sinal ou lançamento de um objeto (projétil) acima da velocidade da luz e o evento B é a recepção deste sinal ou projétil. É fácil ver no referencial inercial (t, x) que o sinal ou objeto viaja acima da velocidade da luz, pois a sua trajetória tem inclinação menor que a inclinação da trajetória de um raio de luz neste referencial. Obtemos então um referencial inercial (t', x') em que esta trajetória é paralela às linhas de simultaneidade neste referencial; isso significa que os eventos A e B ocorrem simultaneamente neste referencial. Ou seja, nos deparamos com um sinal instantâneo (comunicação instantânea) ou com o teletransporte instantâneo de um objeto material. Já no referencial inercial (t'', x'') , o evento B ocorre antes do evento A : o sinal ou objeto foi primeiro recebido e depois foi enviado.



3.3 Geodésicas no Espaço de Minkowski

Considere os eventos A e B no espaço-tempo e as duas trajetórias possíveis α, β de dois observadores de um evento a outro de acordo com diagrama espaço-tempo abaixo, em relação ao referencial inercial do observador que segue a trajetória α .



O observador que segue a trajetória α permanece em repouso durante todo o tempo, isto é, os eventos A e B ocorrem na mesma posição em seu referencial inercial. Portanto, como já vimos antes, o tempo próprio medido no seu referencial é exatamente

$$t_B - t_A.$$

O observador que segue a trajetória curva β no espaço-tempo claramente não está em um referencial inercial, já que o próprio fato da sua trajetória ser curva em relação a um referencial inercial mostra que ele sofre aceleração ao longo de sua trajetória. Por outro lado, faz sentido comparar o tempo próprio que ele mede (em seu referencial não inercial), ou seja, quanto tempo se passa medido no seu relógio de pulso ou no seu relógio biológico. Dividindo a sua trajetória em segmentos infinitesimais e usando a métrica de Lorentz, concluímos como na figura que em cada um destes segmentos da sua trajetória vale

$$(\Delta\tau)^2 = (\Delta t)^2 - (\Delta x)^2 \leq (\Delta t)^2,$$

ou seja,

$$\Delta\tau \leq \Delta t.$$

Somando (integrando) as contribuições destes segmentos infinitesimais ao longo da trajetória toda, concluímos que o tempo próprio medido seguindo-se a trajetória β é *menor* que o tempo próprio medido ao longo da trajetória α , onde o observador encontra-se em repouso. Mais precisamente, podemos escrever

$$(\Delta\tau)^2 = (\Delta t)^2 - (\Delta x)^2 = \left[1 - \left(\frac{\Delta x}{\Delta t} \right)^2 \right] (\Delta t)^2,$$

de modo que

$$\Delta\tau = \sqrt{1 - v^2} \Delta t,$$

onde

$$\Delta\tau_{AB} = \int_{t_A}^{t_B} \sqrt{1 - v^2} dt.$$

Daí, como $\sqrt{1 - v^2} < 1$, segue que

$$\Delta\tau_{AB} < \int_{t_A}^{t_B} dt = \Delta t_{AB}.$$

Assim, trajetórias curvas diminuem o tempo próprio, de modo que

retas tem tempo próprio máximo.

Embora acabamos de ver isso apenas no caso em que o observador encontra-se em repouso, isto é, eventos A e B ocorrem no mesmo referencial, é fácil obter a mesma conclusão no caso geral, fazendo-se uma mudança de coordenadas.

3.4 Paradoxo dos Gêmeos

Esta seção é adaptada de [Takeuchi], Seção 6.6, pp. 148–153 e [Taylor-Wheeler], Seções 4.6 (pp. 125–126) e 4.9 (pp. 128–131) e, crucialmente, Exercício 5-8 (p. 169).

Uma nave espacial viaja da Terra para a estrela Canopus, percorrendo uma distância de 99 anos-luz a uma velocidade 99/101 da luz. Isso significa que ela viaja os 99 anos-luz de distância, medidos no referencial inercial Terra em 101 anos de tempo. Como vimos na Seção 1.7.2, o tempo que os astronautas medem no referencial Nave, isto é, o tempo próprio decorrido entre os eventos A (partida da espaçonave da Terra) e B (chegada da nave em Canopus) é 20 anos. Segue que se a nave retornar para a Terra com a mesma velocidade, o tempo decorrido entre os eventos A e C (retorno da espaçonave à Terra) é medido no referencial inercial Terra como sendo de 202 anos, enquanto que o tempo medido no referencial Nave será de apenas 40

anos. Para efeitos dramáticos, se um dos astronautas tivesse um irmão gêmeo na época que partiu da Terra e este tivesse permanecido lá, quando ele retornar, o irmão gêmeo terrestre terá envelhecido 202 anos (mas talvez não tenha morrido, pois se foi desenvolvida tecnologia para viajar a velocidades próximas da luz, é possível que também foi desenvolvida tecnologia para permitir às pessoas viverem consideravelmente mais tempo) enquanto que o irmão gêmeo astronauta terá envelhecido apenas 40 anos.

O *paradoxo dos gêmeos* é o seguinte: como todos os referenciais inerciais são equivalentes, na primeira parte da viagem, da Terra a Canopus, o astronauta mede no referencial inercial Nave que a Terra viajou para um ponto situado a 99 anos-luz de distância de sua posição a uma velocidade $99/101$ da luz, levando portanto 101 anos para atingir aquela posição, enquanto que do ponto de vista do referencial Terra, só se passaram 20 anos; na segunda parte da viagem, a Terra retornou à posição que o espaçonave ocupa, tendo gastado novamente 101 anos de acordo com o referencial Nave e 20 anos de acordo com o referencial Terra. Portanto, passaram-se 202 anos na espaçonave e apenas 40 anos na Terra, exatamente o contrário do que havíamos concluído no parágrafo anterior.

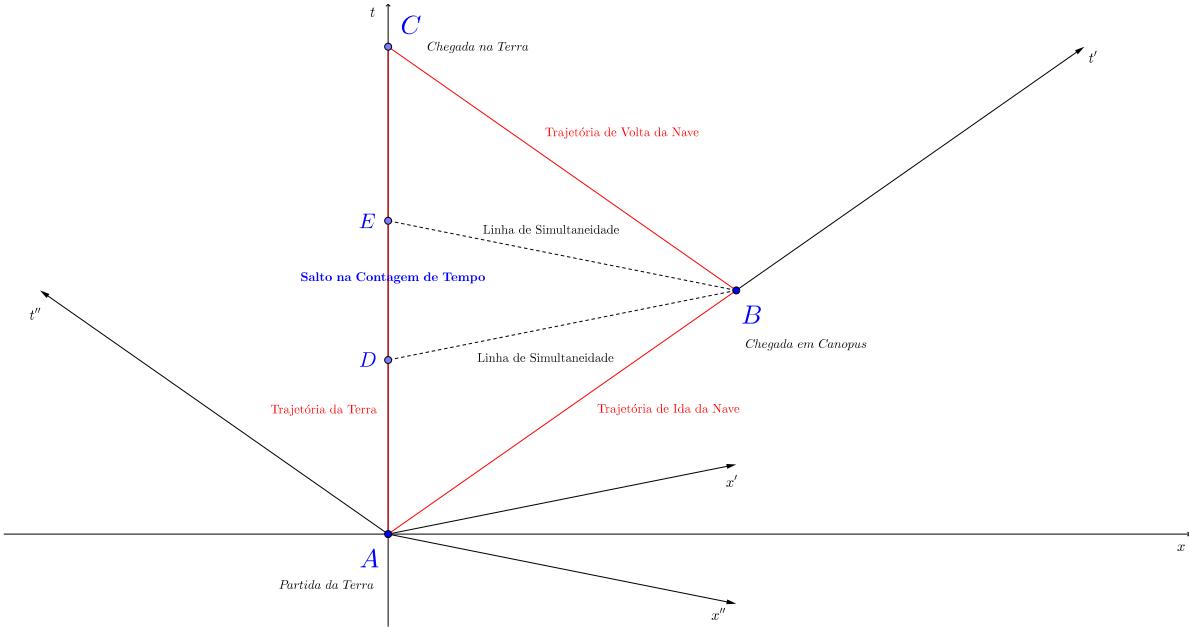
A resolução trivial deste paradoxo é que enquanto o referencial Terra é de fato um referencial inercial pois permanece em repouso (relativamente) durante toda a viagem, o referencial Nave é um referencial inercial na primeira parte da viagem, mas é um referencial inercial *diferente* na segunda parte da viagem. No momento em que ele muda a direção do seu movimento para retornar para a Terra, ele deixa de ser um referencial inercial, ele passa de um referencial inercial para outro.

Mas ainda há duas questões interessantes a resolver:

1. Será que os cálculos que fizemos anteriormente ainda estão corretos, isto é, decorreram-se 202 anos na Terra e apenas 40 anos na espaçonave, ou a aceleração que a espaçonave sofreu quando mudou de direção para retornar à Terra afeta os cálculos?
2. A questão anterior pode ser resolvida no âmbito da Teoria da Relatividade Especial, que trata apenas de referenciais inerciais em movimento retilíneo uniforme um com relação ao outro, ou, já que existe uma aceleração envolvida na mudança da direção da espaçonave, esta questão só pode ser resolvida no âmbito da Teoria da Relatividade Geral que, como veremos, trata também de referenciais acelerados?

Em [Taylor-Wheeler], Box 4-1, p. 132, os autores afirmam categoricamente que não há necessidade de se usar a Teoria da Relatividade Geral. Já em [Stephani], Seção 3.4, pp. 20–21, o autor afirma que para uma resposta confiável, é necessário recorrer à Teoria da Relatividade Geral. Ambos estão corretos.

A Teoria da Relatividade Especial se aplica apenas a observações (medidas) feitas em referenciais inerciais, de modo que não podemos realmente dizer o que aconteceu com o astronauta durante o processo de aceleração. Mas, supondo que podemos tornar o processo de aceleração tão curto quanto quisermos, podemos considerar um salto instantâneo de um referencial para o outro. Acompanhe a nossa discussão no diagrama espaço-tempo a seguir. O referencial inercial da Terra tem coordenadas (t, x) , o referencial inercial da nave na ida tem coordenadas (t', x') e o referencial inercial da nave na volta tem coordenadas (t'', x'') . O salto entre referenciais inerciais ocorre no evento B (chegada da nave a Canopus e, simultaneamente, partida de volta para a Terra). Imediatamente na chegada a Canopus, o evento B é simultâneo ao evento D na Terra. Com a mudança instantânea de referenciais (mudança instantânea da direção da trajetória da nave), o tempo medido na Terra dá um salto de D para E (observe as linhas de simultaneidade; a linha inferior é a linha de simultaneidade no referencial inercial da nave na ida, isto é, os eventos C e D são simultâneos neste referencial, e a linha de simultaneidade superior é a linha de simultaneidade do referencial inercial da nave na volta, isto é, os eventos C e E são simultâneos neste outro referencial; os eventos D e E não estão em escala no eixo t , pois, como veremos a seguir, eles estão relativamente muito próximos dos eventos A e C , respectivamente). Por causa deste efeito, o astronauta será mais jovem que o seu irmão gêmeo quando retornar à Terra.



Vamos determinar a localização no espaçotempo dos eventos D, E nas coordenadas do referencial inercial Terra para calcular o tempo decorrido lá entre a partida da nave e o seu retorno. Para isso vamos encontrar as equações para as linhas de simultaneidade nas coordenadas (t, x) . O fator de dilatação do tempo é

$$\gamma = \frac{1}{\sqrt{1-v^2}} = \frac{1}{\sqrt{1-\left(\frac{99}{101}\right)^2}} = \frac{101}{20}.$$

Como

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} \gamma & -v\gamma \\ -v\gamma & \gamma \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix},$$

obtemos a equação para a linha de simultaneidade inferior

$$t' = \gamma t - v\gamma x = \frac{101}{20}t - \frac{99}{20}x,$$

isto é, como $t'_D = 20$,

$$20 = \frac{101}{20}t - \frac{99}{20}x.$$

Como $x_D = 0$, segue que

$$t_D = \frac{400}{101} = 3,96,$$

ou seja, quando a nave chega em Canopus depois de 20 anos, como medido no seu referencial, ela mede que transcorreram 3,96 anos na Terra. Quando a nave salta de referencial, após mudar de direção, a linha de simultaneidade superior tem inclinação igual e com sinal oposto ao da linha de simultaneidade inferior. Como a posição do evento C no referencial Terra é $(t_C, x_C) = (202, 0)$, pois o ponto médio da trajetória da Terra no espaçotempo tem posição $(101, 0)$, simultâneo ao evento B no referencial inercial Terra, como vimos, segue por simetria que

$$t_E = 202 - 3,96 = 198,04,$$

ou seja, nesta mudança instantânea de referenciais há um salto na medida do tempo transcorrido na Terra conforme a nave mede: o tempo na Terra conforme medido no novo referencial inercial da nave na volta é que transcorreram na Terra 198,04 anos. Assim, quando a nave retorna à Terra, transcorreram na Terra 202 anos, enquanto que na nave transcorreram apenas 40 anos.

É claro que este salto na contagem de tempo é fisicamente impossível, assim como é fisicamente impossível uma mudança instantânea na direção da nave. Numa situação física real, a nave levaria algum tempo para mudar de direção, de modo que a sua trajetória no espaço-tempo seria uma curva diferenciável, com as linhas de simultaneidade mudando de forma contínua à medida que a nave desacelera e acelera para mudar de direção. Teríamos então uma trajetória curva que, como vimos na Seção 3.3, tem um tempo próprio menor que o tempo medido na trajetória reta entre os eventos A e C ; uma análise detalhada do paradoxo dos gêmeos ao longo deste argumento pode ser vista em [Ferraro], Seção 4.5, pp. 103–107.

A história do paradoxo dos gêmeos começa no ano 1911 e foi um dos fenômenos mais discutidos e incomprendidos da teoria, mas a ocorrência do fenômeno foi confirmada experimentalmente um número enorme de vezes, de forma direta depois da invenção dos relógios atômicos em 1949 (relógios atômicos em aviões marcaram um tempo menor do que o decorrido no chão entre decolagem e aterrissagem). Veja [Topper], pp. 77–78, 127–129 (a resolução de Einstein usando a Teoria da Relatividade Geral; aceleração, assim como a gravidade, faz com que o tempo ande mais devagar), pp. 145–146 (uma comprovação através dos inúmeros experimentos realizados em aceleradores de partículas em circuito fechado).

Capítulo 4

Momento-Energia

4.1 4-Vetores

Tanto na Teoria da Relatividade Especial quanto na Geral, é conveniente descrever o movimento de um objeto em um referencial inercial (em Relatividade Especial) ou em um sistema de coordenadas geral (em Relatividade Geral) em coordenadas do espaço-tempo dependendo do seu tempo próprio τ , isto é, de acordo com o tempo medido em seu próprio referencial, e portanto independente de referencial. Assim, em um referencial inercial a trajetória de uma partícula livre é uma curva parametrizada pelo tempo próprio

$$\alpha(\tau) = (t(\tau), x(\tau), y(\tau), z(\tau)),$$

e a geodésica γ que uma partícula livre segue no espaço-tempo curvo seria parametrizada pelo tempo próprio $\gamma(\tau)$.

Vetores no espaço-tempo tem 4 coordenadas e para distingui-los dos vetores espaciais (projeções espaciais), são frequentemente chamados de *quadrivetores* ou 4-vetores. Assim, temos o **quadrivetor deslocamento**, ou seja deslocamento no espaço-tempo

$$\Delta \mathbf{s} = (\Delta t, \Delta x, \Delta y, \Delta z)$$

e o **quadrivetor velocidade**

$$\mathbf{u} = \frac{d\mathbf{s}}{d\tau} = \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right), \quad (4.1)$$

ou $\alpha'(\tau)$ para uma trajetória curva $\alpha(\tau)$ no espaço-tempo. O quadrivetor velocidade é por definição tangente à trajetória da partícula no espaço-tempo.

Observe que se a trajetória do objeto no espaço-tempo é uma curva, o seu referencial próprio obviamente não é um referencial inercial. Entretanto, em cada instante de tempo existe um referencial inercial que tem a mesma velocidade que o objeto; este referencial inercial que *comove* com o objeto é temporário, pois no próximo instante de tempo ele não está mais se comovendo com o objeto. Ele é chamado o seu *referencial comovente temporário* (em inglês: *momentarily comoving reference frame* (MCRF)). O quadrivetor velocidade em um determinado instante de tempo (tangente à trajetória da partícula naquele instante de tempo) é justamente o vetor base e_0 do seu referencial comovente temporário naquele momento.

Como o tempo próprio τ do objeto e o tempo t de um referencial inercial arbitrário estão relacionados através da equação

$$\tau = \sqrt{1 - v^2} t, \quad (4.2)$$

onde \mathbf{v} é a velocidade espacial do objeto em relação a este referencial inercial ($v = \|\mathbf{v}\|$), isto é,

$$\mathbf{v} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right),$$

segue que

$$d\tau = \sqrt{1 - v^2} dt,$$

e portanto as coordenadas do quadrivetor velocidade no referencial inercial também podem ser escritas na forma

$$\mathbf{u} = \frac{1}{\sqrt{1 - v^2}} \left(1, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = \frac{1}{\sqrt{1 - v^2}} (1, \mathbf{v})$$

ou

$$\mathbf{u} = \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z) = \frac{1}{\sqrt{1 - v^2}} (1, \mathbf{v}). \quad (4.3)$$

Observe que quando $v \ll 1$ temos $\tau \approx t$ e

$$\mathbf{u} \approx (1, v^x, v^y, v^z) = (1, \mathbf{v}),$$

e o quadrivetor velocidade de um objeto em repouso em relação ao referencial inercial (trajetória vertical no espaço-tempo) tem neste referencial coordenadas

$$\mathbf{u} = (1, 0, 0, 0).$$

Em particular, segue também que se $v \neq 1$ o quadrivetor velocidade é um *vetor unitário*, isto é,

$$\|\mathbf{u}\|^2 = -1 \quad (4.4)$$

pois

$$\begin{aligned} \langle \mathbf{u}, \mathbf{u} \rangle &= \left\langle \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z), \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z) \right\rangle \\ &= \frac{1}{1 - v^2} \left[-1 + (v^x)^2 + (v^y)^2 + (v^z)^2 \right] \\ &= \frac{1}{1 - v^2} [-1 + v^2] \\ &= -1. \end{aligned}$$

4.2 Momento-Energia

4.1 Definição. Se m é a massa de um objeto, definimos o **quadrivetor momento-energia**

$$\mathbf{P} = m\mathbf{u}. \quad (4.5)$$

□

Em um referencial inercial arbitrário, segue que o quadrivetor momento-energia tem coordenadas

$$\mathbf{P} = \frac{m}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z) = \frac{m}{\sqrt{1 - v^2}} (1, \mathbf{v}).$$

Observe que como velocidade é adimensional em relatividade (tempo e espaço são medidos com a mesma unidade), o quadrivetor momento-energia tem unidades de massa. De fato, se $v \neq 1$

$$\langle \mathbf{P}, \mathbf{P} \rangle = \langle m\mathbf{u}, m\mathbf{u} \rangle = m^2 \langle \mathbf{u}, \mathbf{u} \rangle = -m^2,$$

de modo que

$$\|\mathbf{P}\|^2 = -m^2. \quad (4.6)$$

Note que a massa m de um objeto, sendo a norma do quadrivetor momento-energia (um vetor do espaço-tempo de Minkowski), tem o mesmo valor em qualquer referencial inercial. Em outras palavras, ela não depende da velocidade do referencial: massa não aumenta com a velocidade e na teoria da relatividade não se define “massa relativística” (veja [Okun1] para uma exposição clara, incluindo a origem histórica deste termo e até seu uso inicial por Einstein para inferir a atração da luz por uma massa gravitacional (ele raciocinou que fótons teriam massa $m = E/c^2$), antes de formular a Teoria da Relatividade Geral, quando abandonou essa idéia inconsistente, e uma coletânea de artigos criticando os conceitos de “massa de repouso” e “massa relativística” em [Okun2]). Por outro lado, massa é mais facilmente medida em um referencial em que o objeto está em repouso.

Se $m = 0$, então \mathbf{P} é um vetor do tipo luz, e consequentemente \mathbf{u} também é, donde $v = 1$. Portanto, fótons não possuem massa.

4.2 Definição. A componente temporal do quadrivetor momento-energia é chamada a **energia relativística** E do objeto e sua projeção espacial o seu **momento relativístico** \mathbf{p} . \square

Assim,

$$\begin{aligned} E &= \frac{m}{\sqrt{1-v^2}}, \\ \mathbf{p} &= \frac{m}{\sqrt{1-v^2}} \mathbf{v} = \left(\frac{mv^x}{\sqrt{1-v^2}}, \frac{mv^y}{\sqrt{1-v^2}}, \frac{mv^z}{\sqrt{1-v^2}} \right), \end{aligned}$$

de modo que

$$m^2 = E^2 - p^2. \quad (4.7)$$

Note que tanto a energia relativística quanto o momento relativístico dependem do referencial, já que são componentes (coordenadas) de um quadrivetor. Observe também que massa não é a mesma coisa que energia: energia é a componente temporal do quadrivetor momento-energia, enquanto que massa é a magnitude deste vetor. Apenas que, em relatividade, elas são medidas com as mesmas unidades de medida (quilogramas), da mesma forma que tempo e espaço em relatividade são medidos com as mesmas unidades de medida (metros), apesar de não serem a mesma coisa.

Em unidades do espaço-tempo (ct, x, y, z) em que a velocidade da luz é c e não normalizada para 1, a relação entre o tempo próprio e o tempo em um referencial inercial arbitrário é

$$\tau = \sqrt{1 - \left(\frac{v}{c}\right)^2} t$$

e o quadrivetor velocidade tem coordenadas

$$\mathbf{u} = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, v^x, v^y, v^z) = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, \mathbf{v}),$$

de modo que o momento-energia é

$$\mathbf{P} = \frac{m}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, v^x, v^y, v^z) = \frac{m}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, \mathbf{v}),$$

mas a energia relativística é

$$E = \frac{mc^2}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

(para que as unidades de medida de energia fiquem corretas) e o momento relativístico é

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}},$$

ou seja,

$$\mathbf{P} = \left(\frac{E}{c}, \mathbf{v} \right).$$

Quando $v = 0$, temos

$$E_0 = m$$

ou, quando a velocidade da luz não é normalizada, a célebre equação

$$E_0 = mc^2$$

para a energia de repouso; ou seja, a energia de um objeto não desaparece quando ele está em repouso, como na mecânica de Newton, mas ao contrário é enorme (por causa do fator c^2).

Capítulo 5

Aceleração no Espaçotempo de Minkowski

Capítulo 6

O Tensor do Campo Eletromagnético

Capítulo 7

O Princípio de Equivalência – Geodésicas no Espaçotempo Curvo

7.1 O Pensamento Mais Feliz da Vida de Einstein

A Teoria da Gravitação Universal de Newton, baseada na ação à distância instantânea de uma força gravitacional oculta é completamente incompatível com a Teoria da Relatividade, onde nada pode viajar a uma velocidade superior a da luz. Einstein precisava, portanto, inventar uma nova teoria de gravitação (além da necessidade óbvia de estender a teoria da relatividade para referenciais acelerados). O mais importante insight que Einstein teve na sua busca pela Teoria da Relatividade Geral foi em 1907, conforme relatado pelo mesmo em suas próprias palavras (quotado da tradução do original alemão para o inglês em [Topper], Cap. 12, p. 89):

Quando eu estava ocupado [em 1907] escrevendo um resumo do meu trabalho na teoria da relatividade especial ..., eu também tinha que tentar modificar a teoria da gravitação de Newton para adaptar as suas leis na teoria. Por mais que tentativas nesta direção mostraram a praticabilidade do empreendimento, elas não me satisfaziam porque elas tinham que ser baseadas em hipóteses sem fundamento físico. Naquele momento, eu tive o pensamento mais feliz da minha vida [der Glücklichste Gedanke meines Lebens, que pode ser alternativamente traduzido também como o mais sortudo, ou o mais afortunado ou o mais bem sucedido pensamento da minha vida.], na seguinte forma:

Em um exemplo que vale a pena considerar, o campo gravitacional tem uma existência relativa similar ao campo elétrico gerado por indução eletromagnética. Porque, para um observador em queda livre do telhado de uma casa, durante a sua queda não existe nenhum campo gravitacional, pelo menos em sua vizinhança imediata. Especificamente, se o observador soltar quaisquer corpos, eles permanecem em um estado de repouso ou movimento uniforme, relativamente a ele, independentemente de sua natureza química ou física especial [ignorando a resistência do ar, é claro]. O observador, portanto, está justificado em interpretar seu estado como sendo em repouso.

Em uma palestra em Kyoto, ele caracterizou este pensamento da seguinte maneira: *Repentinamente, uma idéia me ocorreu: ‘Se uma pessoa cair livremente, ela não deve sentir o seu próprio peso.’ Imediatamente eu me surpreendi. Este pensamento simples me deixou com uma impressão profunda, com efeito.*

Ou seja, localmente não existe gravidade, o espaço-tempo é o espaço-tempo de Minkowski, isto é, a métrica é a métrica de Lorentz, da Relatividade Especial. Como vimos na Seção 1.3, localmente no espaço e no tempo um pequeno elevador em queda livre é um referencial inercial. Localmente quer dizer eventualmente em um ponto no espaço-tempo, já que em uma região finita do espaço-tempo sempre aparecem os efeitos de maré, se a resolução é suficientemente boa.

A gravidade surge como um fenômeno global, considerando como os vários referenciais inerciais variam na vizinhança de um ponto, isto é, a gravidade surge quando se cola os espaços-tempos de Lorentz em cada

ponto para obter uma estrutura global.

Assim, para se obter uma Teoria da Gravitação, é necessário primeiro saber como colar estes vários espaçotempos de Minkowski. Por sorte, o maquinário matemático para fazer isso foi desenvolvido no início da segunda metade do século XIX por Riemann. Por azar, quando Einstein era estudante de Física no Instituto Politécnico de Zurich na Suíça, ele “matava” as aulas de Matemática e preferia passar o tempo nos laboratórios, observando os experimentos que eram realizados. Por sorte, ele tinha um bom amigo, Marcel Grossmann, especialista em geometria não euclideana. A partir de 1912, eles trabalharam juntos, publicando uma série de artigos a partir de 1912 que tiveram um papel importante na jornada até a Teoria da Relatividade Geral. Mas os esforços não foram realmente bem sucedidos até que Einstein reintroduziu a *regra da covariância* que havia na Teoria da Relatividade Especial, isto é, que as equações deveriam ter a mesma forma em todos os referenciais inerciais (até então, os dois tinham trabalhado sem assumir que as equações eram covariantes). A Teoria da Relatividade Geral foi anunciada por Einstein sozinho em novembro de 1915 em uma série de 4 palestras na Academia de Ciências da Prússia e publicada em março de 1916. Este artigo, “A Fundação da Teoria da Relatividade Geral” começa com um reconhecimento do trabalho de Minkowski (recorde o final da Seção 2.3), “que foi o primeiro a reconhecer a equivalência formal das coordenadas espaciais e as coordenadas temporais, e usou isso na construção da teoria”. Além disso, Einstein “em reconhecimento” agradeceu “meu amigo, o matemático Grossmann, cuja ajuda não apenas me salvou o esforço de estudar a literatura matemática pertinente, como também me ajudou na busca das equações de campo da gravitação.” ([Topper], Cap. 15, p. 113.)

Assim, o pensamento mais feliz da vida de Einstein, que deu origem à Teoria da Relatividade Geral, é que em cada ponto, o espaço-tempo é de Minkowski. A maneira como colar estes espaços de Minkowski em cada ponto para obter a geometria global do espaço-tempo, isto é, a *curvatura* do espaço-tempo é ditada pela Geometria Riemanniana. Neste momento temos que fazer uma pausa na Física e um grande desvio para aprender esta geometria, que requer uma série de pré-requisitos para ser bem compreendida. É o que faremos nos próximos capítulos.

7.1.1 Inércia

Antes, porém, é necessário considerar o conceito de **inércia** e as trajetórias que as partículas livres percorrem no espaço-tempo. Isso nos guiará na motivação dos conceitos necessários para o desenvolvimento do espaço-tempo curvo (variedade de Lorentz) físico além da simples colagem de espaços tangentes de Minkowski.

Newton aceitou o Princípio da Relatividade Galileana: não existe repouso e movimento uniforme absoluto. Mas existe a *aceleração absoluta* e ela é a justificativa para a existência do espaço absoluto. A aceleração absoluta é em relação a este espaço. Além disso, a *inércia* é uma medida da resistência de um corpo a sofrer uma aceleração, e esta é dada pela massa do corpo.

No espaço-tempo plano de Minkowski as geodésicas são retas. Quaisquer trajetórias curvas implicam na existência de uma aceleração e portanto de uma força agindo. A aceleração é absoluta, mas não é necessário postular um espaço absoluto para entendê-la. Ela existe no espaço-tempo absoluto de Minkowski.

No espaço-tempo curvo da Relatividade Geral, partículas que se movem apenas por inércia, sem sofrer a ação de forças, também se movem ao longo de geodésicas. Na geometria de Lorentz elas não sofrem qualquer alteração de direção, isto é, as partículas que seguem trajetórias geodésicas no espaço-tempo não sofrem aceleração (não estão sujeitas a ação de forças). Elas se movem por inércia, seguindo as trajetórias geodésicas do espaço-tempo. Apenas as geodésicas do espaço-tempo curvado da relatividade geral são curvas, logo na projeção tridimensional existe uma aceleração. Além disso, devido à curvatura do espaço-tempo, as trajetórias geodésicas se aproximam, o que era interpretado como os corpos se atraindo sujeitos à força da gravidade. Isso não existe no espaço-tempo curvo. Portanto, não existe uma força de gravidade: a gravidade é apenas uma manifestação da curvatura do espaço-tempo.

Forças só existem quando as partículas são impedidas de percorrer sua trajetória geodésica no espaço-tempo. Por exemplo, quando um corpo está apoiado sobre o chão, impedido de cair, impedido de seguir a sua trajetória geodésica natural no espaço-tempo curvado.

Capítulo 8

Topologia Diferencial

8.1 Variedades Diferenciáveis

8.1 Definição. Seja M um espaço topológico de Hausdorff com base enumerável. Um **atlas** de dimensão n (ou **sistema de coordenadas**) para M é uma família

$$\Phi = \{\varphi_\alpha\}_{\alpha \in \mathcal{A}}$$

de homeomorfismos $\varphi_\alpha : U_\alpha \rightarrow V_\alpha$ de um aberto $U_\alpha \subset \mathbb{R}^n$ sobre um aberto V_α de M para cada $\alpha \in \mathcal{A}$, satisfazendo as seguintes condições:

(i) Os abertos V_α cobrem M , isto é,

$$\bigcup_{\alpha \in \mathcal{A}} V_\alpha = M.$$

(ii) Para todos índices $\alpha, \beta \in \mathcal{A}$ tais que $V_{\alpha\beta} = V_\alpha \cap V_\beta \neq \emptyset$, as **funções de transição**

$$\begin{aligned}\varphi_{\alpha\beta} &= \varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}(V_{\alpha\beta}) \rightarrow \varphi_\beta^{-1}(V_{\alpha\beta}), \\ \varphi_{\beta\alpha} &= \varphi_\alpha^{-1} \circ \varphi_\beta : \varphi_\beta^{-1}(V_{\alpha\beta}) \rightarrow \varphi_\alpha^{-1}(V_{\alpha\beta}),\end{aligned}$$

são diferenciáveis de classe C^∞ .

Cada aplicação φ_α é chamada uma **carta** (ou uma **parametrização** ou um **sistema de coordenadas locais**) para uma vizinhança de M , denotada $(\varphi_\alpha, U_\alpha)$, e $V_\alpha = \varphi_\alpha(U_\alpha)$ é chamada uma **vizinhança coordenada**.

Se $p = \varphi_\alpha(x_1, \dots, x_n)$, então x_1, \dots, x_n são chamadas as **coordenadas locais** de p na carta φ_α .

Uma **estrutura diferenciável** para M é um atlas maximal.

Uma **variedade diferenciável** de dimensão n é um espaço topológico de Hausdorff com base enumerável munido de uma estrutura diferenciável. \square

Em outras palavras, uma variedade diferenciável é uma *variedade topológica* em que as mudanças de coordenadas de um sistema de coordenadas local para outro são diferenciáveis. Observe que o que definimos como variedade diferenciável é chamado **variedade suave** em outros lugares. Em Geometria Riemanniana, vários conceitos importantes necessitam derivadas de várias ordens, portanto costuma-se trabalhar com variedades suaves desde o início e não com variedades diferenciáveis de classe C^k , para não ter que especificar a todo momento o valor de k necessário para que certo conceito possa ser definido ou para que certo teorema faça sentido. Requerer que o atlas seja maximal é incluir no atlas todas as cartas locais que são **compatíveis** com o atlas, isto é, um atlas $\Phi = \{\varphi_\alpha\}_{\alpha \in \mathcal{A}}$ é maximal se sempre que $\varphi : U \rightarrow V$ é um homeomorfismo de um aberto $U \subset \mathbb{R}^n$ sobre um aberto V de M tal que $\varphi^{-1} \circ \varphi_\alpha$ e $\varphi_\alpha^{-1} \circ \varphi$ são diferenciáveis para todo $\alpha \in \mathcal{A}$, então $\varphi \in \Phi$; por definição, o atlas maximal é único. Esta é uma maneira mais simples de definir o conceito

de estrutura diferenciável, do que definir estruturas diferenciáveis como classes de equivalência de atlases compatíveis, ou seja, de atlases tais que as cartas de são todas compatíveis com as cartas do outro; uma função definida em uma variedade diferenciável é diferenciável (veremos a definição logo a seguir) com relação a dois atlases diferentes se e somente se eles são compatíveis, portanto atlases compatíveis definem a mesma estrutura diferenciável neste sentido. Se uma variedade topológica possui uma estrutura diferenciável (toda variedade topológica de dimensão $n \leq 3$ possui, mas para toda dimensão $n \geq 4$ existem variedades topológicas compactas que não possuem) ela possui uma quantidade não enumerável de estruturas diferenciáveis (veja [Lee 1], p. 30, Problem 1.6). Por outro lado, a estrutura diferenciável de uma variedade diferenciável é única, no sentido que todo atlas diferenciável está contido em um único atlas maximal (para uma demonstração, veja [Lee 1], p. 13, Proposition 1.17).

Surpreendentemente, a condição de ser de Hausdorff (que assegura, entre outras coisas, que conjuntos finitos de pontos são fechados e que limites de sequências convergentes são únicos) não é implicada pela definição (veja Exercício 8.44). A condição de possuir uma base enumerável garante a existência de partições da unidade. Por outro lado, dado um conjunto X , uma estrutura diferenciável sobre X determina uma topologia para X (veja Exercício 8.45). Observe que toda variedade diferenciável é localmente conexa por caminhos e que uma variedade diferenciável é conexa se e somente se ela é conexa por caminhos (Exercício 8.46). Além disso, toda variedade diferenciável é paracompacta, isto é, toda cobertura de uma variedade diferenciável por abertos admite uma subcobertura localmente finita; mas que isso, esta subcobertura pode ser tomada enumerável (veja [Lee 1] para demonstrações destas afirmações).

Quando nos referirmos a uma variedade diferenciável, assumimos que ela está munida de uma estrutura diferenciável. Denotaremos às vezes uma variedade diferenciável M de dimensão n por M^n quando for necessário especificar a dimensão da variedade. Também denotaremos uma carta do atlas por (φ, U) quando a sua imagem não for importante nas considerações.

8.2 Aplicações Diferenciáveis

8.2 Definição. Seja M^n uma variedade diferenciável. Dizemos que uma função $f : M \rightarrow \mathbb{R}^k$ é uma **função diferenciável** se para todo $p \in M$ existe uma carta (φ, U) de uma vizinhança de p tal que

$$f \circ \varphi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$$

é uma função diferenciável de classe C^∞ . \square

Observe que se $f \circ \varphi$ é diferenciável para uma carta (φ, U) de uma vizinhança de p , então para qualquer carta (ψ, V) de uma vizinhança de p temos que $f \circ \psi$ é diferenciável, pois

$$f \circ \psi = f \circ \varphi \circ (\varphi^{-1} \circ \psi)$$

e $\psi^{-1} \circ \varphi$ é um difeomorfismo. A função $f \circ \varphi$ é chamada uma **representação de f em coordenadas**. Frequentemente omitimos a carta φ quando trabalhamos com a representação de f em coordenadas e escrevemos

$$f(x_1, \dots, x_n)$$

ao invés de

$$(f \circ \varphi)(x_1, \dots, x_n).$$

8.3 Definição. Se M é uma variedade diferenciável, definimos o espaço vetorial

$$C^\infty(M) = \{f : M \rightarrow \mathbb{R} : f \text{ é diferenciável}\}.$$

\square

$C^\infty(M)$ é um espaço vetorial porque a combinação linear de funções diferenciáveis é uma função diferenciável.

8.4 Definição. Sejam M^m e N^n variedades diferenciáveis. Dizemos que uma aplicação $F : M \rightarrow N$ é uma **aplicação diferenciável** se para todo $p \in M$ existem cartas (φ, U) de uma vizinhança de p e (ψ, V) de uma vizinhança de $F(p)$ com $F(\varphi(U)) \subset \psi(V)$ tais que

$$\psi^{-1} \circ F \circ \varphi : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$$

é uma aplicação diferenciável de classe C^∞ . \square

Novamente observamos que se $\psi^{-1} \circ F \circ \varphi$ é diferenciável para as cartas $(\varphi, U), (\psi, V)$, então para quaisquer cartas $(\tilde{\varphi}, \tilde{U})$ de uma vizinhança de p e $(\tilde{\psi}, \tilde{V})$ de uma vizinhança de $F(p)$ tais que $F(\varphi(\tilde{U})) \subset \psi(\tilde{V})$ temos que $\tilde{\psi}^{-1} \circ F \circ \tilde{\varphi}$ é diferenciável, pois

$$\tilde{\psi}^{-1} \circ F \circ \tilde{\varphi} = (\tilde{\psi}^{-1} \circ \psi) \circ \psi^{-1} \circ F \circ \varphi \circ (\varphi^{-1} \circ \tilde{\varphi})$$

e $\varphi^{-1} \circ \tilde{\varphi}, \tilde{\psi}^{-1} \circ \psi$ são difeomorfismos. A aplicação $\psi^{-1} \circ F \circ \varphi$ é uma **representação de F em coordenadas**.

Ressaltamos de novo que o que definimos como função diferenciável e aplicação diferenciável são chamadas **função suave** e **aplicação suave** em outros lugares.

8.5 Definição. Dizemos que uma aplicação diferenciável $F : M \rightarrow N$ é um **difeomorfismo** se F é um homeomorfismo e F^{-1} também é diferenciável.

Se existir um difeomorfismo entre duas variedades diferenciáveis M e N , dizemos que elas são **difeomórficas**. \square

Se duas variedades diferenciáveis são difeomórficas, em particular elas possuem a mesma dimensão.

O conceito de difeomorfismo leva naturalmente a perguntar se, dada uma variedade diferenciável, estruturas diferenciáveis diferentes sobre ela definem sempre duas variedades diferenciáveis difeomórfas ou não. A resposta é que para variedades diferenciáveis de dimensão $n \leq 3$ existe apenas uma estrutura diferenciável a menos de difeomorfismo, enquanto que mesmo \mathbb{R}^4 tem um número não enumerável de estruturas diferenciáveis não difeomórfas; sabe-se que esferas de dimensão até $n = 20$, $n \neq 4$, possuem um número finito de estruturas diferenciáveis a menos de difeomorfismo e este número é conhecido (veja referências em Wikipedia e [Lee 1], p. 40). Quantas estruturas diferenciáveis difeomórfas \mathbb{S}^4 possui (ou mesmo se este número é maior que 1 ou finito) é uma questão em aberto, a *conjectura de Poincaré generalizada*.

Uma das aplicações diferenciáveis mais importantes entre variedades são as *curvas diferenciáveis*:

8.6 Definição. Uma **curva diferenciável** em uma variedade diferenciável M é uma aplicação diferenciável $\alpha : I \rightarrow M$ onde $I \subset \mathbb{R}$ é um intervalo. \square

8.2.1 Partições da Unidade

8.7 Definição. Seja $\mathcal{V} = \{V_\alpha\}_{\alpha \in \mathcal{A}}$ uma cobertura por abertos de uma variedade diferenciável M . Uma **partição da unidade subordinada a \mathcal{V}** é uma coleção $\{\rho_\alpha\}_{\alpha \in \mathcal{A}}$ de funções diferenciáveis $\rho_\alpha : M \rightarrow \mathbb{R}$ tais que

$$(i) 0 \leq \rho_\alpha \leq 1.$$

$$(ii) \text{supp } \rho_\alpha \subset V_\alpha.$$

(iii) $\{\text{supp } \rho_\alpha\}_{\alpha \in \mathcal{A}}$ é localmente finita (todo ponto em M possui uma vizinhança que intersecta apenas um número finito destes suportes).

$$(iv) \sum_{\alpha \in \mathcal{A}} \rho_\alpha = 1.$$

8.8 Teorema (Existência de Partições da Unidade). *Toda cobertura por abertos de uma variedade diferenciável possui uma partição da unidade subordinada.*

Prova: Veja [Lee 2], p. 43, Theorem 2.23. \blacksquare

8.9 Corolário (Existência de Funções Bump). *Dados um fechado A e um aberto $V \supset A$ em uma variedade diferenciável M , existe uma função diferenciável $f : M \rightarrow \mathbb{R}$ tal que*

- (i) $0 \leq f \leq 1$.
- (ii) $f \equiv 1$ em A .
- (iii) $\text{supp } f \subset V$.

8.10 Corolário (Lema de Extensão). *Dados um fechado A , um aberto $V \supset A$ em uma variedade diferenciável M e uma função diferenciável $f : A \rightarrow \mathbb{R}^k$ existe uma extensão diferenciável $\tilde{f} : M \rightarrow \mathbb{R}^k$ de f tal que $\text{supp } \tilde{f} \subset V$.*

Em particular, se (φ, U) é uma carta local e $V \subset \subset \varphi(U)$, qualquer função diferenciável $f : V \rightarrow \mathbb{R}^k$ pode ser estendida a uma função diferenciável $\tilde{f} : M \rightarrow \mathbb{R}^k$ com $\text{supp } \tilde{f} \subset \varphi(U)$.

Prova: Para a primeira parte veja [Lee 2], p. 45, Lemma 2.26. A segunda parte segue imediatamente da primeira, já que $\bar{V} \subset U$ é fechado. ■

8.3 Vetores Tangentes

Consideremos agora a questão de como definir a noção de vetor tangente a um ponto em uma variedade diferenciável. Esta noção não é óbvia, já que uma variedade é um espaço abstrato que não se encontra em princípio imerso em um espaço ambiente, ou seja, em um espaço euclidiano, onde operações diferenciais e vetoriais são naturais. Portanto, precisamos procurar uma característica de vetores tangentes em espaços euclidianos que independa do espaço ambiente. Faremos isso em duas etapas, aumentando em abstração até chegar a uma definição que provará ser extremamente conveniente de trabalhar.

8.3.1 Vetores Tangentes a Curvas

No que se segue, denotaremos as derivadas parciais de funções reais f de várias variáveis reais por

$$\frac{\partial f}{\partial x^i} \quad \text{ou} \quad \partial_i f$$

conforme for mais conveniente.

Quando $\alpha : I \rightarrow \mathbb{R}^n$ é uma curva diferenciável em um espaço euclidiano, com $\alpha(t_0) = p$ e $\alpha'(t_0) = v$, escrevendo em coordenadas

$$\alpha(t) = (x^1(t), \dots, x^n(t)),$$

temos que

$$\alpha'(t) = \left(\frac{dx^1}{dt}(t), \dots, \frac{dx^n}{dt}(t) \right),$$

e em particular

$$v = \alpha'(t_0) = \left(\frac{dx^1}{dt}(t_0), \dots, \frac{dx^n}{dt}(t_0) \right).$$

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função diferenciável em p , então a derivada direcional de f em p na direção de v é dada pela regra da cadeia por

$$(f \circ \alpha)'(t_0) = df_{\alpha(t_0)} \alpha'(t_0) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(p) \frac{dx^i}{dt}(t_0) = \left[\sum_{i=1}^n \frac{dx^i}{dt}(t_0) \frac{\partial}{\partial x^i} \right] f(p),$$

o que significa que a derivada direcional na direção de v pode ser vista como um operador linear sobre funções diferenciáveis que depende apenas do vetor v .

8.11 Definição (Preliminar). Seja $\alpha : I \rightarrow M$ uma curva diferenciável com $\alpha(t_0) = p$. O **vetor tangente à curva** α em p é uma função $v_p : C^\infty(M) \rightarrow \mathbb{R}$ definida por

$$v_p(f) = (f \circ \alpha)'(t_0).$$

Um **vetor tangente à variedade** M em p é qualquer vetor tangente à uma curva diferenciável passando por p . \square

Ou seja, cada curva diferenciável em M passando por p dá origem a um vetor tangente em p . É claro que curvas diferenciáveis diferentes $\alpha, \beta : I \rightarrow M$ com $\alpha(t_0) = \beta(s_0) = p$ podem dar origem ao mesmo vetor tangente: basta que

$$(f \circ \alpha)'(t_0) = (f \circ \beta)'(s_0)$$

para todo f .

8.12 Proposição. Um vetor tangente $v_p : C^\infty(M) \rightarrow \mathbb{R}$ é um funcional linear e o conjunto dos vetores tangentes a uma variedade em um ponto formam um espaço vetorial real n -dimensional.

Prova: Seja (φ, U) uma carta de uma vizinhança de p com $\varphi(x_0) = p$. Sejam $\alpha : I \rightarrow M$ uma curva diferenciável com $\alpha(t_0) = p$ e v_p o vetor tangente a α em p . Dado $f \in C^\infty(M)$, temos

$$\begin{aligned} v_p(f) &= (f \circ \alpha)'(t_0) \\ &= (f \circ \varphi \circ \varphi^{-1} \circ \alpha)'(t_0) \\ &= d(f \circ \varphi)_{x_0} (\varphi^{-1} \circ \alpha)'(t_0) \\ &= \sum_{i=1}^n \frac{\partial(f \circ \varphi)}{\partial x^i}(x_0) \frac{dx^i}{dt}(t_0), \end{aligned}$$

onde denotamos $(\varphi^{-1} \circ \alpha)(t) = (x^1(t), \dots, x^n(t))$ em coordenadas locais. Segue que para todos $f, g \in C^\infty(M)$ e para todos $a, b \in \mathbb{R}$ temos

$$\begin{aligned} v_p(af + bg) &= \sum_{i=1}^n \frac{\partial([af + bg] \circ \varphi)}{\partial x^i}(x_0) \frac{dx^i}{dt}(t_0) \\ &= a \sum_{i=1}^n \frac{\partial(f \circ \varphi)}{\partial x^i}(x_0) \frac{dx^i}{dt}(t_0) + b \sum_{i=1}^n \frac{\partial(g \circ \varphi)}{\partial x^i}(x_0) \frac{dx^i}{dt}(t_0) \\ &= av_p(f) + bv_p(g), \end{aligned}$$

de modo que $v_p : C^\infty(M) \rightarrow \mathbb{R}$ é um funcional linear.

Para mostrar que o conjunto dos vetores tangentes a M em p formam um espaço vetorial e que este tem dimensão n , mostraremos que todo vetor tangente é a combinação linear de n vetores tangentes linearmente independentes $\partial_1|_p, \dots, \partial_n|_p$ a serem definidos e que, além disso, qualquer combinação linear dos vetores tangentes $\partial_1|_p, \dots, \partial_n|_p$ é um vetor tangente (embora combinações lineares de funcionais lineares sejam sempre funcionais lineares, nada garante em princípio que um tal funcional linear é um vetor tangente; de fato, depois que provarmos que o espaço vetorial dos vetores tangentes tem dimensão n , segue que ele é um subespaço vetorial próprio do espaço vetorial dos funcionais lineares de $C^\infty(M)$, pois este tem dimensão infinita).

De fato, reescreva a expressão obtida acima para $v_p(f)$ na forma

$$v_p(f) = \sum_{i=1}^n \frac{dx^i}{dt}(t_0) \frac{\partial(f \circ \varphi)}{\partial x^i}(x_0).$$

Denotando por $\mathcal{B} = \{e_1, \dots, e_n\}$ a base canônica de \mathbb{R}^n , seja α_i a curva diferenciável $\alpha_i : I_i \rightarrow M$ definida por

$$\alpha_i(t) = \varphi(x_0 + te_i),$$

onde I_i é um intervalo aberto em torno de t_0 tal que $x_0 + te_i \in U$ para todo $t \in I_i$ e denote por $\partial_i|_p$ o vetor tangente à curva α_i em p . Como

$$\begin{aligned}\frac{\partial(f \circ \varphi)}{\partial x^i}(x_0) &= \lim_{t \rightarrow 0} \frac{(f \circ \varphi)(x_0 + te_i) - (f \circ \varphi)(x_0)}{t} \\ &= (f \circ \alpha_i)'(t_0) \\ &= \partial_i|_p(f),\end{aligned}$$

segue que

$$v_p = \sum_{i=1}^n \frac{dx^i}{dt}(t_0) \partial_i|_p.$$

Reciprocamente, se v é o funcional linear

$$v_p = \sum_{i=1}^n c_i \partial_i|_p,$$

então v é o vetor tangente à curva α em p definida por

$$\alpha(t) = \varphi\left(x_0 + t \left(\sum_{i=1}^n c_i e_i\right)\right),$$

pois, pela regra da cadeia como vimos no início da demonstração,

$$(f \circ \alpha)'(0) = \sum_{i=1}^n \frac{dx^i}{dt}(0) \frac{\partial(f \circ \varphi)}{\partial x^i}(x_0) = \sum_{i=1}^n c_i \partial_i|_p(f) = v(f).$$

Finalmente, se

$$\sum_{i=1}^n c_i \partial_i|_p = 0,$$

então

$$\sum_{i=1}^n c_i \frac{\partial(f \circ \varphi)}{\partial x^i}(x_0) = 0$$

para todo $f \in C^\infty(M)$. Definindo para cada j

$$f_j(x^1, \dots, x^n) = x^j$$

em um aberto $U_0 \subset \subset U$, segue que f_j é diferenciável em U_0 e pelo Corolário 8.10 podemos estender f_j a uma função diferenciável $\tilde{f}_j \in C^\infty(M)$. Como

$$\frac{\partial(\tilde{f}_j \circ \varphi)}{\partial x^i}(x_0) = \frac{\partial(f_j \circ \varphi)}{\partial x^i}(x_0) = \delta_{ij},$$

escolhendo $f = f_j$ obtemos $c_j = 0$ para todo j . ■

8.13 Proposição (Regra do Produto). *O vetor tangente $v_p : C^\infty(M) \rightarrow \mathbb{R}$ satisfaz a propriedade*

$$v_p(fg) = v_p(f)g(p) + f(p)v_p(g).$$

Prova: Seja v_p o vetor tangente à curva α em p . Então

$$\begin{aligned} v_p(fg) &= ((fg) \circ \alpha)'(t_0) = [(f \circ \alpha)(g \circ \alpha)]'(t_0) \\ &= (f \circ \alpha)'(t_0)(g \circ \alpha)(t_0) + (f \circ \alpha)(t_0)(g \circ \alpha)'(t_0) \\ &= v_p(f)g(p) + f(p)v_p(g). \end{aligned}$$

■

8.3.2 Vetores Tangentes como Derivações

8.14 Definição. Seja M uma variedade diferenciável. Um **vetor tangente** a M em p é um funcional linear $v_p : C^\infty(M) \rightarrow \mathbb{R}$ que também é uma **derivação** em p , isto é, ele satisfaz a regra do produto

$$v_p(fg) = v_p(f)g(p) + f(p)v_p(g).$$

□

Note que nesta definição o conjunto dos vetores tangentes a um ponto $p \in M$ forma naturalmente um espaço vetorial real, pois é um subespaço do espaço vetorial dos funcionais lineares em $C^\infty(M)$:

$$\begin{aligned} (\alpha v_p + \beta w_p)(fg) &= \alpha v_p(fg) + \beta w_p(fg) \\ &= \alpha v_p(f)g(p) + \alpha f(p)v_p(g) + \beta w_p(f)g(p) + \beta f(p)w_p(g) \\ &= [(\alpha v_p + \beta w_p)(f)]g(p) + f(p)[(\alpha v_p + \beta w_p)(g)]. \end{aligned}$$

Mas a dimensão deste espaço não é imediatamente óbvia. Além disso, não é claro que todo vetor tangente segundo esta definição é um vetor tangente segundo a definição anterior. Embora seja consequência das Proposições 8.12 e 8.13 que vetores tangentes a curvas são funcionais lineares em $C^\infty(M)$ que são derivações, logo o espaço vetorial dos vetores tangentes segundo a definição anterior é um subespaço vetorial do espaço vetorial dos vetores tangentes segundo a nova definição, ainda não sabemos que todo vetor tangente a M em p é uma derivação. Isso provará ser verdade quando provarmos que o espaço vetorial dos funcionais lineares em $C^\infty(M)$ que são derivações também tem dimensão n (Proposição 8.22). Em outras palavras, as duas definições não são apenas equivalentes, mas de fato definem o mesmo conceito de vetor tangente. De agora em diante, utilizaremos a Definição 8.14 para vetor tangente.

8.15 Proposição. Qualquer vetor tangente $v_p : C^\infty(M) \rightarrow \mathbb{R}$ satisfaz as seguintes propriedades:

- (i) Se f é uma função constante, então $v_p(f) = 0$.
- (ii) Se $f(p) = g(p) = 0$, então $v_p(fg) = 0$.

Prova: Ambas as propriedades seguem imediatamente da regra do produto. (i) Como v_p é linear, basta provar para a função constante $f \equiv 1$. Pela regra do produto,

$$v_p(f) = v_p(f)f(p) + f(p)v_p(f) = 2v_p(f),$$

logo $v_p(f) = 0$. (ii) Pela regra do produto, temos

$$v_p(fg) = v_p(f)g(p) + f(p)v_p(g) = v_p(f)0 + 0v_p(g) = 0 + 0 = 0.$$

■

Apesar dos vetores tangentes (derivações) estarem definidas no espaço global $C^\infty(M)$, o próximo resultado mostra que a sua atuação é local.

8.16 Proposição. Seja $v_p : C^\infty(M) \rightarrow \mathbb{R}$ um vetor tangente. Se $f, g \in C^\infty(M)$ coincidem em uma vizinhança de p , então $v_p(f) = v_p(g)$.

Prova: Seja $h = f - g$, de modo que $h \in C^\infty(M)$ e $h = 0$ em uma vizinhança de p . Seja $\rho \in C^\infty(M)$ uma função cujo suporte está contida em $M \setminus \{p\}$ e que é igual a 1 no suporte de h . Em particular, como $\rho = 1$ onde h é não nula, segue que $\rho h = h$. Daí, $v_p(h) = v_p(\rho h) = 0$ pela propriedade (ii) da Proposição 8.15; o resultado segue agora por linearidade. ■

8.17 Definição. Seja M uma variedade diferenciável. O espaço vetorial dos vetores tangentes a um ponto $p \in M$ é chamado o **espaço tangente** a M em p e denotado $T_p M$. □

8.3.3 Diferencial de uma Aplicação Diferenciável

Para definir a diferencial (derivada) de uma aplicação diferenciável, usaremos a definição de vetores tangentes como derivações:

8.18 Definição. Sejam M e N variedades diferenciáveis e $F : M \rightarrow N$ uma aplicação diferenciável em $p \in M$. A **diferencial** de F em p é aplicação linear

$$dF_p : T_p M \rightarrow T_{F(p)} N$$

definida por

$$[dF_p(v_p)](f) = v_p(f \circ F)$$

para todo $f \in C^\infty(N)$. □

Note que como $f \in C^\infty(N)$ e F é de classe C^∞ , $f \circ F \in C^\infty(M)$. $dF_p(v)$ é uma derivação em $F(p)$ porque

$$\begin{aligned} dF_p(v_p)(fg) &= v_p((fg) \circ F) = v_p((f \circ F)(g \circ F)) \\ &= v_p(f \circ F)g(F(p)) + f(F(p))v_p(g \circ F) \\ &= [dF_p(v_p)(f)]g(F(p)) + f(F(p))[dF_p(v_p)(g)]. \end{aligned}$$

Além disso, dF_p é uma aplicação linear porque v_p é um funcional linear.

8.19 Proposição (Regra da Cadeia). Sejam M, N, P variedades diferenciáveis e $F : M \rightarrow N, G : N \rightarrow P$ aplicações diferenciáveis. Então $G \circ F : M \rightarrow P$ é uma aplicação diferenciável e

$$d(G \circ F)_p = dG_{F(p)} \circ dF_p.$$

Prova: Provaremos a segunda parte; a primeira parte é deixada como exercício. Por definição, para todo $f \in C^\infty(P)$

$$\begin{aligned} [d(G \circ F)_p(v_p)](f) &= v_p(f \circ (G \circ F)) = v_p((f \circ G) \circ F) = dF_p(v_p)(f \circ G) \\ &= [dG_{F(p)}(dF_p(v_p))](f). \end{aligned}$$

■

8.20 Corolário. Se $F : M \rightarrow N$ é um difeomorfismo, então dF_p é um isomorfismo para cada $p \in M$ e $d(F^{-1})_{F(p)} = (dF_p)^{-1}$.

8.21 Lema. Seja M uma variedade diferenciável. Se V é um aberto de M e $i : V \rightarrow M$ é a inclusão, então di_p é um isomorfismo para todo $p \in M$.

Prova: Para provar que $di_p : T_p V \rightarrow T_p M$ é injetivo, suponha que $di_p(v_p) = 0$ para $v_p \in T_p V$. Seja $W \subset\subset V$ uma vizinhança de p . Se $f \in C^\infty(V)$ é uma função diferenciável arbitrária, considere uma extensão $\tilde{f} \in C^\infty(M)$ tal que $\tilde{f} = f$ em \overline{W} . Como f e \tilde{f} coincidem na vizinhança W de p , segue da Proposição 8.16 que

$$v_p(f) = v_p(\tilde{f}|_V) = v_p(\tilde{f} \circ i) = di_p(v_p)(\tilde{f}) = 0.$$

Como $f \in C^\infty(V)$ é arbitrária, isso prova que $v_p = 0$, logo di_p é injetiva.

Para provar que di_p é sobrejetiva, seja $w_p \in T_p M$ um vetor tangente qualquer. Defina uma função $v : C^\infty(V) \rightarrow \mathbb{R}$ por

$$v(f) = w_p(\tilde{f})$$

onde \tilde{f} é uma extensão definida como no início da demonstração. Pela Proposição 8.16, o valor de $w(\tilde{f})$ independe da escolha de \tilde{f} , logo v está bem definida. É fácil ver que v é uma derivação. Para todo $g \in C^\infty(M)$ temos

$$di_p(v)(g) = v(g \circ i) = w_p(\widetilde{g \circ i}) = w_p(g)$$

onde a última igualdade segue do fato que $\widetilde{g \circ i}$ e g coincidem em W . Portanto, $di_p(v) = w_p$. ■

8.22 Proposição. *Se M é uma variedade diferenciável de dimensão n , então $T_p M$ é um espaço vetorial de dimensão n para todo $p \in M$.*

Prova: Seja $\varphi : U \rightarrow V$ uma carta para uma vizinhança $V = \varphi(U) \subset M$ de $p = \varphi(x)$. Como φ é um difeomorfismo, segue que $d\varphi_x : \mathbb{R}^n \rightarrow T_p V$ é um isomorfismo. Como $T_p V$ e $T_p M$ são isomórfos pelo lema, segue o resultado. ■

Conforme a discussão que se segue à Definição 8.14, concluímos que para todo vetor tangente $v_p \in T_p M$ existe uma curva diferenciável $\alpha : I \rightarrow M$ com $\alpha(t_0) = p$ tal que

$$v_p(f) = (f \circ \alpha)'(t_0).$$

8.23 Definição. Seja M uma variedade diferenciável e $\varphi : U \rightarrow M$ uma carta de uma vizinhança de um ponto $p \in M$. A base obtida na demonstração da Proposição 8.12 será chamada a **base coordenada** do espaço tangente $T_p M$ associada à carta φ e denotada por

$$\partial_1|_p, \dots, \partial_n|_p$$

ou por

$$\left. \frac{\partial}{\partial x^1} \right|_p, \dots, \left. \frac{\partial}{\partial x^n} \right|_p$$

quando for conveniente ou necessário explicitar as coordenadas da carta. □

8.4 Coordenadas

8.4.1 Diferencial em Coordenadas

Seja $\mathcal{B} = \{e_1, \dots, e_n\}$ a base canônica de \mathbb{R}^n . Se $\varphi : U \subset \mathbb{R}^n \rightarrow V$ é uma carta para uma vizinhança coordenada V de $p = \varphi(x) \in M$, a base coordenada associada à φ é também dada por

$$\left. \frac{\partial}{\partial x^i} \right|_p = d\varphi_x(e_i).$$

De fato, se $f \in C^\infty(M)$, então

$$\left. \frac{\partial}{\partial x^i} \right|_p(f) = \left. \frac{\partial(f \circ \varphi)}{\partial x^i} \right|_x(x) = e_i(f \circ \varphi) = d\varphi_x(e_i)(f).$$

Assim, podemos *definir*

8.24 Definição.

$$\frac{\partial f}{\partial x^i}(p) = \left. \frac{\partial}{\partial x^i} \right|_p (f) = \frac{\partial(f \circ \varphi)}{\partial x^i}(x).$$

Vamos ver agora como é a diferencial de uma aplicação diferenciável em coordenadas.

Primeiro recordaremos o caso em que as variedades são espaços euclidianos. Denote por $\mathcal{B}^m = \{e_1, \dots, e_m\}$ e $\mathcal{B}^n = \{f_1, \dots, f_n\}$ as bases canônicas de \mathbb{R}^m e \mathbb{R}^n , respectivamente. Observe que se $F : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ é uma aplicação diferenciável, então $dF_x : \mathbb{R}^m \rightarrow \mathbb{R}^n$ é a derivada usual para cada $x \in U$ e pela regra da cadeia

$$dF_x(e_i)(f) = e_i(f \circ F) = \frac{\partial(f \circ F)}{\partial x^i}(x) = \sum_{j=1}^m \frac{\partial f}{\partial x^j}(F(x)) \frac{\partial F^j}{\partial x^i}(x) = \sum_{j=1}^m \frac{\partial F^j}{\partial x^i}(x) f_j(f),$$

ou seja,

$$dF_x(e_i) = \sum_{j=1}^n \frac{\partial F^j}{\partial x^i}(x) f_j.$$

Assim, a matriz da diferencial dF_x em relação às bases $\mathcal{B}^m, \mathcal{B}^n$ é o jacobiano

$$J = \begin{bmatrix} \frac{\partial F^1}{\partial x^1} & \cdots & \frac{\partial F^1}{\partial x^m} \\ \vdots & & \vdots \\ \frac{\partial F^n}{\partial x^1} & \cdots & \frac{\partial F^n}{\partial x^m} \end{bmatrix} =: [dF_x]_{\mathcal{B}^m, \mathcal{B}^n}.$$

Ou seja, se $v = \sum_{i=1}^m v^i e_i$, então

$$dF_x(v) = \sum_{i=1}^m v^i dF_x(e_i) = \sum_{j=1}^n \left[\sum_{i=1}^m v^i \frac{\partial F^j}{\partial x^i}(x) \right] f_j,$$

isto é,

$$[dF_x(v)]_{\mathcal{B}^n} = \begin{bmatrix} \frac{\partial F^1}{\partial x^1} & \cdots & \frac{\partial F^1}{\partial x^m} \\ \vdots & & \vdots \\ \frac{\partial F^n}{\partial x^1} & \cdots & \frac{\partial F^n}{\partial x^m} \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^m \end{bmatrix} = J[v]_{\mathcal{B}^m}.$$

No caso geral, se $F : M^m \rightarrow N^n$ é uma aplicação diferenciável, sejam $\varphi : U \subset \mathbb{R}^m \rightarrow \varphi(U)$, $\psi : V \subset \mathbb{R}^n \rightarrow \psi(V)$ cartas de vizinhanças de $p = \varphi(x)$ em M e de $F(p) = \psi(y)$ em N , respectivamente, de modo que

$$\tilde{F} = \psi^{-1} \circ F \circ \varphi : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n.$$

Escrevendo

$$F \circ \varphi = \psi \circ \tilde{F}$$

temos

$$\begin{aligned} dF_p \left(\frac{\partial}{\partial x^i} \Big|_p \right) &= dF_p[d\varphi_x(e_i)] = d\psi_y \left[d\tilde{F}_x(e_i) \right] = d\psi_y \left[\sum_{j=1}^n \frac{\partial \tilde{F}^j}{\partial x^i}(x) f_j \right] = \sum_{j=1}^n \frac{\partial \tilde{F}^j}{\partial x^i}(x) d\psi_y(f_j) \\ &= \sum_{j=1}^n \frac{\partial \tilde{F}^j}{\partial x^i}(x) \left(\frac{\partial}{\partial y^j} \Big|_{F(p)} \right). \end{aligned}$$

Portanto, se

$$\begin{aligned}\mathcal{B}_p &= \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^m} \Big|_p \right\}, \\ \mathcal{B}_{F(p)} &= \left\{ \frac{\partial}{\partial y^1} \Big|_{F(p)}, \dots, \frac{\partial}{\partial y^n} \Big|_{F(p)} \right\},\end{aligned}$$

são as bases coordenadas de $T_p M$ e $T_{F(p)} N$, respectivamente, então a matriz que representa a diferencial dF_p em relação a estas bases é

$$[dF_p]_{\mathcal{B}_p, \mathcal{B}_{F(p)}} = \begin{bmatrix} \frac{\partial \tilde{F}^1}{\partial x^1} & \cdots & \frac{\partial \tilde{F}^1}{\partial x^m} \\ \vdots & & \vdots \\ \frac{\partial \tilde{F}^n}{\partial x^1} & \cdots & \frac{\partial \tilde{F}^n}{\partial x^m} \end{bmatrix}.$$

8.5 Fibrado Tangente

8.25 Definição. Seja M uma variedade diferenciável de dimensão n com um atlas $\Phi = \{\varphi_\alpha : U_\alpha \rightarrow M\}_{\alpha \in \mathcal{A}}$ de classe C^k . O **fibrado tangente** de M é a variedade diferenciável de dimensão $2n$ e classe C^{k-1}

$$TM = \{(p, v) : p \in M \text{ e } v \in T_p M\}$$

com um atlas

$$\Psi = \{\psi_\alpha : U_\alpha \times \mathbb{R}^n \rightarrow TM\}_{\alpha \in \mathcal{A}}$$

definido por

$$\psi_\alpha(x, v_1, \dots, v_n) = \left(\varphi_\alpha(x), \sum_{i=1}^n v_i \partial_i(x) \right).$$

□

Na definição acima, o próprio atlas Ψ define a topologia necessária em TM (Exercício 8.45).

8.6 Fibrados Vetoriais

8.26 Definição. Um **fibrado vetorial** de ordem k sobre uma variedade diferenciável M é uma variedade diferenciável E juntamente com uma aplicação sobrejetiva diferenciável $\pi : E \rightarrow M$ tal que

- (i) cada **fibra** $E_p = \pi^{-1}(p)$ de E sobre p é um espaço vetorial de dimensão k ;
- (ii) para cada $p \in M$ existe uma vizinhança U de p e um difeomorfismo $\varphi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}^k$, chamado uma **trivialização local** de E , tal que o diagrama seguinte é comutativo:

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\varphi} & U \times \mathbb{R}^k \\ \pi \downarrow & \swarrow \pi_1 & \\ U & & \end{array}$$

($\pi_1 : U \times \mathbb{R}^k \rightarrow U$ é a projeção na primeira variável) e tal que $\varphi|_{E_p} : E_p \rightarrow \{p\} \times \mathbb{R}^k$ é um isomorfismo de espaços vetoriais.

A variedade E é chamada o **espaço total** do fibrado, M a **base** do fibrado e π a sua projeção. □

Frequentemente identificamos o espaço total com o fibrado e dizemos simplesmente que E é o fibrado vetorial sobre M . Fibrados tangentes são exemplos de fibrados vetoriais.

8.27 Definição. Seja E um fibrado vetorial de dimensão k sobre M . Uma **seção** de E é uma aplicação $s : M \rightarrow E$ tal que $\pi \circ s = Id|_M$. \square

Em outras palavras, $s : M \rightarrow E$ é uma seção se e somente se $s(p) \in E_p$ para todo $p \in M$.

8.7 Campos Vetoriais

Considere $\pi : TM \rightarrow M$ a projeção canônica do fibrado tangente de M sobre M , isto é, $\pi(p, v) = p$ para todo $v \in T_p M$.

8.28 Definição. Seja M uma variedade diferenciável. Um **campo vetorial diferenciável** em M é uma aplicação diferenciável $X : M \rightarrow TM$ tal que $\pi \circ X = id_M$. \square

Podemos pensar em campos vetoriais como aplicações que associam a cada ponto $p \in M$ um vetor tangente $X(p) \in T_p M$; frequentemente, denotaremos o vetor tangente $X(p)$ simplesmente por X_p . Em termos de coordenadas locais, se $\mathcal{B} = \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$ é a base do espaço tangente $T_p M$ associada à uma carta $\varphi : U \rightarrow M$ para pontos $p \in \varphi(U)$, então

$$X_p = \sum_{i=1}^n X^i(p) \frac{\partial}{\partial x^i} \Big|_p$$

e o campo vetorial X é diferenciável em $\varphi(U)$ se e somente se as funções coordenadas X^1, \dots, X^n são diferenciáveis.

Outra forma de ver um campo vetorial diferenciável em M é como uma aplicação que associa a cada função $f \in C^\infty(M)$ uma função $Xf \in C^\infty(M)$ através da expressão

$$(Xf)(p) = X_p f$$

onde $X_p : C^\infty(M) \rightarrow \mathbb{R}$ é um vetor tangente.

8.29 Definição. Seja M uma variedade diferenciável. Um **campo vetorial diferenciável** em M é uma aplicação $X : C^\infty(M) \rightarrow C^\infty(M)$ que satisfaz as seguintes propriedades

(i) X é linear:

$$X(\alpha f + \beta g) = \alpha Xf + \beta Xg$$

para todos $\alpha, \beta \in \mathbb{R}$ e $f, g \in C^\infty(M)$.

(ii) X satisfaz a regra do produto:

$$X(fg) = (Xf)g + f(Xg)$$

para todos $f, g \in C^\infty(M)$. \square

As duas definições são equivalentes. Usando a última definição, podemos definir combinações lineares de campos vetoriais de forma natural.

8.30 Notação. Seja M uma variedade diferenciável. O espaço vetorial dos campos vetoriais diferenciáveis em M é denotado por $\mathcal{T}(M)$. \square

8.31 Proposição. Sejam M e N variedades diferenciáveis e $F : M \rightarrow N$ uma aplicação diferenciável. Se $X \in \mathcal{T}(M)$ e $Y \in \mathcal{T}(N)$ são campos vetoriais tais que

$$Y_{F(p)} = dF_p(X_p)$$

para todo $p \in M$, então

$$X(f \circ F) = (Yf) \circ F$$

para todo $f \in C^\infty(N)$.

Prova: Pela definição de diferencial,

$$[dF_p(X_p)](f) = X_p(f \circ F),$$

logo,

$$[X(f \circ F)](p) = X_p(f \circ F) = [dF_p(X_p)](f) = Y_{F(p)}f = (Yf)(F(p)).$$

■

8.32 Definição. Sejam M e N variedades diferenciáveis e $F : M \rightarrow N$ um difeomorfismo. Definimos a aplicação **pushforward**

$$F_* : \mathcal{T}(M) \rightarrow \mathcal{T}(N)$$

por

$$(F_*X)_q = dF_p(X_p)$$

onde $q = F(p)$. □

Equivalentemente,

$$(F_*X)_q = dF_{F^{-1}(q)}(X_{F^{-1}(q)}).$$

8.33 Proposição. Sejam M e N variedades diferenciáveis e $F : M \rightarrow N$ um difeomorfismo. Considere $\mathcal{T}(M)$ e $\mathcal{T}(N)$ como módulos sobre os anéis $C^\infty(M)$ e $C^\infty(N)$, respectivamente. Então o operador pushforward F_* é linear no seguinte sentido:

$$F_*(fX + gY) = (f \circ F^{-1})F_*X + (g \circ F^{-1})F_*Y$$

para todos $X, Y \in \mathcal{T}(M)$ e para todas $f, g \in C^\infty(M)$. Além disso, para toda $f \in C^\infty(N)$ vale

$$[(F_*X)f] \circ F = X(f \circ F)$$

ou, equivalentemente,

$$(F_*X)f = X(f \circ F) \circ F^{-1}.$$

Prova:

$$\begin{array}{ccc} M & \xrightarrow{F} & N \\ \downarrow f & \swarrow f \circ F^{-1} & \\ \mathbb{R} & & \end{array}$$

Primeiro provamos a linearidade de F_* . No que se segue, $q = F(p)$. Temos

$$\begin{aligned} [F_*(X + Y)]_q &= dF_p(X_p + Y_p) \\ &= dF_p(X_p) + dF_p(Y_p) \\ &= (F_*X)_q + (F_*Y)_q \end{aligned}$$

e

$$\begin{aligned}[F_*(fX)]_q &= dF_p \left((fX)_p \right) \\ &= dF_p (f(p) X_p) \\ &= f(p) dF_p (X_p) \\ &= (f \circ F^{-1})(q) (F_* X)_q.\end{aligned}$$

A última afirmativa segue imediatamente da Proposição 8.31, já que $F_* X$ é exatamente o campo Y do enunciado daquela proposição. ■

8.34 Teorema. *Seja $X \in \mathcal{T}(M)$ um campo diferenciável. Dado $p \in M$, existe uma vizinhança V de p em M , $\delta > 0$ e uma aplicação diferenciável*

$$\varphi : (-\delta, \delta) \times V \longrightarrow M$$

tais que $\varphi_q(t) = \varphi(t, q)$ é a única curva diferenciável em M que satisfaz

$$\begin{cases} \frac{d\varphi}{dt}(t, q) = X_{\varphi(t, q)} & \text{para todos } t \in (-\delta, \delta), q \in V, \\ \varphi(0, q) = q. \end{cases}$$

Além disso, para cada t fixado, $\varphi_t = \varphi(t, \cdot)$ é um difeomorfismo e o fluxo é um grupo aditivo a um parâmetro, isto é,

$$\begin{aligned}\varphi_0 &= \text{id}, \\ \varphi_{t+s} &= \varphi_t \varphi_s.\end{aligned}$$

Prova: Veja [Lee 1], Chapter 9, p. 209. ■

φ é chamado o **fluxo local** do campo vetorial X . Note que por causa das propriedades de grupo temos

$$(\varphi_t)^{-1} = \varphi_{-t}.$$

8.8 Colchete de Lie

Embora a Definição 8.29 de campos vetoriais permite também em princípio definir a composta de campos vetoriais e, já que Xf é interpretada como a derivada de f na direção de X , gostaríamos de interpretar naturalmente a expressão

$$X(Yf)$$

como a derivada segunda de f primeiro na direção de Y e em seguida na direção de X , em geral esta composta não é um campo vetorial porque não satisfaz a regra do produto:

$$\begin{aligned}(X \circ Y)(fg) &= X[Y(fg)] = X[(Yf)g + f(Yg)] = X[(Yf)g] + X[f(Yg)] \\ &= [X(Yf)]g + (Yf)(Xg) + (Xf)(Yg) + f[X(Yg)] \\ &= [(X \circ Y)f]g + f[(X \circ Y)g] + (Xf)(Yg) + (Yf)(Xg);\end{aligned}$$

em coordenadas locais (veja Proposição 8.37 a seguir), a composta realmente envolve derivadas parciais de segunda ordem, as quais não são vetores tangentes por não satisfazerem a regra do produto. Para definir cálculo diferencial de ordem superior, é necessário o conceito de *derivada covariante*, que veremos no Capítulo 11.

Por outro lado, a operação

$$X \circ Y - Y \circ X$$

define um campo vetorial.

8.35 Definição. Sejam $X, Y \in \mathcal{T}(M)$. O **colchete de Lie** de X e Y é o campo vetorial

$$[X, Y] = XY - YX.$$

□

Esta expressão deve ser entendida no sentido de

$$[X, Y] = X \circ Y - Y \circ X,$$

ou seja,

$$[X, Y]_p f = X_p(Yf) - Y_p(Xf).$$

O colchete de Lie é de fato um campo vetorial, pois

$$\begin{aligned} [X, Y](\alpha f + \beta g) &= X[Y(\alpha f + \beta g)] - Y[X(\alpha f + \beta g)] \\ &= X[\alpha Yf + \beta Yg] - Y[\alpha Xf + \beta Xg] \\ &= \alpha X(Yf) + \beta X(Yg) - \alpha Y(Xf) - \beta Y(Xg) \\ &= \alpha[X(Yf) - Y(Xf)] + \beta[X(Yg) - Y(Xg)] \\ &= \alpha[X, Y]f + \beta[X, Y]g \end{aligned}$$

e

$$\begin{aligned} [X, Y](fg) &= X[Y(fg)] - Y[X(fg)] \\ &= X[fYg + gYf] - Y[fXg + gXf] \\ &= X[fYg] + X[gYf] - Y[fXg] - Y[gXf] \\ &= fX(Yg) + YgXf + gX(Yf) + YfXg - fY(Xg) - XgYf - gY(Xf) - XfYg \\ &= f[X(Yg) - Y(Xg)] + g[X(Yf) - Y(Xf)] \\ &= f[X, Y](g) + g[X, Y](f). \end{aligned}$$

8.36 Proposição. O colchete de Lie satisfaaz as seguintes propriedades:

(i) (Anticomutatividade)

$$[X, Y] = -[Y, X].$$

Consequentemente,

$$[X, X] = 0.$$

(ii) (Bilinearidade)

$$\begin{aligned} [\alpha X + \beta Y, Z] &= \alpha[X, Z] + \beta[Y, Z], \\ [Z, \alpha X + \beta Y] &= \alpha[Z, X] + \beta[Z, Y]. \end{aligned}$$

(iii) (Identidade de Jacobi)

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

(iv)

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X.$$

(v) Se $F : M \rightarrow N$ é um difeomorfismo, então

$$F_*[X, Y] = [F_*X, F_*Y].$$

Prova: (i) e (ii) são imediatas. Para provar (iii), usando (ii) obtemos

$$\begin{aligned} [[X, Y], Z] &= [XY - YX, Z] = [XY, Z] - [YX, Z] \\ &= XYZ - ZXZ - YXZ + ZYX. \end{aligned}$$

Logo, usando (i) e novamente (ii), segue que

$$\begin{aligned} [[Y, Z], X] + [[Z, X], Y] &= -[X, [Y, Z]] - [Y, [Z, X]] \\ &= -[X, YZ - ZY] - [Y, ZX - XZ] \\ &= -[X, YZ] + [X, ZY] - [Y, ZX] + [Y, XZ] \\ &= -XYZ + YZX + XZY - ZYX - YZX + ZXZ + YXZ - XZY \\ &= -XYZ + ZXZ + YXZ - ZYX \\ &= -[[X, Y], Z]. \end{aligned}$$

A propriedade (iv) segue da regra do produto: se $h \in C^\infty(M)$,

$$\begin{aligned} [fX, gY] h &= f[X(g(Yh))] - g[Y(f(Xh))] \\ &= f[(Xg)(Yh)] + f[gX(Yh)] - g[(Yf)(Xh)] - g[fY(Xh)] \\ &= fgX(Yh) - gfY(Xh) + f[(Xg)(Yh)] - g[(Yf)(Xh)] \\ &= fg(XY - YX)h + [f(Xg)Y]h - [g(Yf)X]h \\ &= [fg[X, Y] + f(Xg)Y - g(Yf)X]h. \end{aligned}$$

(v) segue da Proposição 8.33: para todo $f \in C^\infty(N)$ temos

$$(XY)(f \circ F) = X[Y(f \circ F)] = X[(F_*Y)f \circ F] = (F_*X)(F_*Y)f \circ F$$

e, analogamente,

$$(YX)(f \circ F) = (F_*Y)(F_*X)f \circ F.$$

Logo,

$$\begin{aligned} (F_*[X, Y])f &= [X, Y](f \circ F) \circ F^{-1} \\ &= (XY - YX)(f \circ F) \circ F^{-1} \\ &= [(F_*X)(F_*Y) - (F_*Y)(F_*X)]f \circ F \circ F^{-1} \\ &= [F_*X, F_*Y]f. \end{aligned}$$

■

Uma álgebra de Lie é um espaço vetorial em que se define um produto (ou seja, uma aplicação bilinear) anticomutativo que satisfaz a identidade de Jacobi (veja o Capítulo 10). Portanto, esta proposição mostra que $\mathcal{T}(M)$ com a operação colchete é uma álgebra de Lie.

8.37 Proposição (Colchete de Lie em coordenadas locais). Se $X, Y \in \mathcal{T}(M)$ são campos vetoriais que se expressam em coordenadas locais por

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i} \quad e \quad Y = \sum_{i=1}^n Y^i \frac{\partial}{\partial x^i},$$

então

$$[X, Y] = \sum_{i,j=1}^n \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j},$$

ou, em notação mais sucinta,

$$[X, Y] = \sum_{j=1}^n (X(Y^j) - Y(X^j)) \frac{\partial}{\partial x^j}.$$

Em particular,

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] = 0$$

para todos i, j .

Prova: Temos

$$\begin{aligned} X(Yf) &= X \left(\sum_{i=1}^n Y^i \frac{\partial f}{\partial x^i} \right) = \sum_{i=1}^n X \left(Y^i \frac{\partial f}{\partial x^i} \right) = \sum_{i=1}^n Y^i X \left(\frac{\partial f}{\partial x^i} \right) + \sum_{i=1}^n \frac{\partial f}{\partial x^i} X(Y^i) \\ &= \sum_{i=1}^n Y^i \left(\sum_{j=1}^n X^j \frac{\partial^2 f}{\partial x^j \partial x^i} \right) + \sum_{i=1}^n \frac{\partial f}{\partial x^i} \left(\sum_{j=1}^n X^j \frac{\partial Y^i}{\partial x^j} \right) \\ &= \sum_{i,j=1}^n X^j Y^i \frac{\partial^2 f}{\partial x^j \partial x^i} + \sum_{i,j=1}^n X^j \frac{\partial Y^i}{\partial x^j} \frac{\partial f}{\partial x^i} \end{aligned}$$

e, por simetria,

$$Y(Xf) = \sum_{i,j=1}^n Y^j X^i \frac{\partial^2 f}{\partial x^j \partial x^i} + \sum_{i,j=1}^n Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i} = \sum_{i,j=1}^n X^j Y^i \frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i,j=1}^n Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i}.$$

Como

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i},$$

os termos envolvendo as derivadas parciais de segunda ordem se cancelam ao calcularmos $[X, Y] f = X(Yf) - Y(Xf)$ e a expressão do enunciado é obtida trocando os índices i, j . ■

8.9 Derivada de Lie

Em princípio, é um problema diferenciar campos vetoriais em variedades, já que não podemos tomar a diferença de vetores que moram em espaços tangentes diferentes (não há uma maneira de identificar os espaços tangentes com \mathbb{R}^n de uma maneira que seja invariante por mudança de coordenadas). Uma solução é a seguinte. Dado um campo Y em uma variedade que queremos diferenciar na direção de um vetor tangente X_p no ponto p , primeiro estendemos X_p a um campo vetorial X definido em toda a variedade. O campo vetorial X tem um fluxo local φ_t definido. Usamos o fluxo para levar o vetor $Y_{\varphi_t(p)}$ ao longo da trajetória reversa φ_{-t} do campo X para o espaço tangente $T_p M$ e fazer a diferença lá com o vetor Y_p , tomando em seguida o limite quanto $t \rightarrow 0$. No Capítulo 11 veremos o conceito de *derivada covariante*, que é uma solução diferente para este problema, mais semelhante ao conceito de derivada direcional, porque dependerá apenas do valor do vetor tangente X_p e não do valor de X ao longo de uma curva; no caso desta *derivada de Lie*, ela depende do valor de X ao longo de uma trajetória do campo.

8.38 Definição. Sejam $X, Y \in \mathcal{T}(M)$ campos vetoriais, $p \in M$ e φ_t o fluxo local do campo X em uma vizinhança V de p em M . A **derivada de Lie** do campo Y na direção do campo X em p é definida por

$$(\mathcal{L}_X Y)_p = \lim_{t \rightarrow 0} \frac{[d\varphi_{-t}]_{\varphi_t(p)} (Y_{\varphi_t(p)}) - Y_p}{t} = \frac{d}{dt} [d\varphi_{-t}]_{\varphi_t(p)} Y_{\varphi_t(p)} \Big|_{t=0}.$$

Na linguagem de pushforwards,

$$(\mathcal{L}_X Y)_p = \lim_{t \rightarrow 0} \frac{[(\varphi_{-t})_* Y]_p - Y_p}{t}.$$

A definição de derivada de Lie não é operacionalmente útil, já que em geral é muito difícil e mesmo impossível obter o fluxo explicitamente. Felizmente, como veremos agora, a derivada de Lie coincide com o colchete de Lie e este é muito fácil de calcular.

8.39 Teorema (Interpretação Geométrica do Colchete de Lie). *Se $X, Y \in \mathcal{T}(M)$ são campos vetoriais, $p \in M$ e φ_t é o fluxo local do campo X em uma vizinhança V de p em M então*

$$[X, Y] = \mathcal{L}_X Y.$$

Prova: Primeiro observe que se $g : (-\delta, \delta) \times V \rightarrow \mathbb{R}$ é uma função diferenciável tal que

$$g(0, q) = 0 \quad \text{para todo } q \in V,$$

então existe uma aplicação diferenciável $h : (-\delta, \delta) \times V \rightarrow \mathbb{R}$ tal que

$$g(t, q) = th(t, q).$$

De fato, basta definir

$$h(t, q) = \int_0^1 \frac{\partial g}{\partial s}(ts, q) \, ds$$

e notar que, pelo Teorema Fundamental do Cálculo,

$$\begin{aligned} th(t, q) &= \int_0^1 t \frac{\partial g}{\partial s}(ts, q) \, ds = \int_0^1 \frac{\partial}{\partial s}[g(ts, q)] \, ds \\ &= [g(ts, q)]_{s=0}^{s=1} = g(t, q) - g(0, q) \\ &= g(t, q). \end{aligned}$$

Em particular, segue que

$$\left. \frac{\partial g}{\partial t}(t, q) \right|_{t=0} = h(0, q).$$

Seja agora $f \in C^\infty(M)$. Defina $g : (-\delta, \delta) \times V \rightarrow \mathbb{R}$ por

$$g(t, q) = f(q) - f(\varphi_{-t}(q)),$$

ou, em notação funcional,

$$g(t, \cdot) = f - f \circ \varphi_{-t}.$$

Então $g(0, q) = f(q) - f(\varphi_0(q)) = f(q) - f(q) = 0$, de modo que a observação que fizemos no início da demonstração se aplica e existe uma aplicação diferenciável $h : (-\delta, \delta) \times V \rightarrow \mathbb{R}$ tal que

$$f(\varphi_{-t}(q)) = f(q) - th(t, q)$$

isto é,

$$f \circ \varphi_{-t} = f - th(t, \cdot)$$

e, além disso (por definição de vetor tangente, lembrando que φ_{-t} é uma curva diferenciável, trajetória do fluxo do campo X na direção reversa),

$$\begin{aligned} h(0, q) &= \left. \frac{\partial g}{\partial t}(t, q) \right|_{t=0} = - \left. \frac{\partial(f \circ \varphi_{-t})}{\partial t} \right|_{t=0} \\ &= - \left. \frac{\partial \varphi_{-t}(q)}{\partial t} \right|_{t=0} f = X_{\varphi(0, q)} f \\ &= X_q f. \end{aligned}$$

Daí (na primeira equação na demonstração da Proposição 8.31 substitua F por φ_{-t} e p por $\varphi_t(p)$),

$$\left[[d\varphi_{-t}]_{\varphi_t(p)} (Y_{\varphi_t(p)}) \right] f = Y_{\varphi_t(p)} (f \circ \varphi_{-t}) = Y_{\varphi_t(p)} f - t Y_{\varphi_t(p)} (h(t, \cdot)).$$

Portanto,

$$\begin{aligned} (\mathcal{L}_X Y)_p f &= \lim_{t \rightarrow 0} \frac{\left[[d\varphi_{-t}]_{\varphi_t(p)} (Y_{\varphi_t(p)}) \right] f - Y_p f}{t} \\ &= \lim_{t \rightarrow 0} \frac{Y_{\varphi_t(p)} f - t Y_{\varphi_t(p)} (h(t, \cdot)) - Y_p f}{t} \\ &= \lim_{t \rightarrow 0} \frac{Y_{\varphi_t(p)} f - Y_p f}{t} - \lim_{t \rightarrow 0} Y_{\varphi_t(p)} (h(t, \cdot)) \\ &= \lim_{t \rightarrow 0} \frac{(Y f)(\varphi_t(p)) - (Y f)(p)}{t} - Y_p (h(0, \cdot)) \\ &= \frac{\partial \varphi_t(p)}{\partial t} \Big|_{t=0} (Y f) - Y_p (h(0, \cdot)) \\ &= X_p (Y f) - Y_p (X f) \\ &= [X, Y]_p f. \end{aligned}$$

■
Portanto, o colchete de Lie de dois campos vetoriais é a derivada de Lie. E a derivada de Lie é a “derivada direcional” do segundo campo vetorial ao longo do fluxo do primeiro; ela não é uma derivada direcional no senso exato do termo, porque ela não depende apenas da direção do primeiro campo, ou seja, não podemos usar qualquer curva tangente ao primeiro campo para calculá-la, mas apenas uma trajetória do campo. A principal diferença entre a derivada de Lie e a derivada covariante (que é uma derivada direcional na correta assumpção da palavra) está então resumida nas Proposições 8.37 e 11.2: enquanto que a derivada covariante $(\nabla_X Y)_p$ depende apenas do valor de X em p e do valor de Y ao longo de uma curva tangente a X_p , a derivada de Lie $(\mathcal{L}_X Y)_p$ depende dos valores de X ao longo de uma curva tangente a Y_p e do valor de Y ao longo de uma curva tangente a X_p : de fato,

$$(\mathcal{L}_X Y)_p = [X, Y]_p = \sum_{i=1}^n (X_p (Y^i) - Y_p (X^i)) \frac{\partial}{\partial x^i},$$

e por definição de vetor tangente, os coeficientes $X_p (Y^1), \dots, X_p (Y^n)$ dependem dos valores de Y ao longo de uma curva passando por p cujo vetor tangente em p é X_p e os coeficientes $Y_p (X^1), \dots, Y_p (X^n)$ dependem dos valores de X ao longo de uma curva passando por p cujo vetor tangente em p é Y_p .

Outra diferença importante entre a derivada de Lie e a derivada covariante é que em variedades riemannianas esta é mais natural no seguinte sentido. Denotando $q = \varphi_t(p)$, na derivada de Lie o vetor Y_q é trazido ao longo da trajetória do campo X para o espaço tangente $T_p M$, onde ele é subtraído do vetor X_p através do operador linear $(\varphi_{-t})_*$. Em uma interpretação geométrica que veremos na Proposição 11.27, o vetor Y_q também é trazido ao longo da trajetória do campo X para o espaço tangente $T_p M$, mas através de um operador linear chamado *transporte paralelo*, que é interpretado, como o nome indica, como um operador que não muda a direção do vetor original em um sentido que veremos em maiores detalhes no Capítulo 11. Assim, o conceito de derivada covariante está mais próximo ao conceito de derivada direcional em \mathbb{R}^n , onde identificamos $T_p \mathbb{R}^n$ através de translações, que preservam as direções de vetores.

Mas é importante ressaltar que nenhuma das derivadas, tanto a derivada de Lie quanto a derivada covariante, dependem apenas de X_p e Y_p , ou seja, nenhum deles é um *tensor*, um conceito que veremos no próximo capítulo. Apesar disso, ambos aparecerão na definição do segundo tensor mais importante em Geometria Riemanniana, o *tensor curvatura* (o tensor mais importante é obviamente o *tensor métrica*).

8.10 Campos Vetoriais que Comutam

8.40 Lema. *Sejam M, N variedades diferenciáveis e $F : M \rightarrow N$ um difeomorfismo. Se X é um campo vetorial em M com fluxo local φ_t em uma vizinhança V , então o campo vetorial F_*X em N tem fluxo local $F \circ \varphi_t \circ F^{-1}$ em $F(V)$.*

Prova: Em outras palavras, se $\varphi : (-\delta, \delta) \times V \rightarrow M$ é o fluxo local de X em V , então $\psi : (-\delta, \delta) \times F(V) \rightarrow N$ dado por

$$\psi(t, q) = F(\varphi_t(F^{-1}(q)))$$

é o fluxo local do campo F_*X . Para provar este resultado, note primeiro que se $f \in C^\infty(M)$, então por definição de vetor tangente

$$X_p(f) = \frac{d}{dt} f \circ \varphi_t(p) \Big|_{t=0} = \lim_{t \rightarrow 0} \frac{f(\varphi_t(p)) - f(p)}{t}$$

porque a trajetória $\varphi_t(p)$ é uma curva diferenciável que tem X_p como vetor tangente em $t = 0$. Por definição, se $q = F(p)$, temos

$$\begin{aligned} (F_*X)_q(f) &= [dF_p(X_p)] f \\ &= X_p(f \circ F) \\ &= \lim_{t \rightarrow 0} \frac{(f \circ F)(\varphi_t(p)) - (f \circ F)(p)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(F \circ \varphi_t(F^{-1}(q))) - f(F \circ F)(F^{-1}(q))}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(F \circ \varphi_t \circ F^{-1}(q)) - f(q)}{t}, \end{aligned}$$

o que significa que a curva diferenciável $F \circ \varphi_t \circ F^{-1}$ tem $(F_*X)_q$ como vetor tangente em q , logo é o fluxo local do campo F_*X . ■

8.41 Corolário. *Se M é uma variedade diferenciável e $F : M \rightarrow M$ é um difeomorfismo, então*

$$F_*X = X$$

se e somente se

$$F \circ \varphi_t = \varphi_t \circ F.$$

8.42 Teorema. *Se $X, Y \in \mathcal{T}(M)$ são campos vetoriais e φ_t, ψ_s são os fluxos locais respectivos de X, Y em uma vizinhança V de M , então*

$$\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$$

se e somente se

$$[X, Y] = 0$$

em V .

Prova: Se $\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$, como φ_t é um difeomorfismo, segue do Corolário 8.41 que $(\varphi_t)_*Y = Y$, de modo que

$$[X, Y]_p = (\mathcal{L}_X Y)_p = \lim_{t \rightarrow 0} \frac{[(\varphi_{-t})_* Y]_p - Y_p}{t} = \lim_{t \rightarrow 0} \frac{Y_p - Y_p}{t} = 0$$

para todo $p \in V$.

Reciprocamente, se $[X, Y] = 0$ em V , considere a curva $\alpha : (-\varepsilon, \varepsilon) \rightarrow T_p M$ definida por

$$\alpha(t) = [(\varphi_{-t})_* Y]_p.$$

Temos, observando que o pushforward satisfaz $(F \circ G)_* = F_* \circ G_*$

$$\begin{aligned}
\alpha'(t) &= \lim_{h \rightarrow 0} \frac{\alpha(t+h) - \alpha(t)}{h} \\
&= \lim_{h \rightarrow 0} \frac{[(\varphi_{-t-h})_* Y]_p - [(\varphi_{-t})_* Y]_p}{h} \\
&= \lim_{h \rightarrow 0} \frac{[(\varphi_{-t})_* (\varphi_{-h})_* Y]_p - [(\varphi_{-t})_* Y]_p}{h} \\
&= \lim_{h \rightarrow 0} \frac{(d\varphi_{-t})_{\varphi_t(p)} [(\varphi_{-h})_* Y]_{\varphi_t(p)} - (d\varphi_{-t})_{\varphi_t(p)} Y_{\varphi_t(p)}}{h} \\
&= (d\varphi_{-t})_{\varphi_t(p)} \lim_{h \rightarrow 0} \frac{[(\varphi_{-h})_* Y]_{\varphi_t(p)} - Y_{\varphi_t(p)}}{h} \\
&= (d\varphi_{-t})_{\varphi_t(p)} ([X, Y]_{\varphi_t(p)}) \\
&= (d\varphi_{-t})_{\varphi_t(p)} (0) \\
&= 0.
\end{aligned}$$

Portanto, $\alpha(t) = \alpha(0)$, o que implica $(\varphi_{-t})_* Y = Y$, e o resultado segue do Corolário 8.41. ■

Em particular,

$$\varphi_t \circ \varphi_s \circ \varphi_{-t} \circ \varphi_{-s} = \text{id}.$$

Isso significa o seguinte, em outras palavras: quando $[X, Y] = 0$ em uma vizinhança V de $p \in M$, se a partir de p percorrermos a trajetória do campo X durante um intervalo de tempo t atingindo um certo ponto p_1 , e depois percorrermos a partir de p_1 a trajetória do campo Y durante um intervalo de tempo s atingindo um segundo ponto p_2 , voltarmos a partir de p_2 ao longo da trajetória do campo X durante um intervalo de tempo t atingindo um certo ponto p_3 e finalmente voltarmos também de p_3 ao longo da trajetória do campo Y durante um intervalo de tempo s , chegaremos ao ponto original p (obviamente estamos assumindo que em nenhum momento saímos da vizinhança V , o que será verdade para deslocamentos s, t pequenos para os quais os fluxos locais de X e Y estão definidos em V). Se $[X, Y] \neq 0$, isso não é verdade e terminamos em um ponto q diferente de p . O colchete de Lie portanto mede infinitesimalmente este defeito.

8.43 Teorema. *Se $E_1, \dots, E_k \in \mathcal{T}(M)$ são campos vetoriais linearmente independentes suaves em uma vizinhança de $p \in M$ tais que*

$$[E_i, E_j] = 0$$

para todos $i, j = 1, \dots, k$, então existe uma vizinhança coordenada (x^1, \dots, x^n) de p tal que

$$E_i = \frac{\partial}{\partial x^i}$$

para $i = 1, \dots, k$.

Prova: Sem perda de generalidade, podemos assumir através de uma carta adequada que $M = U \subset \mathbb{R}^n$, $p = 0$ e

$$E_i(0) = e_i$$

para $i = 1, \dots, k$, onde $\{e_1, \dots, e_n\}$ é a base canônica de \mathbb{R}^n . Seja φ_t^i o fluxo gerado pelo campo E_i . Defina

$$\begin{aligned}
\psi(x^1, \dots, x^n) &= \varphi_{x^1}^1 \circ \varphi_{x^2}^2 \circ \dots \circ \varphi_{x^k}^k (0, \dots, 0, x^{k+1}, \dots, x^n) \\
&= \varphi_{x^1}^1 (\varphi_{x^2}^2 (\dots (\varphi_{x^k}^k (0, \dots, 0, x^{k+1}, \dots, x^n)) \dots)).
\end{aligned}$$

[Note que no caso especial em que $k = n$, a aplicação ψ é

$$\begin{aligned}
\psi(x) &= \psi(x^1, \dots, x^n) = \varphi_{x^1}^1 \circ \dots \circ \varphi_{x^n}^n (0) \\
&= \varphi_{x^1}^1 (\varphi_{x^2}^2 (\dots (\varphi_{x^n}^n (0)) \dots)).
\end{aligned}$$

Em outras palavras, para calcular $\psi(x^1, \dots, x^n)$, percorremos sucessivamente as trajetórias dos campos E_n, \dots, E_1 a partir da origem durante os intervalos de tempo x^n, \dots, x^1 : primeiro, saindo da origem, percorremos a trajetória do campo E_n durante o intervalo de tempo x^n , chegando em um certo ponto $\varphi_{x^n}^n(0)$; partindo deste ponto percorremos a trajetória do campo E_{n-1} durante o intervalo de tempo x^{n-1} , chegando em um certo ponto $\varphi_{x^{n-1}}^{n-1}(\varphi_{x^n}^n(0))$; continuamos desta forma sucessivamente até chegar no ponto $\varphi_{x^1}^1(\varphi_{x^2}^2(\dots(\varphi_{x^n}^n(0))\dots))$ que definimos como sendo o ponto $\psi(x)$.]

Afirmamos que $d\psi_0 = I$. De fato, dada $f \in C^\infty(M)$, se $i = 1, \dots, k$

$$\begin{aligned} d\psi_0(e_i)(f) &= \frac{\partial(f \circ \psi)}{\partial x^i}(0) \\ &= \lim_{h \rightarrow 0} \frac{(f \circ \psi)\left(0, \dots, 0, \underset{1 \leq i \leq k}{h}, 0, \dots, 0\right) - f(\psi(0))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f \circ \varphi_0^1 \circ \dots \circ \varphi_0^{i-1} \circ \varphi_h^i \circ \varphi_0^{i+1} \circ \dots \circ \varphi_0^k(0) - f(0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\varphi_h^i(0)) - f(0)}{h} \\ &= E_i(0)(f) \\ &= e_i(f), \end{aligned}$$

enquanto que se $i = k + 1, \dots, n$, temos

$$\begin{aligned} d\psi_0(e_i)(f) &= \frac{\partial(f \circ \psi)}{\partial x^i}(0) \\ &= \lim_{h \rightarrow 0} \frac{(f \circ \psi)\left(0, \dots, 0, \underset{k+1 \leq i \leq n}{h}, 0, \dots, 0\right) - f(\psi(0))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f \circ \varphi_0^1 \circ \varphi_0^2 \circ \dots \circ \varphi_0^k \circ \dots \circ \varphi_0^k\left(0, \dots, 0, \underset{k+1 \leq i \leq n}{h}, 0, \dots, 0\right) - f(0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f\left(0, \dots, 0, \underset{k+1 \leq i \leq n}{h}, 0, \dots, 0\right) - f(0)}{h} \\ &= \frac{\partial f}{\partial x^i}(0) \\ &= e_i(f), \end{aligned}$$

portanto

$$d\psi_0(e_i) = e_i$$

para $i = 1, \dots, n$. Segue que $\psi(x) = (x^1, \dots, x^n)$ é um difeomorfismo local e portanto uma carta para uma vizinhança de $p = 0$.

Observe que podemos calcular explicitamente $d\psi_x(e_1)$ para todo x e não somente na origem, obtendo

$$\begin{aligned} d\psi_x(e_1)(f) &= \frac{\partial(f \circ \psi)}{\partial x^1}(x) \\ &= \lim_{h \rightarrow 0} \frac{(f \circ \psi)(x^1 + h, x^2, \dots, x^n) - f(\psi(x))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f \circ \varphi_{x^1+h}^1 \circ \varphi_{x^2}^2 \circ \dots \circ \varphi_{x^k}^k(0, \dots, 0, x^{k+1}, \dots, x^n) - f(\psi(x))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\varphi_{x^1+h}^1(\varphi_{x^2}^2(\dots(\varphi_{x^n}^n(0))\dots))) - f(0)}{h} \\ &= E_1(x)(f), \end{aligned}$$

ou seja,

$$E_1(x) = d\psi_x(e_1) = \left. \frac{\partial}{\partial x^1} \right|_x$$

para todo x onde a carta está definida. Isso prova o resultado para $i = 1$.

Mas, pelo teorema anterior, como $[E_i, E_j] = 0$, temos que os fluxos dos campos E_1, \dots, E_k comutam, isto é,

$$\varphi_t^i \circ \varphi_t^j = \varphi_t^j \circ \varphi_t^i$$

para todos $i, j = 1, \dots, k$. Logo, para $i = 2, \dots, k$ podemos escrever

$$\psi(x^1, \dots, x^n) = \varphi_{x^i}^i \circ \varphi_{x^1}^1 \circ \varphi_{x^2}^2 \circ \dots \circ \widehat{\varphi_{x^i}^i} \circ \dots \circ \varphi_{x^k}^k(0, \dots, 0, x^{k+1}, \dots, x^n).$$

Pelo mesmo argumento anterior no caso $i = 1$, concluímos que

$$E_i(x) = \left. \frac{\partial}{\partial x^i} \right|_x$$

para $i = 2, \dots, k$, para todo x onde a carta ψ está definida, terminando a demonstração do resultado. ■ Assim, o colchete de Lie mede em algum sentido o quanto as trajetórias de campos linearmente independentes E_1, \dots, E_n podem ser usadas para formar as “retas coordenadas” de um sistema de coordenadas (veja [Spivak], Vol. I, pp. 220-221, para uma afirmação mais precisa deste resultado).

8.11 Exercícios

8.44 Exercício. Considere o subconjunto $X = (\mathbb{R} \times \{0\}) \cup (\overline{\mathbb{R}}_+ \times \{1\})$ de \mathbb{R}^2 com a seguinte base para a sua topologia:

- (a) intervalos do tipo $(a, b) \times \{0\}$, $a < b$;
- (b) intervalos do tipo $(a, b) \times \{1\}$, $0 \leq a < b$;
- (c) uniões de intervalos do tipo $[(a, 0) \times \{0\}] \cup [(0, b) \times \{1\}]$, $a < 0 < b$.

Verifique que X possui uma base enumerável, que conjuntos finitos de pontos são fechados em X e que todo ponto de X possui uma vizinhança difeomorfa a um conjunto aberto de \mathbb{R} . Observe que X não é, no entanto, um espaço de Hausdorff, porque os pontos $(0, 0)$ e $(0, 1)$ não possuem vizinhanças abertas disjuntas.

8.45 Exercício. Seja X um conjunto. Suponha que exista uma família

$$\Phi = \{\varphi_\alpha\}_{\alpha \in \mathcal{A}}$$

de aplicações injetivas $\varphi_\alpha : U_\alpha \longrightarrow V_\alpha$ de um aberto $U_\alpha \subset \mathbb{R}^n$ sobre um subconjunto V_α de X para cada $\alpha \in \mathcal{A}$, satisfazendo as seguintes condições:

(1) Os abertos V_α cobrem X , isto é,

$$\bigcup_{\alpha \in A} V_\alpha = X.$$

(2) Para todos índices $\alpha, \beta \in A$ tais que $V_{\alpha\beta} = V_\alpha \cap V_\beta \neq \emptyset$, os conjuntos $\varphi_\alpha^{-1}(V_{\alpha\beta}), \varphi_\beta^{-1}(V_{\alpha\beta})$ são abertos em \mathbb{R}^n e as aplicações

$$\begin{aligned}\varphi_{\alpha\beta} &= \varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}(V_{\alpha\beta}) \longrightarrow \varphi_\beta^{-1}(V_{\alpha\beta}), \\ \varphi_{\beta\alpha} &= \varphi_\alpha^{-1} \circ \varphi_\beta : \varphi_\beta^{-1}(V_{\alpha\beta}) \longrightarrow \varphi_\alpha^{-1}(V_{\alpha\beta}),\end{aligned}$$

são diferenciáveis de classe C^∞ .

Então existe uma única topologia \mathcal{T} em X relativa à qual Φ é um atlas diferenciável de dimensão n para X .

\mathcal{T} não é necessariamente de Hausdorff nem precisa possuir base enumerável (ela é apenas localmente de Hausdorff, pois se $p, q \in V_\alpha$, $p \neq q$, eles possuem vizinhanças disjuntas porque V_α é homeomorfo a um aberto de \mathbb{R}^n). Para isso, são necessárias condições adicionais:

A topologia \mathcal{T} é de Hausdorff se e somente se

3) Para todos índices $\alpha, \beta \in A$ tais que $V_{\alpha\beta} = V_\alpha \cap V_\beta \neq \emptyset$, não existe nenhuma sequência $\{x_k\}_{k \in \mathbb{N}} \subset \varphi_\alpha^{-1}(V_{\alpha\beta})$ tal que

$$\begin{aligned}x_k &\longrightarrow x \in U_\alpha \setminus \varphi_\alpha^{-1}(V_{\alpha\beta}), \\ \varphi_\beta^{-1} \circ \varphi_\alpha(x_k) &\longrightarrow y \in U_\beta \setminus \varphi_\beta^{-1}(V_{\alpha\beta}).\end{aligned}$$

Ela possui uma base enumerável se e somente se

(4) A cobertura $\{V_\alpha\}_{\alpha \in A}$ possui uma cobertura enumerável.

8.46 Exercício. Mostre que uma variedade diferenciável M é conexa se e somente se ela é conexa por caminhos.

Além disso, verifique que as componentes conexas de M são as suas componentes conexas por caminhos. Finalmente, prove que M possui no máximo um número enumerável de componentes conexas.

Capítulo 9

Tensores

9.1 Vetores Contravariantes e Covariantes

Considere o conceito de *vetor* em \mathbb{R}^n , por exemplo o *vetor velocidade* de uma curva descrita no sistema de coordenadas (x^1, \dots, x^n) por

$$x(t) = (x^1(t), \dots, x^n(t)).$$

Temos

$$\frac{dx}{dt} = \left(\frac{dx^1}{dt}, \dots, \frac{dx^n}{dt} \right).$$

Em um outro sistema de coordenadas (y^1, \dots, y^n) a curva é descrita por:

$$y(t) = (y^1(t), \dots, y^n(t)),$$

de modo que seu vetor velocidade neste sistema de coordenadas é dado por

$$\frac{dy}{dt} = \left(\frac{dy^1}{dt}, \dots, \frac{dy^n}{dt} \right).$$

A regra da cadeia nos dá como as coordenadas do vetor velocidade mudam de um sistema de coordenadas para o outro:

$$\frac{dy^i}{dt} = \sum_{j=1}^n \frac{\partial y^i}{\partial x^j} \frac{dx^j}{dt} \quad (9.1)$$

para $i = 1, \dots, n$.

Considere agora o conceito do *gradiente* de uma função, usualmente identificado com um vetor. No sistema de coordenadas (x^1, \dots, x^n) , o gradiente é definido por

$$\nabla_x f(x) = \left(\frac{\partial f}{\partial x^1}(x), \dots, \frac{\partial f}{\partial x^n}(x) \right)$$

enquanto que no sistema de coordenadas (y^1, \dots, y^n) o gradiente é dado por

$$\nabla_y f(x) = \left(\frac{\partial f}{\partial y^1}(x), \dots, \frac{\partial f}{\partial y^n}(x) \right)$$

Novamente, a regra da cadeia nos dá como as coordenadas do gradiente mudam de um sistema de coordenadas para o outro:

$$\frac{\partial f}{\partial y^i} = \sum_{j=1}^n \frac{\partial x^j}{\partial y^i} \frac{\partial f}{\partial x^j} \quad (9.2)$$

para $i = 1, \dots, n$.

Comparando as expressões (9.1) e (9.2), vemos que elas são bem diferentes. Isso fica ainda mais claro se considerarmos o Jacobiano da mudança de coordenadas $y = y(x)$,

$$J = \left[\frac{\partial y^i}{\partial x^j} \right] \quad (9.3)$$

ou seja,

$$J = \begin{bmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial y^n}{\partial x^1} & \cdots & \frac{\partial y^n}{\partial x^n} \end{bmatrix}.$$

Temos

$$\frac{dy}{dt} = J \frac{dx}{dt} \quad (9.4)$$

enquanto que

$$\nabla_y f = (J^{-1})^T \nabla_x f, \quad (9.5)$$

pois

$$J^{-1} = \left[\frac{dx^i}{dy^j} \right] = \begin{bmatrix} \frac{\partial x^1}{\partial y^1} & \cdots & \frac{\partial x^1}{\partial y^n} \\ \vdots & & \vdots \\ \frac{\partial x^n}{\partial y^1} & \cdots & \frac{\partial x^n}{\partial y^n} \end{bmatrix}.$$

Note que as leis de transformação não são exatamente uma a inversa da outra no sentido matricial, já que é necessário transpor a matriz de mudança de coordenadas. Observe também que para as fórmulas concidirem, teríamos que ter

$$J = (J^{-1})^T,$$

isto é, J precisaria ser uma transformação ortogonal, o que equivale a requerer que os dois sistemas de coordenadas (x^1, \dots, x^n) e (y^1, \dots, y^n) sejam ortonormais, o que raramente ocorre.

O fato de que o gradiente de uma função sob uma mudança de coordenadas transformar-se de uma maneira diferente da de um vetor mostra que ele é um tipo diferente de vetor. Como veremos os motivos na próxima seção, vetores que se transformam de acordo com a expressão (9.1) são chamados *vetores contravariantes*, enquanto que vetores que se transformam de acordo com a expressão (9.2) são chamados *vetores covariantes* (ou simplesmente *covetores*).

As coordenadas de um vetor contravariante são convencionalmente denotadas por *superescritos*:

$$v = (v^1, \dots, v^n), \quad (9.6)$$

porque, como nos casos dos vetores deslocamento, velocidade, aceleração, etc., ou seja, vetores cujas dimensões estão diretamente relacionadas às dimensões das coordenadas, o deslocamento aparece no numerador (*acima* da barra da fração), enquanto que as coordenadas de um vetor covariante são convencionalmente denotadas por *subescritos*:

$$v = (v_1, \dots, v_n), \quad (9.7)$$

porque, como no caso do covetor gradiente, o deslocamento aparece no denominador (*abaixo* da barra da fração), ou seja, vetores tais como o gradiente tem dimensões que são inversas às dimensões das coordenadas.

9.1.1 Significado Real do Gradiente

A derivada de uma função real $f : \mathbb{R}^n \rightarrow \mathbb{R}$ em um ponto $x \in \mathbb{R}^n$ é um **funcional linear** $df_x : \mathbb{R}^n \rightarrow \mathbb{R}$. O gradiente realmente não é um vetor, mas sim um *funcional linear* ou uma *1-forma* (estes termos são sinônimos). Como veremos daqui a pouco, funcionais lineares são vetores (no espaço dual) que se comportam com relação a mudança de coordenadas como covetores.

Assim, embora diferenciemos entre a diferencial df_x de uma função real $f : \mathbb{R}^n \rightarrow \mathbb{R}$, que é um funcional linear, e a função **gradiente**, que associa a cada ponto x um vetor $\nabla f(x)$ ou $\text{grad } f(x)$, com a propriedade especial que

$$df_x(v) = \langle \nabla f(x), v \rangle$$

para todo $v \in \mathbb{R}^n$, onde $\langle \cdot, \cdot \rangle$ denota o produto interno canônico de \mathbb{R}^n , de qualquer forma, devido à sua definição o vetor gradiente é um covetor e se comporta como tal.

Este fato não é apenas um acidente restrito à forma especial com que ele se comporta com relação a uma mudança de coordenadas, mas também é uma consequência do significado geométrico de funcionais lineares e do gradiente. Através do produto interno, qualquer funcional linear $\omega : \mathbb{R}^n \rightarrow \mathbb{R}$ é identificado com um único vetor v de \mathbb{R}^n : v é o único vetor tal que

$$\omega(w) = \langle v, w \rangle$$

para todo $w \in \mathbb{R}^n$. Este vetor v é portanto perpendicular ao hiperplano $\ker \omega$, o núcleo do funcional ω . A ação do funcional linear ω sobre um vetor arbitrário w pode ser então vista da seguinte forma: ω determina uma família de hiperplanos, os hiperplanos paralelos a $\ker \omega$; $\omega(w)$ é então o número de hiperplanos que a “seta” do vetor w “perfura” por unidade de distância (esta é medida exatamente pelo produto interno). Para vetores com o mesmo comprimento de v , o vetor v é o que perfura o maior número de hiperplanos, já que é perpendicular a todos estes hiperplanos (a mesma consideração evidentemente vale para $-v$). Outros vetores diferentes de v e $-v$ formarão um ângulo não reto com estes hiperplanos e, se tiverem o mesmo comprimento que o vetor v , eles perfurarão consequentemente menos hiperplanos. Ou seja, se $\|w\| = \|v\|$ mas $w \neq v$, então

$$\omega(w) < \omega(v);$$

de fato,

$$\omega(w) = \langle v, w \rangle = \|v\| \|w\| \cos \theta < \|v\| \|w\| = \|v\|^2 = \omega(v).$$

Se w é ortogonal a v , então w está no núcleo de ω e não perfura nenhum hiperplano da família; assim, $\omega(w) = 0$.

O vetor gradiente $\nabla f(x)$ também se comporta geometricamente desta forma. Exceto que no caso do gradiente substituímos a família de hiperplanos paralelos ao núcleo do funcional pela família das hiperfícies de nível da função f . O vetor gradiente é perpendicular às hiperfícies de nível de f . Isto funciona porque segue da definição que o gradiente é perpendicular ao espaço tangente à hiperfície de nível: se $\alpha : I \rightarrow \mathbb{R}^n$ é uma curva contida em uma hiperfície de nível, então

$$f(\alpha(t)) \equiv c$$

para todo $t \in I$ algum valor real c ; derivando esta equação em relação a t , obtemos

$$df_{\alpha(t)}(\alpha'(t)) = 0,$$

ou seja,

$$\langle \nabla f(\alpha(t)), \alpha'(t) \rangle = 0.$$

Como isso vale para todas tais curvas, concluímos que $\nabla f(x)$ é perpendicular a $T_x \{f^{-1}(c)\}$. Portanto, o gradiente “perfura” as hiperfícies de nível. Como no caso linear, a direção do vetor gradiente é a direção em que mais hiperfícies de nível são perfuradas por unidade de distância. Isso é quase equivalente a dizer que o gradiente aponta na direção em que a função cresce com maior rapidez (como se demonstra em Cálculo), pois perfurar as hiperfícies de nível equivale a subir ou descer a montanha de contornos do gráfico de f , dependendo do sentido escolhido.

9.2 Convenção da Soma de Einstein

As escolhas que faremos para a notação neste capítulo e no futuro são necessárias para que a *convenção da soma de Einstein* funcione: ao invés de usar o sinal de somatório \sum para denotar uma soma, *convencionamos* que sempre que uma expressão contém um índice como um superescrito e o mesmo índice como subescrito, uma soma é implícita sobre todos os valores que este índice pode tomar. Alguns exemplos:

Convenção da Soma de Einstein	Notação de Somatório
$v^i e_i$	$\sum_{i=1}^n v^i e_i$
$e_i = A_i^j f_j$	$e_i = \sum_{j=1}^n A_i^j f_j$
$\omega_i e^i$	$\sum_{i=1}^n \omega_i e^i$
$e^k = (A^{-1})_l^k f^l$	$e^k = \sum_{l=1}^n (A^{-1})_l^k f^l$
$\frac{\partial}{\partial x^i} = \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$	$\frac{\partial}{\partial x^i} = \sum_{j=1}^n \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$
$T = T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l}$	$T = \sum_{\substack{i_1, \dots, i_k=1 \\ j_1, \dots, j_l=1}}^n T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l},$

9.3 Vetores e Covetores

9.3.1 Mudança de Coordenadas em Espaços Vetoriais

Dado um espaço vetorial real de dimensão finita V munido de uma base

$$\mathcal{B} = \{e_1, \dots, e_n\},$$

denotaremos por $[v]_{\mathcal{B}}$ o vetor coluna cujos elementos são as coordenadas do vetor v em relação à base \mathcal{B} , ou seja, se

$$v = v^i e_i,$$

então

$$[v]_{\mathcal{B}} = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}.$$

Também abusaremos esta notação às vezes, escrevendo $[v]_{\mathcal{B}} = (v^1, \dots, v^n)$.

9.1 Definição. Sejam V um espaço vetorial real e

$$\begin{aligned} \mathcal{B}_1 &= \{e_1, \dots, e_n\}, \\ \mathcal{B}_2 &= \{f_1, \dots, f_n\}, \end{aligned}$$

duas bases para V . A **matriz de mudança de coordenadas** da base B_1 para a base B_2 é a matriz A tal que

$$[v]_{\mathcal{B}_2} = A [v]_{\mathcal{B}_1}. \tag{9.8}$$

Quando necessário, ela será denotada por $A_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$. \square

9.2 Notação. Denotaremos o elemento que ocupa a i -ésima linha e a j -ésima coluna de A por A_{ij}^i . \square

9.3 Proposição. Sejam

$$\begin{aligned}\mathcal{B}_1 &= \{e_1, \dots, e_n\}, \\ \mathcal{B}_2 &= \{f_1, \dots, f_n\},\end{aligned}$$

duas bases para um espaço vetorial real V . Se $A = (A_j^i)$ é a matriz de mudança de coordenadas da base \mathcal{B}_1 para a base \mathcal{B}_2 , então os elementos desta matriz são definidos por

$$e_i = A_i^j f_j. \quad (9.9)$$

Ou seja, as colunas de A são as coordenadas dos vetores da base \mathcal{B}_1 em relação à base \mathcal{B}_2 .

Prova: De fato, se vale (9.9), então

$$v = v^i e_i = v^i (A_i^j f_j) = (A_i^j v^i) f_j,$$

que é exatamente (9.8):

$$[v]_{\mathcal{B}_2} = \begin{bmatrix} A_1^1 & \dots & A_n^1 \\ \vdots & & \vdots \\ A_1^n & \dots & A_n^n \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix} = A [v]_{\mathcal{B}_1}.$$

■

Observe agora que enquanto a lei de transformação dos *vetores* da base \mathcal{B}_2 para a base \mathcal{B}_1 é

$$e_i = A_i^j f_j, \quad (9.10)$$

a lei de transformação das *coordenadas* de vetores na base \mathcal{B}_2 para a base \mathcal{B}_1 é *contrária*: como

$$[v]_{\mathcal{B}_1} = A^{-1} [v]_{\mathcal{B}_2},$$

segue que se

$$\begin{aligned}[v]_{\mathcal{B}_1} &= (v^1, \dots, v^n), \\ [v]_{\mathcal{B}_2} &= (w^1, \dots, w^n),\end{aligned}$$

então

$$v^i = (A^{-1})_j^i w^j. \quad (9.11)$$

A lei de transformação (9.10) é considerada a *lei de transformação fundamental*. Portanto, a observação acima motiva a seguinte definição:

9.4 Definição. Vetores que se transformam de maneira contrária à lei (9.10) são chamados **vetores contravariantes**. □

Assim, os vetores do próprio espaço vetorial V são vetores contravariantes.

9.3.2 Covetores

9.5 Definição. Seja V um espaço vetorial real de dimensão finita. Um **covetor** de V é qualquer funcional linear $\omega : V \rightarrow \mathbb{R}$.

O espaço vetorial dos covetores de V , com as definições naturais de soma de covetores e multiplicação de covetores por escalares reais é chamado o **espaço dual** de V e denotado por V^* . □

Portanto, *covetor* de V nada mais é que um sinônimo para *funcional linear* sobre V .

9.6 Definição. Seja $\mathcal{B} = \{e_1, \dots, e_n\}$ uma base para o espaço vetorial V . Definimos a **base dual**

$$\mathcal{B}^* = \{e^1, \dots, e^n\}$$

de V^* por

$$e^i(e_j) = \delta_i^j, \quad i, j = 1, \dots, n. \quad (9.12)$$

□

Um covetor arbitrário $\omega \in V^*$ expressa-se em coordenadas com relação à base dual \mathcal{B}^* na forma

$$\omega = \omega_i e^i.$$

Observe que se

$$v = v^i e_i,$$

então

$$e^i(v) = v^i. \quad (9.13)$$

9.7 Definição. Sejam V, W espaços vetoriais. Dada uma aplicação linear $A : V \rightarrow W$, definimos a **aplicação linear dual ou transposta** $A^* : W^* \rightarrow V^*$ de A por

$$(A^* \omega)v = \omega(Av)$$

para todo $\omega \in W^*$ e para todo $v \in V$. □

9.8 Proposição. Sejam

$$\begin{aligned} \mathcal{B}_1 &= \{e_1, \dots, e_n\}, \\ \mathcal{B}_2 &= \{f_1, \dots, f_n\} \end{aligned}$$

duas bases para o espaço vetorial V e

$$\begin{aligned} \mathcal{B}_1^* &= \{e^1, \dots, e^n\}, \\ \mathcal{B}_2^* &= \{f^1, \dots, f^n\} \end{aligned}$$

as respectivas bases duais para V^* . Se A é a matriz de mudança de coordenadas da base \mathcal{B}_1 para a base \mathcal{B}_2 , então

$$e^i = (A^{-1})_j^i f^j. \quad (9.14)$$

Consequentemente, $(A^{-1})^T$ é a matriz de mudança de coordenadas da base dual \mathcal{B}_1^* para a base dual \mathcal{B}_2^* , isto é,

$$[\omega]_{\mathcal{B}_2^*} = (A^{-1})^T [\omega]_{\mathcal{B}_1^*}. \quad (9.15)$$

Prova. Pela Proposição (9.3),

$$e_i = A_i^j f_j.$$

Seja $B = (B_l^k)$ a matriz de transformação da base \mathcal{B}_1^* para a base \mathcal{B}_2^* , isto é,

$$e^k = B_l^k f^l.$$

Então

$$\begin{aligned} \delta_i^k &= e^k(e_i) = e^k(A_i^j f_j) = A_i^j e^k(f_j) = A_i^j B_l^k f^l(f_j) \\ &= A_i^j B_l^k \delta_j^l = A_i^j B_j^k = B_j^k A_i^j, \end{aligned}$$

de modo que $BA = I$, donde $B = A^{-1}$ e portanto

$$e^k = (A^{-1})_l^k f^l.$$

(9.15) segue da aplicação da Proposição (9.3) a (9.14), substituindo V por V^* e $\mathcal{B}_1, \mathcal{B}_2$ por $\mathcal{B}_1^*, \mathcal{B}_2^*$. ■
Portanto, assim como a lei de transformação dos vetores da base \mathcal{B}_2 para a base \mathcal{B}_1 (lei de transformação fundamental) é

$$e_i = A_i^j f_j,$$

a lei de transformação das *coordenadas* de vetores na base \mathcal{B}_2^* para a base \mathcal{B}_1^* é a mesma: como

$$[\omega]_{\mathcal{B}_1^*} = A^T [\omega]_{\mathcal{B}_2^*},$$

segue que se

$$\begin{aligned} [\omega]_{\mathcal{B}_1^*} &= (\omega_1, \dots, \omega_n), \\ [\omega]_{\mathcal{B}_2^*} &= (\sigma_1, \dots, \sigma_n), \end{aligned}$$

então

$$\omega_i = A_i^j \sigma_j. \quad (9.16)$$

Ou seja, covetores variam (se transformam) da mesma forma *como* variam (se transformam) os vetores da base do espaço vetorial, que convencionamos ser a lei de transformação fundamental. Esta observação motiva a seguinte definição:

9.9 Definição. Vetores que se transformam da mesma forma que a lei (9.10) são chamados **vetores covariantes**. □

Assim, os covetores do espaço dual V^* são vetores covariantes.

9.3.3 O Espaço Bidual

9.10 Definição. Seja V um espaço vetorial real de dimensão finita. O espaço dual $(V^*)^*$ do espaço dual de V é chamado o **espaço bidual** de V e denotado V^{**} . □

Uma importante identificação natural (isto é, um isomorfismo definido independentemente de bases e baseado apenas na estrutura linear) existe entre um espaço vetorial e seu espaço bidual:

9.11 Proposição. A aplicação $\Phi : V \rightarrow V^{**}$ definida por

$$\Phi(v)(\omega) = \omega(v)$$

é um isomorfismo natural entre V e V^{**} .

Prova. Como $\dim V = \dim V^{**}$, para verificar que Φ é um isomorfismo basta mostrar que ele é injetivo, isto é, que seu núcleo é o subespaço nulo. Seja $e_1 \in V$ um vetor não nulo qualquer. Estenda este vetor a uma base $\mathcal{B} = \{e_1, \dots, e_n\}$ para V . Seja $\mathcal{B}^* = \{e^1, \dots, e^n\}$ a correspondente base dual de V^* . Então $\Phi(e_1) \neq 0$ porque

$$\Phi(e_1)(e^1) = e^1(e_1) = 1.$$

■

Em vista desta identificação, um vetor $v \in V$ pode ser visto como um funcional linear sobre V^* cuja ação em covetores de V^* é dada por

$$v(\omega) = \omega(v). \quad (9.17)$$

Em particular,

$$e_i(e^j) = \delta_i^j \quad (9.18)$$

e se

$$\omega = \omega_i e^i,$$

então

$$e_i(\omega) = \omega_i. \quad (9.19)$$

9.4 Vetores e Covetores Tangentes

9.4.1 Mudança de Coordenadas no Espaço Tangente $T_p M$

Se $\varphi : U \rightarrow \varphi(U)$ e $\psi : V \rightarrow \psi(V)$ são duas cartas para vizinhanças coordenadas de $p = \varphi(x) = \varphi(x^1, \dots, x^n) = \psi(y) = \psi(y^1, \dots, y^n)$ em M , abusando a notação frequentemente escrevemos

$$(\psi^{-1} \circ \varphi)(x) = (\psi^{-1} \circ \varphi)(x^1, \dots, x^n) = (y^1(x^1, \dots, x^n), \dots, y^n(x^1, \dots, x^n)),$$

isto é, denotamos as função coordenadas $(\psi^{-1} \circ \varphi)^j(x)$ por $y^j(x)$.

9.12 Proposição. *Seja M uma variedade diferenciável n -dimensional e*

$$\begin{aligned}\varphi : U &\longrightarrow \varphi(U), \\ \psi : V &\longrightarrow \psi(V),\end{aligned}$$

duas cartas para vizinhanças coordenadas de $p = \varphi(x) = \psi(y)$ em M . Sejam

$$\begin{aligned}\mathcal{B}_x &= \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}, \\ \mathcal{B}_y &= \left\{ \frac{\partial}{\partial y^1} \Big|_p, \dots, \frac{\partial}{\partial y^n} \Big|_p \right\},\end{aligned}$$

as bases coordenadas de $T_p M$ induzidas pelas cartas φ e ψ , respectivamente. Denote

$$\frac{\partial y^j}{\partial x^i}(x) := \frac{\partial (\psi^{-1} \circ \varphi)^j}{\partial x_i}(x)$$

Então a matriz de mudança de coordenadas da base \mathcal{B}_x para a base \mathcal{B}_y é definida por

$$\frac{\partial}{\partial x^i} \Big|_p = \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \Big|_p. \quad (9.20)$$

Prova: Por definição e pela regra da cadeia,

$$\begin{aligned}\frac{\partial}{\partial x^i} \Big|_p &= d\varphi_x(e_i) = d\psi_y [d(\psi^{-1} \circ \varphi)_x(e_i)] = d\psi_y \left[\frac{\partial (\psi^{-1} \circ \varphi)^j}{\partial x^i}(x) f_j \right] = \frac{\partial (\psi^{-1} \circ \varphi)^j}{\partial x^i} d\psi_y(f_j) \\ &= \frac{\partial y^j}{\partial x_i} \frac{\partial}{\partial y^j} \Big|_p.\end{aligned}$$

■

Portanto, se um vetor $v \in T_p M$ se escreve em coordenadas em relação às bases \mathcal{B}_x e \mathcal{B}_y nas formas

$$\begin{aligned}v &= v_x^i \frac{\partial}{\partial x^i} \Big|_p, \\ v &= v_y^j \frac{\partial}{\partial y^j} \Big|_p,\end{aligned}$$

então, pelas Proposições 9.3 e 9.12, a lei de transformação de coordenadas é dada por

$$v_x^i = \frac{\partial x^i}{\partial y^j} v_y^j. \quad (9.21)$$

9.4.2 Covetores Tangentes

Enquanto que o conceito de vetores tangentes em variedades permite uma interpretação livre de coordenadas de derivadas de curvas, diferenciais de funções reais em variedades (ou seja, o análogo do gradiente em \mathbb{R}^n) são interpretadas de maneira mais natural como *covetores tangentes* (compare a Proposição 9.8 com a discussão na introdução deste capítulo).

9.13 Definição. Seja M uma variedade diferenciável. Para cada $p \in M$ definimos o **espaço cotangente** T_p^*M a M em p por

$$T_p^*M = (T_pM)^*.$$

Elementos de T_p^*M são chamados **covetores tangentes** a M em p . \square

Assim, o espaço cotangente a M em p é o dual do espaço tangente a M em p .

9.14 Definição. Seja M uma variedade diferenciável e $\varphi : U \rightarrow \varphi(U)$ uma carta de uma vizinhança coordenada de um ponto $p \in M$. A base coordenada

$$\mathcal{B}_p = \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$$

do espaço tangente T_pM associada à carta φ dá origem a uma **base dual coordenada** para o espaço cotangente T_p^*M associada à carta φ que denotaremos por

$$\mathcal{B}_p^* = \left\{ dx^1 \Big|_p, \dots, dx^n \Big|_p \right\}. \quad (9.22)$$

\square

Portanto, qualquer covetor $\omega \in T_p^*M$ pode ser escrito de maneira única como

$$\omega = \omega_i dx^i \Big|_p, \quad (9.23)$$

onde

$$\omega_i = \omega \left(\frac{\partial}{\partial x^i} \Big|_p \right). \quad (9.24)$$

Vamos investigar agora como as coordenadas de um covetor tangente se transformam quando há uma mudança de bases coordenadas, de uma carta para outra.

9.15 Proposição. Seja M uma variedade diferenciável n -dimensional e $\varphi : U \rightarrow \varphi(U), \psi : V \rightarrow \psi(V)$ duas cartas para vizinhanças coordenadas de $p = \varphi(x) = \psi(y)$ em M . Sejam

$$\begin{aligned} \mathcal{B}_x &= \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}, \\ \mathcal{B}_y &= \left\{ \frac{\partial}{\partial y^1} \Big|_p, \dots, \frac{\partial}{\partial y^n} \Big|_p \right\} \end{aligned}$$

as bases coordenadas de T_pM induzidas pelas cartas φ e ψ , respectivamente. Denote por

$$\begin{aligned} \mathcal{B}_x^* &= \left\{ dx^1 \Big|_p, \dots, dx^n \Big|_p \right\}, \\ \mathcal{B}_y^* &= \left\{ dy^1 \Big|_p, \dots, dy^n \Big|_p \right\} \end{aligned}$$

as respectivas bases duais. Então a matriz de mudança de coordenadas da base \mathcal{B}_x^* para a base \mathcal{B}_y^* é dada por

$$dx^i \Big|_p = \frac{\partial x^i}{\partial y^j} dy^j \Big|_p. \quad (9.25)$$

Prova: Pela Proposição 9.12, a mudança de coordenadas da base \mathcal{B}_x para a base \mathcal{B}_y é dada por

$$\frac{\partial}{\partial x^i} \Big|_p = \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \Big|_p.$$

O resultado segue então da Proposição 9.8. ■

Obtemos também da discussão que se segue à Proposição 9.8 que se

$$\begin{aligned} [\omega]_{\mathcal{B}_x^*} &= (\omega_1^x, \dots, \omega_n^x), \\ [\omega]_{\mathcal{B}_y^*} &= (\omega_1^y, \dots, \omega_n^y), \end{aligned}$$

então

$$\omega_x^i = \frac{\partial y^j}{\partial x^i} \omega_j^y.$$

Podemos agora entender a terminologia antiga em que vetores tangentes eram chamados *vetores contravariantes*, enquanto que covetores tangentes eram chamados *vetores covariantes*. É importante ressaltar que esta terminologia nada tem a ver com functores covariantes e contravariantes da teoria de categorias.

9.5 Tensores

9.5.1 Definição

9.16 Definição. Seja V um espaço vetorial real de dimensão finita e V^* seu espaço dual.

Um **k -tensor covariante** em V (ou *tensor covariante de ordem k*) é uma função real k -linear

$$T : \underbrace{V \times \dots \times V}_{k \text{ vezes}} \longrightarrow \mathbb{R}.$$

Um **l -tensor contravariante** em V (ou *tensor contravariante de ordem l*) é uma função real l -linear

$$T : \underbrace{V^* \times \dots \times V^*}_{l \text{ vezes}} \longrightarrow \mathbb{R}.$$

Um **tensor do tipo (k, l)** é um tensor k -covariante e l -contravariante, isto é, uma função real multilinear

$$T : \underbrace{V \times \dots \times V}_{k \text{ vezes}} \times \underbrace{V^* \times \dots \times V^*}_{l \text{ vezes}} \longrightarrow \mathbb{R}.$$

O espaço vetorial real dos k -tensores covariantes sobre V será denotado por $T^k(V)$; o espaço vetorial dos l -tensores contravariantes sobre V será denotado por $T_l(V)$ e o espaço vetorial dos (k, l) tensores sobre V será denotado por $T_l^k(V)$. Estes espaços vetoriais são chamados **espaços tensoriais**. □

9.17 Exemplo. Um 1-tensor covariante é simplesmente um covetor. Formas bilineares, entre elas o produto interno, são 2-tensores covariantes. Determinantes são n -tensores covariantes em \mathbb{R}^n . □

Algumas identificações naturais (isto é, independente de especificação de bases):

- 0-tensores são números reais:

$$T^0(V) = \mathbb{R};$$

- tensores do tipo $(k, 0)$ são k -tensores covariantes:

$$T_0^k(V) = T^k(V);$$

- tensores do tipo $(0, l)$ são l -tensores contravariantes:

$$T_l^0(V) = T_l(V);$$

- 1-tensores covariantes são covetores:

$$T^1(V) = V^*$$

- 1-tensores contravariantes são vetores:

$$T_1(V) = V^{**} = V.$$

9.18 Proposição. Seja $\text{End}(V)$ o espaço vetorial dos operadores lineares sobre V . Então existe um isomorfismo natural

$$T_1^1(V) \cong \text{End}(V).$$

Prova. Um isomorfismo natural $\Phi : \text{End}(V) \rightarrow T_1^1(V)$ pode ser definido por

$$\Phi(A)(v, \omega) = \omega(Av).$$

■

9.19 Proposição. Considere o espaço vetorial $\mathcal{L}\left(V^k \times (V^*)^l; V\right)$ das aplicações multilineares

$$T : \underbrace{V \times \dots \times V}_{k \text{ vezes}} \times \underbrace{V^* \times \dots \times V^*}_{l \text{ vezes}} \rightarrow V.$$

Então existe um isomorfismo natural

$$T_{l+1}^k(V) \cong \mathcal{L}\left(V^k \times (V^*)^l; V\right).$$

Prova. Este pode ser definido por

$$(\Phi T)(v_1, \dots, v_k, \omega^1, \dots, \omega^l, \omega^{l+1}) = \omega_{l+1}(T(v_1, \dots, v_k, \omega^1, \dots, \omega^l)).$$

■

9.5.2 Produto Tensorial

9.20 Definição. Sejam T e S tensores de tipos (k, l) e (p, q) , respectivamente. Seu **produto tensorial** é o tensor $T \otimes S$ do tipo $(k+p, l+q)$ definido por

$$(T \otimes S)(v_1, \dots, v_{k+p}, \omega^1, \dots, \omega^{l+q}) = T(v_1, \dots, v_k, \omega^1, \dots, \omega^l) S(v_{k+1}, \dots, v_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}).$$

□

9.21 Exemplo. Sejam ω^1, ω^2 dois covetores (1-tensores covariantes). Então

$$\omega^1 \otimes \omega^2(v_1, v_2) = \omega^1(v_1) \omega^2(v_2)$$

é um 2-tensor covariante (uma forma bilinear). □

Usando produtores tensoriais, podemos obter uma base para o espaço tensorial $T_l^k(V)$:

9.22 Proposição. Se

$$\mathcal{B} = \{e_1, \dots, e_n\}$$

é uma base para o espaço vetorial V e

$$\mathcal{B}^* = \{e^1, \dots, e^n\}$$

é a correspondente base dual para V^* , então

$$\mathcal{B}_l^k = \{e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l}\}_{\substack{1 \leq i_1, \dots, i_k \leq n \\ 1 \leq j_1, \dots, j_l \leq n}} \quad (9.26)$$

é uma base para o espaço tensorial $T_l^k(V)$. Além disso, qualquer tensor $T \in T_l^k(V)$ se escreve na forma

$$T = T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l}, \quad (9.27)$$

onde

$$T_{i_1 \dots i_k}^{j_1 \dots j_l} = T(e_{i_1}, \dots, e_{i_k}, e^{j_1}, \dots, e^{j_l}). \quad (9.28)$$

Em particular, $\dim T_l^k(V) = n^{k+l}$.

Prova. Primeiro mostraremos que \mathcal{B}_l^k gera o espaço tensorial $T_l^k(V)$. Seja $T \in T_l^k(V)$ um tensor qualquer e defina

$$T_{i_1 \dots i_k}^{j_1 \dots j_l} = T(e_{i_1}, \dots, e_{i_k}, e^{j_1}, \dots, e^{j_l}).$$

Se $v_1, \dots, v_k \in V$, $\omega^1, \dots, \omega^l \in V^*$ são vetores e covetores arbitrários, expressos em coordenadas por

$$v_r = v_r^{i_r} e_{i_r} \quad \text{e} \quad \omega^s = \omega_{j_s}^s e^{j_s}$$

para $r = 1, \dots, k$ e $s = 1, \dots, l$, segue da multilinearidade que

$$\begin{aligned} & T(v_1, \dots, v_k, \omega^1, \dots, \omega^l) \\ &= T(v_1^{i_1} e_{i_1}, \dots, v_k^{i_k} e_{i_k}, \omega_{j_1}^1 e^{j_1}, \dots, \omega_{j_l}^l e^{j_l}) \\ &= v_1^{i_1} \dots v_k^{i_k} \omega_{j_1}^1 \dots \omega_{j_l}^l T(e_{i_1}, \dots, e_{i_k}, e^{j_1}, \dots, e^{j_l}) \\ &= T_{i_1 \dots i_k}^{j_1 \dots j_l} v_1^{i_1} \dots v_k^{i_k} \omega_{j_1}^1 \dots \omega_{j_l}^l \\ &= T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1}(v_1) \dots e^{i_k}(v_k) e_{j_1}(\omega^1) \dots e_{j_l}(\omega^l) \\ &= T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l} (v_1, \dots, v_k, \omega^1, \dots, \omega^l). \end{aligned}$$

Para mostrar que \mathcal{B}_l^k é linearmente independente, suponha que exista uma combinação linear nula

$$T = C_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l} = 0$$

para algumas constantes $C_{i_1 \dots i_k}^{j_1 \dots j_l} \in \mathbb{R}$. Como

$$\begin{aligned} & e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l} (e_{r_1}, \dots, e_{r_k}, e^{s_1}, \dots, e^{s_l}) \\ &= e^{i_1}(e_{r_1}) \dots e^{i_k}(e_{r_k}) e_{j_1}(e^{s_1}) \dots e_{j_l}(e^{s_l}) \\ &= \delta_{i_1}^{r_1} \dots \delta_{i_k}^{r_k} \delta_{s_1}^{j_1} \dots \delta_{s_l}^{j_l}, \end{aligned}$$

segue que

$$0 = T(e_{r_1}, \dots, e_{r_k}, e^{s_1}, \dots, e^{s_l}) = C_{r_1 \dots r_k}^{s_1 \dots s_l}$$

para todos os índices $r_1, \dots, r_k, s_1, \dots, s_l = 1, \dots, n$. ■

Este resultado mostra que um tensor é completamente determinado pela sua ação em todas as sequências possíveis de covetores e vetores das bases de V^* e V .

Observe que, se $F \in T_l^k(V)$, $G \in T_q^p(V)$ e $T = F \otimes G \in T_{l+q}^{k+p}(V)$, então

$$\begin{aligned} T_{i_1 \dots i_k i_{k+1} \dots i_{k+p}}^{j_1 \dots j_l j_{l+1} \dots j_{l+q}} &= T(e_{i_1}, \dots, e_{i_k}, e_{i_{k+1}}, \dots, e_{i_{k+p}}, e^{j_1}, \dots, e^{j_l}, e^{j_{l+1}}, \dots, e^{j_{l+q}}) \\ &= F(e_{i_1}, \dots, e_{i_k}, e^{j_1}, \dots, e^{j_l}) G(e_{i_{k+1}}, \dots, e_{i_{k+p}}, e^{j_{l+1}}, \dots, e^{j_{l+q}}) \end{aligned}$$

de modo que

$$T_{i_1 \dots i_k i_{k+1} \dots i_{k+p}}^{j_1 \dots j_l j_{l+1} \dots j_{l+q}} = F_{i_1 \dots i_k}^{j_1 \dots j_l} G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}}. \quad (9.29)$$

9.5.3 Mudança de Base

9.23 Proposição. Sejam $\mathcal{B}_1 = \{e_1, \dots, e_n\}$, $\mathcal{B}_2 = \{f_1, \dots, f_n\}$ duas bases para o espaço vetorial V e $\mathcal{B}_1^* = \{e^1, \dots, e^n\}$, $\mathcal{B}_2^* = \{f^1, \dots, f^n\}$ as respectivas bases duais para V^* . Sejam A a matriz de mudança de coordenadas da base \mathcal{B}_1 para a base \mathcal{B}_2 , e $(A^{-1})^T$ a matriz de mudança de coordenadas da base dual \mathcal{B}_1^* para a base dual \mathcal{B}_2^* , isto é,

$$e_i = A_i^j f_j \quad \text{e} \quad e^k = (A^{-1})_l^k f^l.$$

Sejam

$$\begin{aligned} T &= E_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l} \\ &= F_{i_1 \dots i_k}^{j_1 \dots j_l} f^{i_1} \otimes \dots \otimes f^{i_k} \otimes f_{j_1} \otimes \dots \otimes f_{j_l} \end{aligned}$$

as expressões em coordenadas para um tensor $T \in T_l^k(V)$ em relação a estas bases. Então

$$E_{i_1 \dots i_k}^{j_1 \dots j_l} = A_{i_1}^{r_1} \dots A_{i_k}^{r_k} (A^{-1})_{s_1}^{j_1} \dots (A^{-1})_{s_l}^{j_l} F_{r_1 \dots r_k}^{s_1 \dots s_l}. \quad (9.30)$$

Prova. Segue da última proposição e por multilinearidade que

$$\begin{aligned} E_{i_1 \dots i_k}^{j_1 \dots j_l} &= T(e_{i_1}, \dots, e_{i_k}, e^{j_1}, \dots, e^{j_l}) \\ &= T(A_{i_1}^{r_1} f_{r_1}, \dots, A_{i_k}^{r_k} f_{r_k}, (A^{-1})_{s_1}^{j_1} f^{s_1}, \dots, (A^{-1})_{s_l}^{j_l} f^{s_l}) \\ &= A_{i_1}^{r_1} \dots A_{i_k}^{r_k} (A^{-1})_{s_1}^{j_1} \dots (A^{-1})_{s_l}^{j_l} T(f_{r_1}, \dots, f_{r_k}, f^{s_1}, \dots, f^{s_l}) \\ &= A_{i_1}^{r_1} \dots A_{i_k}^{r_k} (A^{-1})_{s_1}^{j_1} \dots (A^{-1})_{s_l}^{j_l} F_{r_1 \dots r_k}^{s_1 \dots s_l}. \end{aligned}$$

■

9.24 Definição. Seja M uma variedade diferenciável. Para cada $p \in M$ definimos o **espaço tensorial tangente** $T_p^k(T_p M)$ a M em p . Seja $\varphi : U \rightarrow \varphi(U)$ uma carta de uma vizinhança de um ponto $p \in M$. A base coordenada

$$\mathcal{B}_p = \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$$

do espaço tangente $T_p M$ associada à carta φ e sua respectiva base dual

$$\mathcal{B}_p^* = \left\{ dx^1 \Big|_p, \dots, dx^n \Big|_p \right\}$$

dão origem à **base coordenada** associada à carta φ para o espaço tensorial tangente $T_p^k(T_p M)$

$$(\mathcal{B}_p^k)_p = \left\{ dx^{i_1} \Big|_p \otimes \dots \otimes dx^{i_k} \Big|_p \otimes \frac{\partial}{\partial x^{j_1}} \Big|_p \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \Big|_p \right\}_{\substack{1 \leq i_1, \dots, i_k \leq n \\ 1 \leq j_1, \dots, j_l \leq n}} \quad (9.31)$$

□

9.25 Corolário. Seja M uma variedade diferenciável n -dimensional e $\varphi : U \rightarrow \varphi(U), \psi : V \rightarrow \psi(V)$ duas cartas para vizinhanças de $p = \varphi(x) = \psi(y)$ em M . Sejam

$$\begin{aligned}\mathcal{B}_x &= \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}, \\ \mathcal{B}_y &= \left\{ \frac{\partial}{\partial y^1} \Big|_p, \dots, \frac{\partial}{\partial y^n} \Big|_p \right\}\end{aligned}$$

as bases coordenadas de $T_p M$ induzidas pelas cartas φ e ψ , respectivamente, e

$$\begin{aligned}\mathcal{B}_x^* &= \left\{ dx^1 \Big|_p, \dots, dx^n \Big|_p \right\}, \\ \mathcal{B}_y^* &= \left\{ dy^1 \Big|_p, \dots, dy^n \Big|_p \right\}\end{aligned}$$

sus suas respectivas bases duais. Sejam

$$\begin{aligned}T_p &= E_{i_1 \dots i_k}^{j_1 \dots j_l}(p) dx^{i_1} \Big|_p \otimes \dots \otimes dx^{i_k} \Big|_p \otimes \frac{\partial}{\partial x^{j_1}} \Big|_p \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \Big|_p \\ &= F_{i_1 \dots i_k}^{j_1 \dots j_l}(p) dy^{i_1} \Big|_p \otimes \dots \otimes dy^{i_k} \Big|_p \otimes \frac{\partial}{\partial y^{j_1}} \Big|_p \otimes \dots \otimes \frac{\partial}{\partial y^{j_l}} \Big|_p\end{aligned}$$

as expressões em coordenadas para um tensor $T_p \in T_l^k(T_p M)$ em relação a estas bases. Então

$$E_{i_1 \dots i_k}^{j_1 \dots j_l}(p) = \frac{\partial y^{r_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{r_k}}{\partial x^{i_k}} \frac{\partial x^{j_1}}{\partial y^{s_1}} \cdots \frac{\partial x^{j_l}}{\partial y^{s_l}} F_{r_1 \dots r_k}^{s_1 \dots s_l}(p). \quad (9.32)$$

Prova: Segue das Proposições 9.12, 9.15 e 9.23. ■

9.5.4 Traço de Tensores

O traço de uma matriz $A = (A_j^i)_{n \times n}$ é definido por

$$\text{tr } A = A_i^i.$$

A partir disso pode-se definir o traço de um operador linear sobre um espaço vetorial real de dimensão finita como sendo o traço de qualquer uma de suas representações matriciais com respeito a uma base fixada pois pode-se provar que o traço independe da base escolhida, ou seja, que o traço é uma noção independente de coordenadas. Usando o isomorfismo natural entre o espaço vetorial $\text{End}(V)$ dos operadores lineares sobre V e $T_1^1(V)$, podemos definir logo de início o traço para operadores lineares independentemente de coordenadas. Além da vantagem óbvia de se ter uma definição que não se refere a coordenadas, a maior vantagem é que ela será naturalmente generalizada para definir o traço de tensores.

Observe que é uma consequência da Proposição 9.22 que os produtos tensoriais da forma $\omega \otimes v$, $\omega \in V^*$, $v \in V$, geram $T_1^1(V)$; em outras palavras, todo $(1, 1)$ -tensor é uma combinação linear de tais produtos tensoriais.

9.26 Definição. O traço de $(1, 1)$ -tensores é o funcional linear $\text{tr} : T_1^1(V) \rightarrow \mathbb{R}$ definido por

$$\text{tr}(\omega \otimes v) = \omega(v)$$

em produtos tensoriais e estendido linearmente a todo $T_1^1(V)$.

Se $\Phi : \text{End}(V) \rightarrow T_1^1(V)$ é o endomorfismo natural, então o **traço** de um operador linear $A \in \text{End}(V)$ é definido por

$$\text{tr } A = \text{tr}(\Phi(A)).$$

□

9.27 Proposição. Se $T \in T_1^1(V)$ se escreve em coordenadas na forma

$$T = T_i^j e^i \otimes e_j,$$

então

$$\operatorname{tr} T = T_i^i. \quad (9.33)$$

Se $A \in \operatorname{End}(V)$, então

$$\operatorname{tr} A = A_i^i. \quad (9.34)$$

Prova: Por definição,

$$\operatorname{tr} T = T_i^j \operatorname{tr} (e^i \otimes e_j) = T_i^j e^i (e_j) = T_i^j \delta_j^i = T_i^i.$$

Daí, como

$$\operatorname{tr} A = [\Phi(A)]_i^i,$$

e, pela Proposição 9.18,

$$\begin{aligned} [\Phi(A)]_i^j &= \Phi(A)(e_i, e^j) = e^j(Ae_i) = e^j(A_i^k e_k) \\ &= A_i^k e^j(e_k) = A_i^k \delta_{jk} = A_i^j, \end{aligned}$$

segue a segunda expressão. ■

O conceito de traço pode ser generalizado para tensores de qualquer tipo, produzindo uma operação que diminui a ordem total do tensor em 2, 1 para a parte covariante e 1 para a parte contravariante. Antes observe que, dado um tensor T do tipo (k, l) e índices p, q , cada $(k-1, l-1)$ -upla fixada

$$(v_1, \dots, v_{p-1}, v_{p+1}, \dots, v_k, \omega^1, \dots, \omega^{q-1}, \omega^{q+1}, \dots, \omega^l) \in V^{k-1} \times (V^*)^{l-1}$$

define um tensor $S \in T_1^1(V)$, que depende da $(k-1, l-1)$ -upla escolhida, através da expressão

$$S(v, \omega) = T(v_1, \dots, v_{p-1}, v, v_{p+1}, \dots, v_k, \omega^1, \dots, \omega^{q-1}, \omega, \omega^{q+1}, \dots, \omega^l).$$

Em outras palavras, fixados $v_1, \dots, v_{p-1}, v_{p+1}, \dots, v_k, \omega^1, \dots, \omega^{q-1}, \omega^{q+1}, \dots, \omega^l$,

$$T(v_1, \dots, v_{p-1}, \cdot, v_{p+1}, \dots, v_k, \omega^1, \dots, \omega^{q-1}, \cdot, \omega^{q+1}, \dots, \omega^l)$$

é um $(1, 1)$ -tensor.

9.28 Definição. Dado um tensor T do tipo (k, l) e índices p, q , o **traço** de T com respeito aos índices p, q (índice covariante p e índice contravariante q) é o tensor $\operatorname{tr} T$ do tipo $(k-1, l-1)$ definido por

$$\begin{aligned} (\operatorname{tr} T)(v_1, \dots, v_{p-1}, v_p, \dots, v_{k-1}, \omega^1, \dots, \omega^{q-1}, \omega^q, \dots, \omega^{l-1}) \\ = \operatorname{tr} T(v_1, \dots, v_{p-1}, \cdot, v_p, \dots, v_{k-1}, \omega^1, \dots, \omega^{q-1}, \cdot, \omega^q, \dots, \omega^{l-1}). \end{aligned}$$

Se for necessário explicitar os índices em relação aos quais foi tomado o traço, denotaremos $\operatorname{tr}_q^p T$. □

9.29 Proposição. Se $T \in T_l^k(V)$ se escreve em coordenadas na forma

$$T = T_{i_1 \dots i_k}^{j_1 \dots j_l} e^{i_1} \otimes \dots \otimes e^{i_k} \otimes e_{j_1} \otimes \dots \otimes e_{j_l}.$$

então as coordenadas de

$$\operatorname{tr} T = (\operatorname{tr} T)_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} e^{i_1} \otimes \dots \otimes e^{i_{k-1}} \otimes e_{j_1} \otimes \dots \otimes e_{j_{l-1}}$$

são dadas por

$$(\operatorname{tr} T)_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} = T_{i_1 \dots i_{p-1} i i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} i j_q \dots j_{k-1}}. \quad (9.35)$$

Prova: Por definição, se S é o tensor $T(e_{i_1}, \dots, e_{i_{p-1}}, \cdot, e_{i_p}, \dots, e_{i_{k-1}}, e^{j_1}, \dots, e^{j_{q-1}}, \cdot, e^{j_q}, \dots, e^{j_{l-1}})$, então

$$\begin{aligned} (\text{tr } T)_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} &= (\text{tr } T)(e_{i_1}, \dots, e_{i_{p-1}}, e_{i_p}, \dots, e_{i_{k-1}}, e^{j_1}, \dots, e^{j_{q-1}}, e^{j_q}, \dots, e^{j_{l-1}}) \\ &= \text{tr } S \\ &= S_i^i \\ &= T(e_{i_1}, \dots, e_{i_{p-1}}, e_i, e_{i_p}, \dots, e_{i_{k-1}}, e^{j_1}, \dots, e^{j_{q-1}}, e^i, e^{j_q}, \dots, e^{j_{l-1}}) \\ &= T_{i_1 \dots i_{p-1} i i_p \dots i_{k-1}}^{j_1 \dots j_{q-1} i j_q \dots j_{l-1}}. \end{aligned}$$

■

9.6 Fibrados Tensoriais

9.30 Definição. Seja M uma variedade diferenciável de dimensão n com um atlas $\Phi = \{\varphi_\alpha : U_\alpha \rightarrow M\}_{\alpha \in \mathcal{A}}$. O **fibrado (k, l) -tensorial** de M é a variedade diferenciável de dimensão $n + n^{k+l}$

$$T_l^k M = \{(p, T) : p \in M \text{ e } T \in T_l^k(T_p M)\}$$

com um atlas

$$\Psi = \left\{ \psi_\alpha : U_\alpha \times \mathbb{R}^{n^{k+l}} \rightarrow T_l^k TM \right\}_{\alpha \in \mathcal{A}}$$

definido por

$$\begin{aligned} \psi_\alpha &\left(x, \left(T_{i_1 \dots i_k}^{j_1 \dots j_l} \right)_{\substack{i_1, \dots, i_k=1, \dots, n \\ j_1, \dots, j_l=1, \dots, n}} \right) \\ &= \left(\varphi_\alpha(x), T_{i_1 \dots i_k}^{j_1 \dots j_l} dx^{i_1}|_p \otimes \dots \otimes dx^{i_k}|_p \otimes \frac{\partial}{\partial x^{j_1}}|_p \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}}|_p \right). \end{aligned}$$

□

Fibrados tensoriais são fibrados vetoriais (veja o Exercício 1.3). Note que

$$\begin{aligned} T^0 M &= C^\infty(M), \\ T_1 M &= TM, \\ T^1 M &= T^* M, \\ T_0^k M &= T^k M, \\ T_l^0 M &= T_l M. \end{aligned}$$

O fibrado $T^1 M$ é chamado o **fibrado cotangente**.

9.7 Campos Tensoriais

9.31 Definição. Um **campo tensorial** é uma seção do fibrado tensorial. Um **campo tensorial diferenciável** é uma seção diferenciável do fibrado tensorial.

O espaço vetorial dos campos (k, l) -tensoriais diferenciáveis é denotado por $\mathcal{T}_l^k(M)$. □

A menos que seja dito o contrário, lidaremos apenas com campos tensoriais diferenciáveis. Note que

$$\begin{aligned}\mathcal{T}^0(M) &= C^\infty(M), \\ \mathcal{T}_1(M) &= \mathcal{T}(M), \\ \mathcal{T}_0^k(M) &= \mathcal{T}^k(M), \\ \mathcal{T}_l^0(M) &= \mathcal{T}_l(M).\end{aligned}$$

e $\mathcal{T}^1 M$ é o espaço vetorial dos campos covetoriais.

9.32 Proposição. *Seja $T : M \rightarrow T_l^k M$ um campo tensorial. Para cada carta $\varphi : U \rightarrow V$ uma vizinhança V de M , denote a base coordenada associada para o espaço tensorial $T_l^k(T_p M)$ por*

$$(\mathcal{B}_l^k)_p = \left\{ dx^{i_1} \Big|_p \otimes \dots \otimes dx^{i_k} \Big|_p \otimes \frac{\partial}{\partial x^{j_1}} \Big|_p \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \Big|_p \right\}_{\substack{1 \leq i_1, \dots, i_k \leq n \\ 1 \leq j_1, \dots, j_l \leq n}}$$

para todo $p \in V$, de modo que nesta base coordenada o campo tensorial T se escreve na forma

$$T_p = T_{i_1 \dots i_k}^{j_1 \dots j_l}(p) dx^{i_1} \Big|_p \otimes \dots \otimes dx^{i_k} \Big|_p \otimes \frac{\partial}{\partial x^{j_1}} \Big|_p \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \Big|_p. \quad (9.36)$$

Então T é um campo tensorial diferenciável se e somente se para toda carta φ as funções $T_{i_1 \dots i_k}^{j_1 \dots j_l} : V \rightarrow \mathbb{R}$ são diferenciáveis para todos os índices $i_1, \dots, i_k, j_1, \dots, j_l = 1, \dots, n$.

9.8 Exercícios

9.33 Exercício. Defina explicitamente o *fibrado cotangente* e mostre que ele é um fibrado vetorial. Defina explicitamente o conceito de *campos covetoriais*.

9.34 Exercício. Mostre que o fibrado tensorial definido pela Definição 9.30 é de fato uma variedade diferenciável.

9.35 Exercício. Mostre que um fibrado tensorial é um fibrado vetorial.

9.36 Exercício. Demonstre a Proposição 9.32.

9.37 Exercício. Seja $T : M \rightarrow T_l^k M$ uma seção do fibrado tensorial. Mostre que T é diferenciável (e, portanto, um campo tensorial diferenciável) se e somente se para toda vizinhança $V \subset M$ e para todos os campos vetoriais X_1, \dots, X_k e para todas os campos covetoriais $\omega^1, \dots, \omega^l$ a função $T(X_1, \dots, X_k, \omega^1, \dots, \omega^l) : V \rightarrow \mathbb{R}$ definida por

$$T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)(p) = T_p(X_1(p), \dots, X_k(p), \omega^1(p), \dots, \omega^l(p))$$

é diferenciável.

Capítulo 10

Variedades Métricas

10.1 Definição e Exemplos

10.1 Definição. Seja M uma variedade diferenciável de dimensão n . Uma **métrica** (ou **tensor métrica**) de índice k em M é um campo diferenciável tensorial 2-covariante g tal que g_p é um produto interno com o mesmo índice k para todo $p \in M$. Uma variedade diferenciável munida de uma métrica g é chamada uma **variedade métrica**.

Uma variedade diferenciável M com uma métrica g dada é chamada uma **variedade métrica**.

Se g é um produto interno positivo definido (ou seja, se o seu índice é 0), dizemos que g é uma **métrica riemanniana** e M é uma **variedade riemanniana**.

Se o índice de g é diferente de zero, dizemos que g é uma **métrica semi-riemanniana** e M é uma **variedade semi-riemanniana**.

Se g é um produto interno de Lorentz (ou seja, se o índice é 1), dizemos que g é uma **métrica de Lorentz** e M é uma **variedade de Lorentz**. \square

Em outras palavras, uma métrica em M é uma aplicação que associa a cada ponto $p \in M$ um produto interno

$$g_p = \langle \cdot, \cdot \rangle_p$$

no espaço tangente $T_p M$ que varia diferenciavelmente com p no sentido de que se $\varphi : U \rightarrow V$ é uma carta para uma vizinhança coordenada V de M e $\mathcal{B}_p = \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$ é a base coordenada de $T_p M$ associada a esta carta para cada $p \in V$, então as funções $g_{ij} : V \rightarrow \mathbb{R}$

$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x^i} \Big|_p, \frac{\partial}{\partial x^j} \Big|_p \right\rangle_p \quad (10.1)$$

são diferenciáveis. De fato, escrevendo o tensor métrica em coordenadas, temos

$$g_p = g_{ij}(p) dx^i \Big|_p \otimes dx^j \Big|_p, \quad (10.2)$$

e as funções componentes g_{ij} do tensor métrica g são diferenciáveis para toda parametrização φ se e somente se g é diferenciável.

Omitindo o símbolo do ponto de atuação p , como frequentemente faremos, escrevemos simplesmente

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \quad (10.3)$$

e notamos que a simetria do tensor métrica implica que

$$g_{ij} = g_{ji}. \quad (10.4)$$

Em particular, quando consideramos a matriz

$$G = (g_{ij}) \quad (10.5)$$

segue que G é uma matriz simétrica. Observe que devido à simetria existem apenas

$$\frac{n(n+1)}{2}$$

componentes potencialmente distintos do tensor métrica, ao invés dos n^2 componentes distintos para um tensor 2-covariante geral.

Usando o *produto simétrico* de tensores (veja [Lee 1], Cap. 12, p. 315), que no caso de covetores é simplesmente

$$\omega\eta := \frac{1}{2}(\omega \otimes \eta + \eta \otimes \omega) \quad (10.6)$$

e a simetria do tensor métrica, podemos escrever a expressão

$$g = g_{ij}dx^i \otimes dx^j$$

na forma mais familiar

$$g = g_{ij}dx^i dx^j, \quad (10.7)$$

já que

$$\begin{aligned} g &= g_{ij}dx^i \otimes dx^j = \frac{1}{2}(g_{ij} + g_{ji})dx^i \otimes dx^j \\ &= \frac{1}{2}g_{ij}dx^i \otimes dx^j + \frac{1}{2}g_{ji}dx^i \otimes dx^j \\ &= \frac{1}{2}g_{ij}dx^i \otimes dx^j + \frac{1}{2}g_{ij}dx^j \otimes dx^i \quad (\text{permutando os índices } i, j) \\ &= g_{ij}\frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i) \\ &= g_{ij}dx^i dx^j. \end{aligned}$$

Estritamente falando, uma variedade métrica é um par (M, g) , onde M é uma variedade diferenciável e g a métrica riemanniana, já que uma mesma variedade diferenciável pode admitir diferentes métricas riemannianas, como veremos no decorrer deste texto. Contudo, quando não houver perigo de confusão, nós vamos nos referir à variedade riemanniana simplesmente por M .

10.2 Exemplo (Métrica Euclidiana). A variedade riemanniana mais simples é o **espaço euclidiano** \mathbb{R}^n com a métrica euclidiana $g_{ij} = \delta_{ij}$, que denotaremos \mathbb{E}^n . \square

10.3 Exemplo (Métrica Semieuclidiana). A variedade semi-riemanniana mais simples de índice k é o **espaço semi-euclidiano** \mathbb{R}^n com a métrica semieuclidiana $g_{ij} = \eta_{ij}^k$, que denotaremos \mathbb{M}_k^n . \square

10.4 Exemplo (Métrica de Lorentz). A variedade de Lorentz mais simples de índice 1 é o **espaço de Minkowski n-dimensional** \mathbb{R}^n , $n \geq 2$, com a métrica de Lorentz $g_{ij} = \eta_{ij}$. Denotaremos este espaço por \mathbb{M}^n , enquanto que o espaço-tempo de Minkowski da relatividade especial \mathbb{M}^4 continuará a ser denotado geralmente apenas por \mathbb{M} . \square

10.5 Proposição. *Toda variedade diferenciável possui uma métrica riemanniana.*

Prova: Seja $\{\varphi_\alpha : U_\alpha \rightarrow V_\alpha\}_\alpha$ um atlas para M e $\{f_\alpha\}_\alpha$ uma partição da unidade de M subordinada à cobertura $\{V_\alpha\}_\alpha$.

Em cada V_α podemos definir uma métrica riemanniana, aquela induzida pela carta: dado $p \in M$ e vetores $v, w \in T_p M$, eles se escrevem em coordenadas com relação à base $\mathcal{B}_p = \left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right\}$ associada à carta φ_α por

$$v = v^i \frac{\partial}{\partial x^i} \Big|_p \quad \text{e} \quad w = w^j \frac{\partial}{\partial x^j} \Big|_p$$

e definimos o produto interno

$$\langle v, w \rangle_p^\alpha = v_i w_i.$$

Esta é uma métrica riemanniana na subvariedade V_α com $g_{ij}^\alpha = \delta_{ij}$. Para obter uma métrica riemanniana global em M , usamos a partição da unidade, definindo

$$\langle v, w \rangle_p = \sum_\alpha f_\alpha(p) \langle v, w \rangle_p^\alpha.$$

De fato, esta soma é finita em uma vizinhança de p , portanto define um tensor diferenciável em M ; além disso, como uma combinação linear finita positiva de produtos internos é um produto interno, ela define um produto interno em $T_p M$. ■

Observe que a demonstração acima não funciona para métricas semi-riemannianas. De fato, uma combinação linear finita de métricas semi-riemannianas, mesmo positiva, não é necessariamente uma métrica semi-riemanniana. Por exemplo, a soma das métricas semi-riemannianas em \mathbb{R}^4

$$\begin{aligned} \langle v, w \rangle_1 &= v_1 w_1 + v_2 w_2 - v_3 w_3 - v_4 w_4, \\ \langle v, w \rangle_2 &= -v_1 w_1 - v_2 w_2 + v_3 w_3 + v_4 w_4, \end{aligned}$$

é o tensor nulo. Mais ainda, a soma das métricas de Lorentz em \mathbb{R}^4

$$\begin{aligned} \langle v, w \rangle_1 &= -v_1 w_1 + v_2 w_2 + v_3 w_3 + v_4 w_4, \\ \langle v, w \rangle_2 &= v_1 w_1 - v_2 w_2 + v_3 w_3 + v_4 w_4, \end{aligned}$$

é o tensor

$$T(v, w) = 2(v_3 w_3 + v_4 w_4),$$

que é degenerado, pois se $v_0 = (1, 1, 0, 0)$, então

$$T(v_0, w) = 0$$

para todo $w \in \mathbb{R}^4$. A realidade é que existem variedades diferenciáveis que não admitem métricas semi-riemannianas. Por exemplo, é possível provar, usando técnicas de Topologia Diferencial, que esferas de dimensão par não admitem métricas de Lorentz (em particular \mathbb{S}^4 não pode ser uma variedade de Lorentz; o espaço-tempo não pode ser uma esfera 4-dimensional); mas geralmente, é possível provar que toda variedade diferenciável compacta de dimensão ímpar possui uma métrica de Lorentz e que uma variedade diferenciável compacta de dimensão par admite uma métrica de Lorentz se e somente se ela tem característica de Euler nula (veja [Steenrod], Theorem 40.13, p. 207). Por outro lado, todas as variedades diferenciáveis não compactas admitem métricas de Lorentz (veja [BEE], p. 50).

10.6 Exemplo (Métrica Produto). Se (M_1, g_1) e (M_2, g_2) são duas variedades métricas, então definimos a **métrica produto** $g = g_1 \oplus g_2$ na variedade produto $M_1 \times M_2$ por

$$\langle (v_1, v_2), (w_1, w_2) \rangle_{(p_1, p_2)} = \langle v_1, w_1 \rangle_{p_1} + \langle v_2, w_2 \rangle_{p_2} \quad (10.8)$$

para todos $(v_1, v_2), (w_1, w_2) \in T_{p_1}M_1 \oplus T_{p_2}M_2 \cong T_{(p_1, p_2)}(M_1 \times M_2)$. É fácil ver que g é simétrica. Para verificar a sua não-degenericidade, suponha que

$$\langle (v_1, v_2), (w_1, w_2) \rangle_{(p_1, p_2)} = 0$$

para todo $(w_1, w_2) \in T_{p_1}M_1 \oplus T_{p_2}M_2$. Então em particular

$$\begin{aligned}\langle (v_1, v_2), (0, w_2) \rangle_{(p_1, p_2)} &= 0, \\ \langle (v_1, v_2), (w_1, 0) \rangle_{(p_1, p_2)} &= 0.\end{aligned}$$

para todos $w_1 \in T_{p_1}M_1$ e para todo $w_2 \in T_{p_2}M_2$. Como

$$\begin{aligned}\langle (v_1, v_2), (0, w_2) \rangle_{(p_1, p_2)} &= \langle v_1, 0 \rangle_{p_1} + \langle v_2, w_2 \rangle_{p_2} = \langle v_2, w_2 \rangle_{p_2}, \\ \langle (v_1, v_2), (w_1, 0) \rangle_{(p_1, p_2)} &= \langle v_1, w_1 \rangle_{p_1} + \langle v_2, 0 \rangle_{p_2} = \langle v_1, w_1 \rangle_{p_1},\end{aligned}$$

segue que

$$\langle v_1, w_1 \rangle_{p_1} = 0$$

para todo $w_1 \in T_{p_1}M_1$ e

$$\langle v_2, w_2 \rangle_{p_2} = 0$$

para todo $w_2 \in T_{p_2}M_2$, e a não-degenericidade de g_1, g_2 implica que

$$v_1 = v_2 = 0.$$

Observe que a matriz associada à métrica G é a matriz diagonal em blocos

$$G = \left[\begin{array}{cc} G_1 & 0 \\ 0 & G_2 \end{array} \right] = \left[\begin{array}{cc} (g_1)_{ij} & 0 \\ 0 & (g_2)_{ij} \end{array} \right].$$

□

10.7 Definição. Sejam M, N variedades métricas. Um difeomorfismo $F : M \rightarrow N$ é uma **isometria** se

$$\langle v, w \rangle_p = \langle dF_p v, dF_p w \rangle_{F(p)} \quad (10.9)$$

para todo $p \in M$ e para todos $v, w \in T_p M$. Se existir uma isometria entre M e N , dizemos que M e N são **isométricas**.

Dizemos que M e N são **localmente isométricas** se para todo $p \in M$ existe uma vizinhança V_p de p em M e uma isometria $F : V_p \rightarrow F(V_p) \subset N$ e reciprocamente para todo $q \in N$ existe uma vizinhança W_q de q em N e uma isometria $G : W_q \rightarrow G(W_q) \subset M$.

Dizemos que uma variedade métrica (M, g) é **plana**, se ela é localmente isométrica ao espaço euclidiano ou ao espaço semi-euclidiano. □

Observe que o conjunto das isometrias em uma variedade métrica possui uma estrutura natural de grupo em que o produto de isometrias é definido como a composição das aplicações. Este grupo é denotado por

$$\text{Isom}(M).$$

Observe que o pullback de um difeomorfismo $F : M \rightarrow N$ em que N é uma variedade semi-riemanniana não define necessariamente uma métrica semi-riemanniana em M , pois poderíamos até mesmo ter (na pior hipótese) que $T_p M$ seja um subespaço de vetores nulos para o 2-tensor $g = F^*h$. Por exemplo, isso é exatamente o que acontece na imersão $F : \mathbb{R} \rightarrow \mathcal{M}^n$ definida por $F(t) = (t, t, x_3, \dots, x_n)$. Além disso, se isso fosse verdade, pelo Teorema do Mergulho de Whitney, que afirma que toda variedade diferenciável de dimensão n pode ser mergulhada em \mathbb{R}^{2n} (um mergulho é uma imersão injetiva), toda variedade diferenciável teria uma métrica semi-riemanniana. $g = F^*h$ será uma métrica semi-riemanniana em M se e somente se $g|_{T_p M}$ é não-degenerado e tem o mesmo índice para todo $p \in M$.

10.8 Exemplo. O grupo de isometrias de \mathbb{E}^n consiste das composições de operadores ortogonais e translações. O grupo de isometrias de \mathbb{M} é o grupo de Poincaré. \square

É fácil ver que *isometria* é uma relação de equivalência na classe das variedades métricas. Geometria Riemanniana é principalmente o estudo das propriedades que são invariantes por isometrias. Uma excelente referência para o estudo de grupos de isometrias de variedades riemannianas é [Kobayashi].

Dizemos que uma aplicação diferenciável $F : M \rightarrow N$ entre variedades diferenciáveis é uma **imersão** se dF_p é injetiva para todo $p \in M$.

10.9 Definição. Sejam M uma variedade diferenciável, (N, h) uma variedade riemanniana e $F : M \rightarrow N$ uma imersão. A **métrica induzida** por F em M (também chamada a **métrica do pullback**), denotada por

$$g = F^*h,$$

é definida por

$$\langle v, w \rangle_p := \langle dF_p v, dF_p w \rangle_{F(p)} \quad (10.10)$$

para todo $p \in M$ e para todos $v, w \in T_p M$. \square

Com esta métrica definida em M , a imersão F torna-se uma **imersão isométrica**. Na linguagem do pullback, um difeomorfismo F entre duas variedades métricas (M, g) e (N, h) é uma isometria se

$$g = F^*h.$$

10.10 Exemplo (Superfícies n -dimensionais em \mathbb{R}^N). Seja $M \subset \mathbb{R}^{n+k}$ uma variedade diferenciável de dimensão n , isto é, uma superfície n -dimensional. A aplicação inclusão $i : M \rightarrow \mathbb{R}^{n+k}$ é uma imersão, de modo que, se assumirmos a métrica euclidiana em \mathbb{R}^{n+k} , ela induz em M uma métrica riemanniana. Neste caso, a inclusão passa a ser uma imersão isométrica. Daí, como a diferencial di_p da inclusão é a inclusão natural de $T_p M$ em \mathbb{R}^{n+k} , segue que

$$\langle v, w \rangle_p = \langle v, w \rangle_{\mathbb{R}^{n+k}} \quad (10.11)$$

onde $\langle \cdot, \cdot \rangle_{\mathbb{R}^{n+k}}$ é o produto interno canônico de \mathbb{R}^{n+k} . Uma demonstração alternativa de que toda variedade diferenciável possui uma métrica riemanniana segue então do Teorema do Mergulho de Whitney: a métrica induzida pela métrica euclidiana em \mathbb{R}^n . Diferentes cartas podem ser usadas para a mesma superfície n -dimensional, cada uma dando origem a componentes g_{ij} mais ou menos simples.

Um exemplo de superfície n -dimensional é o gráfico de uma função real. Se $U \subset \mathbb{R}^n$ é um aberto e $f : U \rightarrow \mathbb{R}$ é uma função, então o **gráfico** de f

$$\text{graf}(f) = \{(x, f(x)) : x \in U\}$$

é uma variedade diferenciável com a topologia induzida de \mathbb{R}^{n+1} de dimensão n . Uma carta global para o gráfico de f é $\varphi : \mathbb{R}^n \rightarrow \text{graf}(f)$ definida por

$$\varphi(x^1, \dots, x^n) = (x^1, \dots, x^n, f(x^1, \dots, x^n)).$$

Como

$$\frac{\partial \varphi^k}{\partial x^j} = \begin{cases} \delta_{kj} & \text{se } k \neq n+1, \\ \frac{\partial f}{\partial x^j} & \text{se } k = n+1, \end{cases}$$

ou seja,

$$\frac{\partial \varphi}{\partial x^j}(x) = \left(0, \dots, 0, \underset{j}{1}, 0, \dots, 0, \frac{\partial f}{\partial x^j}(x) \right)$$

segue que

$$g_{ij}(x) = \left\langle \frac{\partial \varphi}{\partial x^i}(x), \frac{\partial \varphi}{\partial x^j}(x) \right\rangle_{(x, f(x))} = \left\langle \frac{\partial \varphi}{\partial x^i}(x), \frac{\partial \varphi}{\partial x^j}(x) \right\rangle_{\mathbb{R}^{n+1}} = \delta_{ij} + \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j}. \quad (10.12)$$

Outro exemplo de superfície 2-dimensional é uma **superfície de revolução gerada por uma curva**. Especificamente seja $\gamma : I \rightarrow \mathbb{R}^2$, $\gamma(t) = (\alpha(t), \beta(t))$ uma curva parametrizada regular tal que $\beta'(t) \neq 0$ para todo $t \in I$; podemos imaginar γ contida no plano yz definindo

$$\gamma(t) = (0, \alpha(t), \beta(t)).$$

Se girarmos esta curva ao redor do eixo z obteremos uma superfície parametrizada regular S . A imagem de S é a imagem da aplicação $\varphi : I \times \mathbb{R} \rightarrow \mathbb{R}^3$ dada por

$$\varphi(t, \theta) = (\alpha(t) \cos \theta, \alpha(t) \sin \theta, \beta(t));$$

a partir de φ podemos obter cartas locais restringindo o parâmetro θ a um intervalo aberto de comprimento 2π . Daí,

$$\begin{aligned}\frac{\partial \varphi}{\partial t}(t, \theta) &= (\alpha'(t) \cos \theta, \alpha'(t) \sin \theta, \beta'(t)), \\ \frac{\partial \varphi}{\partial \theta}(t, \theta) &= (-\alpha(t) \sin \theta, \alpha(t) \cos \theta, 0),\end{aligned}$$

onde

$$\begin{aligned}g_{11}(t, \theta) &= \left\langle \frac{\partial \varphi}{\partial t}(t, \theta), \frac{\partial \varphi}{\partial t}(t, \theta) \right\rangle_{\varphi(t, \theta)} = \left\langle \frac{\partial \varphi}{\partial t}, \frac{\partial \varphi}{\partial t} \right\rangle_{\mathbb{R}^3} = [\alpha'(t)]^2 + [\beta'(t)]^2, \\ g_{12}(t, \theta) &= \left\langle \frac{\partial \varphi}{\partial t}(t, \theta), \frac{\partial \varphi}{\partial \theta}(t, \theta) \right\rangle_{\varphi(t, \theta)} = \left\langle \frac{\partial \varphi}{\partial t}, \frac{\partial \varphi}{\partial \theta} \right\rangle_{\mathbb{R}^3} = 0, \\ g_{22}(t, \theta) &= \left\langle \frac{\partial \varphi}{\partial \theta}(t, \theta), \frac{\partial \varphi}{\partial \theta}(t, \theta) \right\rangle_{\varphi(t, \theta)} = \left\langle \frac{\partial \varphi}{\partial \theta}, \frac{\partial \varphi}{\partial \theta} \right\rangle_{\mathbb{R}^3} = [\alpha(t)]^2.\end{aligned}$$

Portanto

$$G(t, \theta) = \begin{bmatrix} [\alpha'(t)]^2 + [\beta'(t)]^2 & 0 \\ 0 & [\alpha(t)]^2 \end{bmatrix}.$$

□

10.11 Exemplo (Esfera). A métrica euclidiana induz uma métrica na **esfera** de raio r

$$\mathbb{S}_r^n = \left\{ x \in \mathbb{R}^{n+1} : \|x\|^2 = (x^1)^2 + \dots + (x^{n+1})^2 = r^2 \right\}$$

que chamaremos a *métrica canônica* de \mathbb{S}_r^n . Denotaremos a esfera unitária por \mathbb{S}^n , simplesmente. Vamos ver os coeficientes g_{ij} para diferentes cartas da esfera.

a) Como gráfico de função:

O hemisfério superior da esfera é o gráfico da função $f : B_r \subset \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f(x^1, \dots, x^n) = \sqrt{r^2 - (x^1)^2 - \dots - (x^n)^2}$. Como

$$\frac{\partial f}{\partial x^i}(x) = \frac{-x_i}{\sqrt{r^2 - \|x\|^2}},$$

segue que

$$g_{ij}(x) = \delta_{ij} + \frac{x_i x_j}{r^2 - \|x\|^2}. \tag{10.13}$$

Similarmente para o hemisfério inferior. Estas cartas não cobrem o equador da esfera.

b) Como superfície de revolução:

A parametrização da esfera de raio r como superfície de revolução é

$$(x, y, z)(\phi, \theta) = (r \sin \phi \cos \theta, r \sin \phi \sin \theta, r \cos \phi).$$

Segue que

$$G(\phi, \theta) = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \sin^2 \phi \end{bmatrix}.$$

c) *Através da projeção estereográfica:*

Na projeção estereográfica a partir do polo norte $N = (0, \dots, 0, r)$, a reta a partir de N que intercepta o plano $x^{n+1} = 0$ em um ponto $\tilde{x} = (x^1, \dots, x^n, 0)$, intercepta a esfera em um ponto $\varphi(x)$. Portanto, a carta *projeção estereográfica a partir do polo norte* $\varphi : \mathbb{R}^n \rightarrow \mathbb{S}_r^n \setminus \{N\}$ é definida por

$$\varphi(x) = N + t(\tilde{x} - N) = (tx^1, \dots, tx^n, (1-t)r)$$

onde $t > 0$ é tal que $\|\varphi(x)\| = r$. Ou seja, t é tal que

$$t^2 \|x\|^2 + (1-t)^2 r^2 = r^2,$$

onde

$$t = \frac{2r^2}{r^2 + \|x\|^2}.$$

Logo,

$$\varphi(x^1, \dots, x^n) = \left(\frac{2r^2 x^1}{r^2 + \|x\|^2}, \dots, \frac{2r^2 x^n}{r^2 + \|x\|^2}, r \frac{\|x\|^2 - r^2}{r^2 + \|x\|^2} \right), \quad (10.14)$$

onde

$$\frac{\partial \varphi^k}{\partial x^j}(x) = \begin{cases} \frac{2r^2 \delta_{kj}}{r^2 + \|x\|^2} - \frac{4r^2 x^j x^k}{(r^2 + \|x\|^2)^2} & \text{se } k \neq n+1, \\ \frac{4r^3 x^j}{(r^2 + \|x\|^2)^2} & \text{se } k = n+1. \end{cases}$$

Segue que as componentes do tensor métrico nas coordenadas dadas pela carta φ são

$$\begin{aligned} g_{ij}(x) &= \left\langle \frac{\partial \varphi}{\partial x^i}(x), \frac{\partial \varphi}{\partial x^j}(x) \right\rangle_{\varphi(x)} = \left\langle \sum_{k=1}^{n+1} \frac{\partial \varphi^k}{\partial x^i}(x) e_k, \sum_{l=1}^{n+1} \frac{\partial \varphi^l}{\partial x^j}(x) e_l \right\rangle_{\mathbb{R}^{n+1}} \\ &= \sum_{k=1}^{n+1} \frac{\partial \varphi^k}{\partial x^i} \frac{\partial \varphi^l}{\partial x^j} \langle e_k, e_l \rangle_{\mathbb{R}^{n+1}} = \sum_{l=1}^{n+1} \frac{\partial \varphi^k}{\partial x^i} \frac{\partial \varphi^l}{\partial x^j} \delta_{kl} = \sum_{k=1}^{n+1} \frac{\partial \varphi^k}{\partial x^i} \frac{\partial \varphi^k}{\partial x^j} \\ &= \sum_{k=1}^n \left(\frac{2r^2 \delta_{ki}}{r^2 + \|x\|^2} - \frac{4r^2 x^i x^k}{(r^2 + \|x\|^2)^2} \right) \left(\frac{2r^2 \delta_{kj}}{r^2 + \|x\|^2} - \frac{4r^2 x^j x^k}{(r^2 + \|x\|^2)^2} \right) + \frac{16r^6 x^i x^j}{(r^2 + \|x\|^2)^4} \\ &= \sum_{k=1}^n \left(\frac{4r^4 \delta_{ki} \delta_{kj}}{(r^2 + \|x\|^2)^2} - \frac{8r^4 (\delta_{ki} x^j x^k + \delta_{kj} x^i x^k)}{(r^2 + \|x\|^2)^3} + \frac{16r^4 x^i x^j (x^k)^2}{(r^2 + \|x\|^2)^4} \right) + \frac{16r^6 x^i x^j}{(r^2 + \|x\|^2)^4} \\ &= \frac{4r^4 \delta_{ij}}{(r^2 + \|x\|^2)^2} - \frac{16r^4 x^i x^j}{(r^2 + \|x\|^2)^3} + \frac{16r^4 x^i x^j \|x\|^2}{(r^2 + \|x\|^2)^4} + \frac{16r^6 x^i x^j}{(r^2 + \|x\|^2)^4} \\ &= \frac{4r^4 \delta_{ij}}{(r^2 + \|x\|^2)^2} - \frac{16r^4 x^i x^j}{(r^2 + \|x\|^2)^3} + \frac{16r^4 x^i x^j}{(r^2 + \|x\|^2)^3} \\ &= \frac{4r^4 \delta_{ij}}{(r^2 + \|x\|^2)^2}. \end{aligned}$$

Vamos anotar este resultado para futura referência:

$$g_{ij}(x) = \frac{4r^4}{(r^2 + \|x\|^2)^2} \delta_{ij}. \quad (10.15)$$

Observe que

$$G(x) = \begin{bmatrix} \frac{4r^4}{(r^2 + \|x\|^2)^2} & & \\ & \ddots & \\ & & \frac{4r^4}{(r^2 + \|x\|^2)^2} \end{bmatrix} = \frac{4r^4}{(r^2 + \|x\|^2)^2} I.$$

Usando a projeção estereográfica a partir do polo sul obtemos duas cartas que cobrem toda a esfera. \square

10.12 Exemplo (Espaço Hiperbólico). Considere o semiespaço superior de \mathbb{R}^n

$$\mathbb{H}^n = \{(x^1, \dots, x^n) \in \mathbb{R}^n : x_n > 0\}.$$

Com a topologia induzida como aberto de \mathbb{R}^n , \mathbb{H}^n é uma superfície diferenciável de dimensão n . Se definirmos diretamente em \mathbb{H}^n a métrica

$$g_{ij}(x^1, \dots, x^n) = \frac{\delta_{ij}}{(x^n)^2}, \quad (10.16)$$

então \mathbb{H}^n é uma variedade riemanniana chamada o **espaço hiperbólico** n -dimensional. Observe que

$$G = \begin{bmatrix} \frac{1}{(x^n)^2} & & \\ & \ddots & \\ & & \frac{1}{(x^n)^2} \end{bmatrix} = \frac{1}{(x^n)^2} I.$$

\square

10.13 Exemplo (Toros). O **toro** mergulhado em \mathbb{R}^3 é uma superfície de revolução gerada pelo círculo. Tomando o círculo com centro em $(R, 0)$ e raio $r < R$ com parametrização $\gamma(t) = (R + r \cos t, r \sin t)$ obtemos a parametrização para o toro bidimensional como superfície de revolução

$$\varphi(t, \theta) = ((R + r \cos t) \cos \theta, (R + r \cos t) \sin \theta, r \sin t)$$

cuja respectiva métrica é dada por

$$G(t, \theta) = \begin{bmatrix} r^2 & 0 \\ 0 & (R + r \cos t)^2 \end{bmatrix}.$$

Outra métrica induzida de \mathbb{R}^N importante para o toro, não localmente isométrica à métrica dada acima (como veremos depois) é a **métrica plana** do toro: considerando o toro como a superfície n -dimensional $\mathbb{T}^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1 \subset \mathbb{R}^{2n}$, a métrica euclidiana de \mathbb{R}^{2n} induz uma métrica no toro da seguinte forma. Escrevendo

$$\mathbb{T}^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1 = \left\{ x \in \mathbb{R}^{2n} : (x^1)^2 + (x^2)^2 = (x^3)^2 + (x^4)^2 = \dots = (x^{2n-1})^2 + (x^{2n})^2 = 1 \right\},$$

vemos que uma parametrização $\varphi : \mathbb{R}^n \rightarrow \mathbb{T}^n$ para este toro é dada por

$$\varphi(\theta) = \varphi(\theta^1, \dots, \theta^n) = (\cos \theta^1, \sin \theta^1, \dots, \cos \theta^n, \sin \theta^n).$$

Temos, portanto,

$$\frac{\partial \varphi^k}{\partial \theta^j}(\theta) = \begin{cases} -\sin \theta^j & \text{se } k = 2j - 1, \\ \cos \theta^j & \text{se } k = 2j, \\ 0 & \text{se } k \neq 2j - 1, 2j. \end{cases}$$

ou seja,

$$\frac{\partial \varphi}{\partial \theta^j} = \left(0, \dots, 0, -\sin \theta^{j-1}, \cos \theta^j, 0, \dots, 0 \right)$$

Daí

$$g_{ij}(\theta) = \left\langle \frac{\partial \varphi}{\partial \theta^i}(\theta), \frac{\partial \varphi}{\partial \theta^j}(\theta) \right\rangle_{\varphi(\theta)} = \left\langle \frac{\partial \varphi}{\partial \theta^i}(\theta), \frac{\partial \varphi}{\partial \theta^j}(\theta) \right\rangle_{\mathbb{R}^{2n}} = \delta_{ij}, \quad (10.17)$$

que são os mesmos componentes da métrica euclidiana. Portanto, o toro plano é localmente isométrico ao plano \mathbb{R}^n . Observe que considerando \mathbb{T}^2 como uma superfície de \mathbb{R}^3 ou como uma superfície de \mathbb{R}^4 define duas métricas completamente diferentes para a mesma superfície. \square

10.2 Comprimentos e Volumes

10.2.1 Comprimentos de Curvas Parametrizadas

Em variedades métricas podemos definir e calcular o comprimento de **curvas parametrizadas**, isto é, curvas diferenciáveis por partes $\gamma : I \rightarrow M$, onde $I \subset \mathbb{R}$ é um intervalo real; curvas parametrizadas podem possuir autointerseções e até mesmo cúspides ou quinas. Um **segmento** de uma curva parametrizada γ é uma restrição de γ a um intervalo fechado $[a, b] \subset I$. Se M é uma variedade métrica, a **norma** ou comprimento de um vetor $v \in T_p M$ é a norma induzida pelo produto interno:

$$\|v\|_p = \sqrt{|\langle v, v \rangle_p|}. \quad (10.18)$$

10.14 Definição. Seja M uma variedade riemanniana e $\gamma : I \rightarrow M$ uma curva parametrizada. O **comprimento** do segmento de γ definido no intervalo $[a, b] \subset I$ é definido por

$$\ell(\gamma) = \int_a^b \|\gamma'(t)\|_{\gamma(t)} dt. \quad (10.19)$$

\square

10.15 Exemplo. Considere a curva parametrizada $\gamma(t) = (0, t)$ no semiespaço positivo \mathbb{R}_+^2 ; temos $\gamma'(t) = (0, 1) = e_2$. Se \mathbb{R}_+^2 é considerado uma subvariedade riemanniana do plano euclidiano, então

$$\ell(\gamma) = \int_a^b \|e_2\|_{(0,t)} dt = \int_a^b dt = b - a.$$

Se \mathbb{R}_+^2 é o plano hiperbólico \mathbb{H}^2 , então (para $a, b > 0$)

$$\ell(\gamma) = \int_a^b \|e_2\|_{(0,t)} dt = \int_a^b \frac{1}{t} dt = \log b - \log a.$$

Em particular, fixando $b = 1$ temos $\ell(\gamma) \rightarrow +\infty$ quando $a \rightarrow 0$. \square

10.2.2 Volumes em Variedades Riemannianas Orientáveis

A métrica riemanniana também permite definir uma noção de volume em variedades orientadas que permite integrar funções, não apenas formas diferenciais. Seja M uma **variedade riemanniana orientada**. Dado $p \in M$, seja $\mathcal{B}_p = \{e_1, \dots, e_n\}$ uma base ortonormal positiva para $T_p M$. Seja $\varphi : U \rightarrow \varphi(U)$ uma parametrização positiva (isto é, na mesma orientação de M ; para detalhes, veja por exemplo [Carmo], p. 18) de uma vizinhança $\varphi(U)$ de p em M e escreva os vetores da base coordenada de $T_p M$ associada à carta φ em coordenadas em relação à base ortonormal positiva \mathcal{B}_p na seguinte forma:

$$\frac{\partial}{\partial x^i} \Big|_p = A_i^k e_k$$

para $i = 1, \dots, n$. Então

$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x^i} \Big|_p, \frac{\partial}{\partial x^j} \Big|_p \right\rangle = \langle A_i^k e_k, A_j^l e_l \rangle_p = A_i^k A_j^l \langle e_k, e_l \rangle_p = \delta_{kl} A_i^k A_j^l = A_i^k A_j^k.$$

Ou seja, definindo as matrizes $G = (g_{ij})$ e $A = (A_j^i)$, temos

$$G(p) = A^T A$$

onde

$$\det G = (\det A)^2.$$

Denotando por $\text{vol}[v_1, \dots, v_n]$ o volume do paralelepípedo formado pelos vetores v_1, \dots, v_n , sabemos que

$$\text{vol} \left[\frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right] = \det A \text{vol}[e_1, \dots, e_n] = \det A = \sqrt{\det G}(p),$$

já que $\text{vol}[e_1, \dots, e_n] = 1$. Seja $\psi : V \rightarrow \psi(V)$ outra carta positiva de uma vizinhança $\psi(V)$ de p em M e escreva os vetores da base coordenada associada à carta φ em termos dos vetores da base coordenada de $T_p M$ associada à parametrização ψ

$$\frac{\partial}{\partial x^i} \Big|_p = J_i^j \frac{\partial}{\partial y^j} \Big|_p \tag{10.20}$$

com $J_i^j = \partial y^j / \partial x^i$. Denote

$$h_{ij}(p) = \left\langle \frac{\partial}{\partial y^i} \Big|_p, \frac{\partial}{\partial y^j} \Big|_p \right\rangle$$

e

$$H = (h_{ij}).$$

Segue que

$$\begin{aligned} \sqrt{\det G}(p) &= \text{vol} \left[\frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p \right] = \det J \text{vol} \left[\frac{\partial}{\partial y^1} \Big|_p, \dots, \frac{\partial}{\partial y^n} \Big|_p \right] \\ &= \det J \sqrt{\det H}(p). \end{aligned} \tag{10.21}$$

Podemos agora definir o volume.

10.16 Definição. Seja M^n uma variedade riemanniana e $\Omega \subset M$ um conjunto aberto, conexo e com fecho compacto, tal que Ω está contida em uma vizinhança coordenada $\varphi(U)$ de uma parametrização $\varphi : U \rightarrow \varphi(U)$ e a fronteira de $\varphi^{-1}(\Omega)$ tem medida nula em \mathbb{R}^n . O **volume** de Ω é definido por

$$\text{vol } \Omega = \int_{\varphi^{-1}(\Omega)} \sqrt{\det G} dx^1 \dots dx^n. \tag{10.22}$$

Se $\Omega \subset M$ é um compacto, tome qualquer cobertura finita $\{V_i\}_{i=1,\dots,n}$ de Ω por vizinhanças parametrizadas de M e considere uma partição da unidade $\{\rho_i\}_{i=1,\dots,n}$ subordinada a esta cobertura; se $\varphi_i : U_i \rightarrow V_i$, $i = 1, \dots, n$, são parametrizações destas vizinhanças, definimos

$$\text{vol } \Omega = \sum_{i=1}^n \int_{\varphi_i^{-1}(\Omega)} \rho_i \sqrt{\det G} dx^1 \dots dx^n. \quad (10.23)$$

Se $f : M \rightarrow \mathbb{R}$ é uma função contínua com suporte compacto Ω , definimos

$$\int_M f dV_g = \sum_{i=1}^n \int_{\varphi_i^{-1}(\Omega)} f(\varphi_i^{-1}(x)) \sqrt{\det G} dx^1 \dots dx^n. \quad (10.24)$$

□

Segue da fórmula de mudança de variáveis para integrais múltiplas e de (10.21) que o volume está bem definido, isto é, não depende da carta. Na linguagem de formas, o *elemento de volume riemanniano*

$$dV_g = \sqrt{\det G} dx^1 \dots dx^n = \sqrt{\det G} dx^1 \wedge \dots \wedge dx^n \quad (10.25)$$

é uma n -forma (para um tratamento usando formas, veja [Lee 2] e especialmente [Lee 1], Cap. 16, p. 422 em diante, que traz o teorema de Stokes e suas várias versões para variedades riemannianas).

10.3 Grupos de Lie e Álgebras de Lie

O grupo de isometrias de uma variedade Riemanniana é um grupo de Lie, como pode ser visto em [Kobayashi].

10.17 Definição. Um **grupo de Lie** G é uma variedade diferenciável que possui uma estrutura algébrica de grupo tal que a aplicação

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\mapsto gh^{-1} \end{aligned}$$

é diferenciável. □

10.18 Proposição. Seja G um grupo que é uma variedade diferenciável. Então G é um grupo de Lie se e somente se as aplicações

$$\begin{aligned} G &\longrightarrow G \\ g &\mapsto g^{-1} \end{aligned}$$

e

$$\begin{aligned} G \times G &\longrightarrow G \\ (g, h) &\mapsto gh \end{aligned}$$

são diferenciáveis.

Prova: Suponha que G é um grupo de Lie. Então a primeira aplicação é diferenciável porque é a composta das aplicações diferenciáveis

$$\begin{array}{ccccccc} G & \longrightarrow & G \times G & \longrightarrow & G \\ x & \mapsto & (e, g) & \mapsto & eg^{-1} = g^{-1} \end{array}$$

(lembre-se que a inclusão na variedade produto sempre é uma aplicação diferenciável). A segunda é diferenciável porque é a composta das aplicações diferenciáveis

$$\begin{array}{ccccccc} G \times G & \longrightarrow & G \times G & \longrightarrow & G \\ (g, h) & \mapsto & (g, h^{-1}) & \mapsto & g(h^{-1})^{-1} = gh \end{array}$$

(lembre-se que uma aplicação de $G \times G$ em $G \times G$ é uma aplicação diferenciável se e somente se cada aplicação coordenada $G \times G \rightarrow G$ é, e as projeções da variedade produto sobre suas componentes sempre são aplicações diferenciáveis). A recíproca é óbvia. ■

10.19 Exemplo. As seguintes variedades diferenciáveis são grupos de Lie sob as operações indicadas.

- a) \mathbb{R}^n , adição vetorial.
- b) $\mathbb{C} \setminus \{0\}$, multiplicação.
- c) \mathbb{S}^1 , multiplicação induzida de \mathbb{C} .
- d) $G \times H$, variedade produto de dois grupos de Lie G, H , com estrutura de grupo do produto direto dos grupos $(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2)$.
- e) $\mathbb{T}^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$, variedade produto e produto direto n vezes do grupo de Lie \mathbb{S}^1 .
- f) $\mathrm{GL}(\mathbb{R}^n)$ (o grupo *linear geral* das matrizes reais invertíveis $n \times n$, subvariedade de \mathbb{R}^{n^2}), multiplicação matricial.

□

10.20 Proposição. Seja G um grupo de Lie. Dado $g \in G$, então as aplicações translação à esquerda por g

$$\begin{array}{ccc} L_g : & G & \longrightarrow & G \\ & g & \mapsto & gh \end{array} \tag{10.26}$$

e translação à direita por g

$$\begin{array}{ccc} R_g : & G & \longrightarrow & G \\ & h & \mapsto & yg \end{array} \tag{10.27}$$

são difeomorfismos.

Observando que

$$\begin{aligned} (L_g)^{-1} &= L_{g^{-1}}, \\ (R_g)^{-1} &= R_{g^{-1}}, \end{aligned}$$

temos pela regra da cadeia

$$[(dL_g)_h]^{-1} = (dL_{g^{-1}})_{L_g h} = (dL_{g^{-1}})_{gh} \tag{10.28}$$

e similarmente

$$[(dR_g)_h]^{-1} = (dR_{g^{-1}})_{gh}. \tag{10.29}$$

10.21 Definição. Uma **álgebra de Lie** (sobre \mathbb{R}) é um espaço vetorial \mathcal{G} munido de uma aplicação bilinear, chamada o **colchete de Lie**,

$$[\cdot, \cdot] : V \times V \longrightarrow \mathbb{R}$$

que satisfaz

- (i) (anticomutatividade)

$$[X, Y] = -[Y, X]; \tag{10.30}$$

- (ii) (identidade de Jacobi)

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \tag{10.31}$$

para todos $X, Y, Z \in \mathcal{G}$. □

10.22 Exemplo. Seja M uma variedade diferenciável. O espaço vetorial $\mathcal{T}(M)$, equipado com o colchete de Lie é uma álgebra de Lie. \square

10.23 Exemplo. O espaço vetorial das matrizes reais $n \times n$ com a operação colchete definida por

$$[A, B] = AB - BA \quad (10.32)$$

é uma álgebra de Lie. De fato, bilinearidade e anticomutatividade claramente valem e

$$\begin{aligned} & [[A, B], C] + [[B, C], A] + [[C, A], B] \\ &= (AB - BA)C - C(AB - BA) + (BC - CB)A - A(BC - CB) \\ &+ (CA - AC)B - B(CA - AC) \\ &= ABC - BAC - CAB + CBA + BCA - CBA - ABC + ACB \\ &+ CAB - ACB - BCA + BAC \\ &= 0. \end{aligned}$$

Esta álgebra de Lie é denotada por $\mathfrak{gl}(\mathbb{R}^n)$. \square

10.24 Exemplo. \mathbb{R}^3 com o produto vetorial é uma álgebra de Lie. \square

Veremos agora a relação entre grupos de Lie e álgebras de Lie. Primeiro algumas preliminares. Pela Proposição 10.20, toda translação à esquerda L_x em um grupo de Lie é um difeomorfismo.

10.25 Definição. Seja G um grupo de Lie. Dizemos que um campo vetorial $X \in \mathcal{T}(G)$ é **invariante à esquerda** se

$$dL_g X = X \circ L_g \quad (10.33)$$

para todo $g \in G$. \square

Explicitando a definição acima, temos que para todo $g \in G$ vale

$$(dL_g)_h X_h = X_{L_g h} = X_{gh}. \quad (10.34)$$

para todo $h \in G$. Em particular, um campo invariante à esquerda fica completamente determinado pelo seu valor em algum ponto qualquer de G . Por exemplo, se conhecemos o valor de X_e então, tomando $h = e$, segue que

$$X_g = (dL_g)_e X_e. \quad (10.35)$$

O próximo resultado garante a existência de um número infinito de campos invariantes à esquerda em um grupo de Lie.

10.26 Proposição. *Seja G um grupo de Lie. Todo vetor tangente X_e no espaço tangente $T_e G$ possui uma extensão a um campo invariante à esquerda $X \in T(G)$.*

Prova: Basta definir

$$X_g = (dL_g)_e X_e.$$

Como L_g é um difeomorfismo C^∞ , claramente $X \in \mathcal{T}(G)$. Para ver que X é um campo invariante à esquerda, seja $h \in G$ qualquer. Como

$$L_h \circ L_g = L_{hg}$$

segue que

$$(dL_h)_g X_g = (dL_h)_g (dL_g)_e X_e = [d(L_h \circ L_g)]_e X_g = (dL_{hg})_e X_g = X_{hg}.$$

■

10.27 Proposição. Seja G um grupo de Lie. Então o colchete de Lie de campos invariantes à esquerda é invariante à esquerda.

Em particular, o subespaço dos campos invariantes à esquerda é uma álgebra de Lie.

Prova: Sejam $X, Y \in \mathcal{T}(G)$ campos invariantes à esquerda. Temos, para toda $f \in C^\infty(G)$,

$$\begin{aligned} (dL_g)_h [X, Y]_h f &= [X, Y]_h (f \circ L_g) \\ &= XY_h (f \circ L_g) - YX_h (f \circ L_g) \\ &= X (dL_g)_h Y_h (f) - Y (dL_g)_h X_h (f) \\ &= XY_{gh} (f) - YX_{gh} (f) \\ &= [X, Y]_{gh} f. \end{aligned}$$

Como o subconjunto dos campos invariantes à esquerda é um subespaço vetorial de $\mathcal{T}(G)$, segue que ele é uma (sub)álgebra de Lie. ■

Podemos agora definir uma operação colchete de Lie no espaço tangente (espaço vetorial) $T_e G$ que o transforma em uma álgebra de Lie:

10.28 Definição. Seja G um grupo de Lie. A **álgebra de Lie** \mathfrak{g} de G é o espaço tangente $T_e G$ munido do colchete de Lie

$$[X_e, Y_e] := [X, Y]_e. \quad (10.36)$$

onde $X, Y \in \mathcal{T}(G)$ são quaisquer extensões invariantes à esquerda dos vetores tangentes X_e, Y_e . □

Pode-se mostrar que a álgebra de Lie dos campos invariantes à esquerda é isomorfa ao espaço tangente $T_e G$, em particular como espaço vetorial ela tem dimensão finita, igual à dimensão de G (veja [Lee 1], Teorema 8.37, e [Warner], p.86, Definição 3.1).

Podemos agora introduzir uma métrica em G com certas propriedades algébricas.

10.29 Definição. Seja G um grupo de Lie. Dizemos que uma métrica $\langle \cdot, \cdot \rangle_g$ em G é **invariante à esquerda** se

$$\langle v, w \rangle_h = \langle (dL_g)_h v, (dL_g)_h w \rangle_{L_g h} \quad (10.37)$$

para todos $g, h \in G$ e para todos $v, w \in T_e G$. Analogamente, definimos uma métrica **invariante à direita**.

Uma métrica que é ao mesmo tempo invariante à esquerda e à direita é chamada uma **métrica bi-invariante**. □

Em outras palavras, em uma métrica invariante à esquerda, toda translação à esquerda L_g é uma isometria, enquanto que em uma métrica invariante à direita, toda translação à direita R_g é uma isometria. Em uma métrica bi-invariante todas as translações são isometrias. A existência de métricas bi-invariantes para grupos de Lie compactos é estabelecida no Exercício 7 de [Carmo]. A existência de métrica invariantes à esquerda ou à direita em qualquer grupo de Lie é estabelecida através da seguinte definição.

10.30 Proposição. Seja G um grupo de Lie. Suponha que $\langle \cdot, \cdot \rangle$ é algum produto interno em $T_e G$. Então a métrica em G definida por

$$\langle v, w \rangle_g = \left\langle (dL_{g^{-1}})_g v, (dL_{g^{-1}})_g w \right\rangle_e \quad (10.38)$$

para todo $g \in G$ e para todos $v, w \in T_e G$, é invariante à esquerda.

Analogamente, a métrica em G definida por

$$\langle v, w \rangle_g = \left\langle (dR_{g^{-1}})_g v, (dR_{g^{-1}})_g w \right\rangle_e$$

para todo $g \in G$ e para todos $v, w \in T_e G$, é invariante à direita.

Prova: Temos, por definição,

$$\begin{aligned}
 \langle (dL_g)_h v, (dL_g)_h w \rangle_{L_g h} &= \left\langle \left(dL_{(L_g h)^{-1}} \right)_{L_g h} (dL_g)_h v, \left(dL_{(L_g h)^{-1}} \right)_{L_g h} (dL_g)_h w \right\rangle_e \\
 &= \left\langle (dL_{h^{-1}g^{-1}})_{gh} (dL_g)_h v, (dL_{h^{-1}g^{-1}})_{gh} (dL_g)_h w \right\rangle_e \\
 &= \langle (dL_{h^{-1}})_h v, (dL_{h^{-1}})_h w \rangle_e \\
 &= \langle v, w \rangle_h
 \end{aligned}$$

lembrando que $L_{h^{-1}g^{-1}} \circ L_g = L_{h^{-1}g^{-1}g} = L_{h^{-1}}$. Analogamente prova-se a invariância à direita da segunda métrica. ■

Há uma relação entre o produto interno e o colchete de Lie em $\mathcal{G} = T_e G$ que caracteriza as métricas bi-invariantes de G que enunciaremos sem prova.

10.31 Teorema. *Seja G um grupo de Lie com álgebra de Lie \mathcal{G} . A métrica invariante à esquerda definida na proposição anterior é bi-invariante se e somente se o produto escalar \langle , \rangle em $\mathcal{G} = T_e G$ usado para definir a métrica satisfaz*

$$\langle [V, X], W \rangle = -\langle V, [W, X] \rangle$$

para todos $V, W, X \in \mathcal{G}$.

10.4 Exercícios

10.32 Exercício. Mostre que a métrica produto é de fato uma métrica. Porque

$$g_{(p_1, p_2)}((v_1, w_1), (v_2, w_2)) = g_{p_1}(v_1, v_2) g_{p_2}(w_1, w_2)$$

não define uma métrica na variedade produto $M_1 \times M_2$?

Capítulo 11

Conexões

É possível definir geodésicas e estudar suas propriedades sem falar de curvatura. Na verdade é até possível falar em geodésicas sem falar de métrica. Geodésicas são generalizações das retas da geometria euclidiana. Embora seja possível definir geodésicas como curvas que minimizam distâncias em variedades riemannianas e curvas que maximizam o tempo próprio em variedades de Lorentz, pelo menos localmente (e neste caso a noção de geodésica estaria também fundamentalmente ligada à noção de métrica) esta propriedade é tecnicamente difícil de trabalhar como definição. Ao invés, escolheremos uma propriedade diferente das retas para generalizar: retas são caracterizadas como curvas com aceleração nula. Esta propriedade não faz nenhuma referência à métrica e pode ser utilizada mesmo em variedades diferenciáveis que não tenham uma estrutura métrica. E em relatividade geral ela faz mais sentido, já que partículas livres seguem trajetórias geodésicas no espaço-tempo curvo, já que não estão sujeitas a forças externas e portanto não sofrem aceleração, assim como partículas livres na mecânica newtoniana seguem trajetórias retas com velocidade uniforme. Para usar esta propriedade, precisaremos primeiro definir o conceito de derivadas de campos tangentes à curvas segundo uma direção. Como em geral não existe um espaço ambiente \mathbb{R}^n onde a variedade está mergulhada, não é imediatamente óbvio como defini-lo. Se $\alpha : I \rightarrow M$ é uma curva diferenciável em uma variedade M não podemos simplesmente definir

$$\alpha''(t_0) = \lim_{t \rightarrow t_0} \frac{\alpha'(t) - \alpha'(t_0)}{t - t_0}$$

porque $\alpha'(t) \in T_{\alpha(t)}M$ e $\alpha'(t_0) \in T_{\alpha(t_0)}M$ vivem em diferentes espaços vetoriais, logo sua diferença não faz sentido. Já vimos que a derivada de Lie não é uma boa solução para este problema, já que ela não é uma derivada direcional. A definição do conceito de *conexão* atende esta necessidade de definir uma noção de derivação intrínseca para campos vetoriais e se comporta exatamente como uma derivada direcional. O nome *conexão* se refere exatamente à idéia de conectar localmente os espaços tangentes de uma variedade.

11.1 Conexões e Derivada Covariante

Consideremos o conjunto $\mathcal{T}(M)$ dos campos vetoriais diferenciáveis em uma variedade diferenciável M como um módulo sobre o anel $C^\infty(M)$ das funções suaves definidas em M .

11.1 Definição. Seja M uma variedade diferenciável de dimensão n . Uma **conexão** ∇ em M é uma aplicação

$$\nabla : \mathcal{T}(M) \times \mathcal{T}(M) \longrightarrow \mathcal{T}(M),$$

denotada por $(X, Y) \mapsto \nabla_X Y$ que satisfaz as seguintes propriedades:

- (i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$,
- (ii) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$,
- (iii) $\nabla_X (fY) = f\nabla_X Y + (Xf)Y$.

para todos os campos $X, Y, Z \in \mathcal{T}(M)$ e para todas as funções $f, g \in C^\infty(M)$.

Dizemos que $\nabla_X Y$ é a **derivada covariante** do campo Y na direção de X . \square

Uma explicação para o nome derivada *covariante* será dada no Capítulo 14. O símbolo $\nabla_X Y$ deve ser interpretado como a derivada direcional do campo Y na direção X . O resultado a seguir reforça esta interpretação:

11.2 Proposição (Conexão em Coordenadas). *Seja ∇ uma conexão em uma variedade diferenciável M . Se $X, Y \in \mathcal{T}(M)$ são campos vetoriais que se expressam em coordenadas locais por*

$$X = X^i \partial_i \quad e \quad Y = Y^j \partial_j,$$

então

$$\nabla_X Y = X(Y^j) \partial_j + X^i Y^j \nabla_{\partial_i} \partial_j. \quad (11.1)$$

Em particular, $(\nabla_X Y)_p$ depende apenas do valor de X em p e do valor de Y ao longo de uma curva tangente a X_p .

Prova: Usando as propriedades de uma conexão, obtemos

$$\begin{aligned} \nabla_X Y &= \nabla_X (Y^j \partial_j) \\ &= Y^j \nabla_X \partial_j + X(Y^j) \partial_j \\ &= Y^j \nabla_{X^i \partial_i} \partial_j + X(Y^j) \partial_j \\ &= X^i Y^j \nabla_{\partial_i} \partial_j + X(Y^j) \partial_j. \end{aligned}$$

Em particular,

$$(\nabla_X Y)_p = X^i(p) Y^j(p) (\nabla_{\partial_i} \partial_j)_p + [X_p(Y^j)](p) \partial_j|_p.$$

Os coeficientes $X^1(p), \dots, X^n(p)$ dependem apenas do valor de X em p ; os coeficientes $X_p(Y^1), \dots, X_p(Y^n)$, por definição de vetor tangente, dependem apenas dos valores de Y ao longo de uma curva passando por p cujo vetor tangente em p é X_p . ■

Da expressão (11.1), escrevendo os campos vetoriais $\nabla_{\partial_i} \partial_j$ em termos dos campos base ∂_k na forma

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k, \quad (11.2)$$

obtemos a seguinte expressão local para o campo $\nabla_X Y$:

$$\nabla_X Y = (X(Y^k) + X^i Y^j \Gamma_{ij}^k) \partial_k. \quad (11.3)$$

11.3 Definição. As funções suaves Γ_{ij}^k definidas pela expressão (11.3) são chamadas os **símbolos de Christoffel** associados à carta particular utilizada. \square

Observe que em princípio precisamos obter n^3 símbolos de Christoffel para determinar uma conexão. No caso de conexões riemannianas, como veremos, a sua simetria diminuirá o número de símbolos diferentes.

11.4 Proposição. *Toda variedade diferenciável possui uma conexão.*

Prova: Se V é uma vizinhança coordenada de M , dadas n^3 funções arbitrárias $\Gamma_{ij}^k \in C^\infty(V)$, a fórmula (11.3) define uma conexão em V , vista como subvariedade de M . Se $\{V_\alpha\}$ é uma cobertura de M por vizinhanças coordenadas, cada uma com uma conexão ∇^α definida, então podemos definir uma conexão global em M , usando uma partição da unidade $\{\rho_\alpha\}$ subordinada a esta cobertura, por

$$\nabla_X Y = \sum_\alpha \rho_\alpha \nabla_X^\alpha Y.$$

As propriedades de uma conexão são facilmente verificadas; apenas a regra do produto merece atenção especial, já que combinações lineares de conexões não são conexões em geral, exatamente por deixarem de satisfazer a regra do produto. Mas combinações lineares convexas de conexões *são* conexões e no nosso caso temos

$$\begin{aligned}\nabla_X(fY) &= \sum_{\alpha} \rho_{\alpha} \nabla_X^{\alpha}(fY) = \sum_{\alpha} \rho_{\alpha} [(Xf)Y + f\nabla_X^{\alpha}Y] \\ &= (Xf)Y \sum_{\alpha} \rho_{\alpha} + f \sum_{\alpha} \rho_{\alpha} \nabla_X^{\alpha}Y \\ &= (Xf)Y + f\nabla_XY.\end{aligned}$$

■

11.5 Exemplo (Conexão Euclidiana). Identificando espaços tangentes em \mathbb{R}^n com o próprio \mathbb{R}^n , vetores tangentes com vetores em \mathbb{R}^n e campos vetoriais em \mathbb{R}^n com aplicações suaves $\mathbb{R}^n \rightarrow \mathbb{R}^n$, nós definimos a **conexão euclidiana** $\nabla : \mathcal{T}(\mathbb{R}^n) \times \mathcal{T}(\mathbb{R}^n) \rightarrow \mathcal{T}(\mathbb{R}^n)$ por

$$(\nabla_X Y)_p = dY_p(X_p), \quad (11.4)$$

ou seja, a derivada direcional do campo Y em p na direção de X_p . Em coordenadas, usando a definição de diferencial em \mathbb{R}^n ,

$$dY_p(X_p) = X^i \frac{\partial Y^j}{\partial x^i} e_j,$$

ou seja,

$$\nabla_X Y = X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j}. \quad (11.5)$$

Outra maneira de obter a mesma expressão em coordenadas, usando a regra da cadeia,

$$dY_p(X_p)(f) = X_p(f \circ Y) = X^i \frac{\partial(f \circ Y)}{\partial x^i} = X^i \frac{\partial f}{\partial x^j} \frac{\partial Y^j}{\partial x^i} = X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j}(f).$$

Em notação mais sucinta, a expressão em coordenadas da conexão euclidiana que obtemos a partir de (11.5) é

$$\nabla_X Y = X(Y^j) \frac{\partial}{\partial x^j}. \quad (11.6)$$

Segue de (11.3) e da observação no início da demonstração da Proposição 11.4 que a conexão euclidiana é de fato uma conexão com símbolos de Christoffel $\Gamma_{ij}^k = 0$.

Esta é uma conexão tanto para o espaço euclidiano, para o espaço semieuclidiano e para o espaço de Minkowski, já que apenas a estrutura diferencial está envolvida. □

11.2 Derivada Covariante ao longo de Curvas

A existência de uma conexão em uma variedade diferenciável M permite derivar campos vetoriais ao longo de curvas na variedade. Em particular, é possível falar em aceleração de uma curva e portanto de geodésicas e, eventualmente, curvatura. Na próxima seção veremos que uma métrica define uma conexão única em uma variedade métrica. Conexões diferentes da conexão induzida pela métrica permitem a definição de estruturas geométricas em variedades diferenciáveis mais gerais que a dada pela métrica; em particular, é possível falar de geodésicas sem uma noção de métrica.

Veremos agora como a conexão permite definir uma noção intrínseca de derivada de um campo vetorial ao longo de uma curva na variedade.

11.6 Definição. Seja $\alpha : I \rightarrow M$ uma curva diferenciável em uma variedade diferenciável M . Um **campo vetorial ao longo da curva** α é um campo vetorial diferenciável $V : I \rightarrow TM$ tal que $V(t) \in T_{\alpha(t)}M$ para todo $t \in I$.

O espaço vetorial dos campos vetoriais ao longo de uma curva α é denotado $\mathcal{T}(\alpha)$. \square

11.7 Proposição. Seja M uma variedade diferenciável com uma conexão ∇ . Existe uma única correspondência que associa a cada campo vetorial diferenciável V ao longo de uma curva diferenciável $\alpha : I \rightarrow M$ um outro campo diferenciável

$$\frac{DV}{dt}$$

ao longo de α tal que

$$\begin{aligned}\frac{D}{dt}(V + W) &= \frac{DV}{dt} + \frac{DW}{dt}, \\ \frac{D}{dt}(fV) &= \frac{df}{dt}V + f\frac{DV}{dt}\end{aligned}$$

para todos os campos diferenciáveis V, W ao longo de α e para toda função diferenciável $f : I \rightarrow \mathbb{R}$, e tal que, se V é induzido por um campo de vetores $X \in \mathcal{T}(M)$, ou seja, $V = X \circ \alpha$, então

$$\frac{DV}{dt} = \nabla_{\alpha'(t)}X.$$

Localmente,

$$\frac{DV}{dt} = \left(\frac{dV^k}{dt} + \frac{d\alpha^i}{dt} \Gamma_{ij}^k V^j \right) \partial_k, \quad (11.7)$$

Prova: Observe que para a expressão $\nabla_{\alpha'(t)}X$ fazer sentido, devemos entender o subscrito $\alpha'(t)$ neste símbolo como qualquer extensão local do campo $\alpha'(t)$ a um campo em M , já que pela Proposição 11.2 só importa o valor da extensão em $\alpha(t)$, isto é, o vetor tangente $\alpha'(t)$, e o valor de X em uma curva tangente a $\alpha'(t)$ em $\alpha(t)$, que pode ser tomada como sendo a própria curva α .

Vamos provar primeiro a unicidade de $\frac{DV}{dt}$. Suponha que exista um tal campo $\frac{DV}{dt}$ satisfazendo todas as propriedades do enunciado. Seja

$$V(t) = V^j(t) \partial_j|_t$$

a expressão local do campo V . Pelas primeiras duas propriedades do enunciado, temos

$$\frac{DV}{dt}|_t = \frac{dV^j}{dt}(t) \partial_j|_t + V^j(t) \frac{D\partial_j}{dt}|_t.$$

Pela terceira propriedade,

$$\frac{D\partial_j}{dt}|_t = (\nabla_{\alpha'(t)}\partial_j)_t = \left(\nabla_{\frac{d\alpha^i}{dt}(t)\partial_i} \partial_j \right)_t = \frac{d\alpha^i}{dt}(t) \nabla_{\partial_i} \partial_j|_t.$$

Portanto, localmente o campo $\frac{DV}{dt}$ se escreve na forma

$$\frac{DV}{dt}|_t = \left[\frac{dV^k}{dt}(t) + \frac{d\alpha^i}{dt}(t) \Gamma_{ij}^k(t) V^j(t) \right] \partial_k|_t,$$

o que mostra que o campo $\frac{DV}{dt}$ é unicamente determinado.

Para determinar a existência de $\frac{DV}{dt}$, dada uma carta (φ, U) para uma vizinhança de $\alpha(t)$, defina o campo $\frac{DV}{dt}$ em $\varphi(U)$ pela expressão (11.7); é imediato verificar que um campo definido desta forma satisfaz todas as propriedades do enunciado. ■

11.8 Definição. O campo diferenciável $\frac{DV}{dt}$ é chamado a **derivada covariante** de V ao longo da curva α . \square

11.3 Transporte Paralelo

11.9 Definição. Seja M uma variedade diferenciável com uma conexão ∇ . Um campo vetorial diferenciável V ao longo de uma curva diferenciável $\alpha : I \rightarrow M$ é chamado um **campo paralelo ao longo** de α se

$$\frac{DV}{dt} \equiv 0.$$

Um campo vetorial $X \in \mathcal{T}(M)$ é chamado um **campo paralelo** se ele é paralelo ao longo de qualquer curva. \square

É fácil ver que um campo vetorial $X \in \mathcal{T}(M)$ é paralelo se e somente se

$$\nabla_Y X = 0$$

para todo campo $Y \in \mathcal{T}(M)$.

11.10 Proposição. Seja M uma variedade diferenciável com uma conexão ∇ . Seja $\alpha : I \rightarrow M$ uma curva diferenciável e V_0 um vetor tangente em $\alpha(t_0)$, $t_0 \in I$. Então existe um único campo paralelo V definido ao longo de α tal que $V_{t_0} = V_0$.

Prova: Usando a expressão (11.7) obtida na Proposição 11.7, o campo derivada covariante $\frac{DV}{dt}$ em coordenadas locais se escreve na forma

$$\frac{DV}{dt} \Big|_t = \left[\frac{dV^k}{dt}(t) + \frac{d\alpha^i}{dt}(t) \Gamma_{ij}^k(t) V^j(t) \right] \partial_k|_t, \quad (11.8)$$

Logo, a existência local do campo $V(t)$ satisfazendo $\frac{DV}{dt} = 0$ para todo t e $V(t_0) = V_0$ corresponde a uma solução do sistema linear de n equações diferenciais

$$\begin{cases} \frac{dV^1}{dt}(t) + \frac{d\alpha^i}{dt}(t) \Gamma_{ij}^1(t) V^j(t) = 0 \\ \vdots \\ \frac{dV^n}{dt}(t) + \frac{d\alpha^i}{dt}(t) \Gamma_{ij}^n(t) V^j(t) = 0 \end{cases}$$

com condição inicial

$$V^1(t_0) = V_0^1, \dots, V^n(t_0) = V_0^n.$$

Se $\alpha(I)$ está inteiramente contida em uma vizinhança coordenada, então o teorema de existência e unicidade para equações diferenciais lineares garante a existência de um único campo V definido em todo o intervalo I . Caso contrário, como $\alpha(I)$ é um conjunto compacto, ela pode ser coberta por um número finito de vizinhanças coordenadas, em cada uma das quais V pode ser definido de maneira única usando o raciocínio acima e esta unicidade garante que o campo é o mesmo nas interseções das vizinhanças. ■

11.11 Definição. O campo V obtido na Proposição 11.10 é chamado o **transporte paralelo** de V_0 ao longo de α .

A aplicação transporte paralelo é a aplicação linear

$$\tau_t : T_{\alpha(t_0)} M \longrightarrow T_{\alpha(t)} M$$

definida em cada vetor $V_0 \in T_{\alpha(t_0)} M$ por

$$\tau_t(V_0) = V_t,$$

isto é, $\tau_t(V_0)$ é o transporte paralelo do vetor V_0 ao longo da curva α . \square

Quando necessário, para $s, t \in I$ denotaremos a aplicação transporte paralelo de vetores em $T_{\alpha(s)} M$ para vetores em $T_{\alpha(t)} M$ por

$$\tau_{s \rightarrow t} : T_{\alpha(s)} M \longrightarrow T_{\alpha(t)} M.$$

A aplicação transporte paralelo é linear porque o transporte paralelo é dado pela solução de um sistema de equações diferenciais lineares. Por unicidade, ela é um isomorfismo com

$$\tau_{s \rightarrow t}^{-1} = \tau_{t \rightarrow s},$$

e da unicidade de solução para um sistema de EDOs segue também que

$$\begin{aligned} \tau_{0 \rightarrow 0} &= \text{id}, \\ \tau_{r \rightarrow t} \circ \tau_{s \rightarrow r} &= \tau_{s \rightarrow t}. \end{aligned}$$

Em geral, o transporte paralelo de um vetor V em $T_p M$ para um vetor em $T_q M$ dependerá da curva α ligando p e q usada; isto é, se $\alpha_1, \alpha_2 : I \longrightarrow M$ são duas curvas diferenciáveis tais que

$$\begin{aligned} \alpha_1(s) &= \alpha_2(s) = p, \\ \alpha_1(t) &= \alpha_2(t) = q, \end{aligned}$$

então em geral

$$\tau_{s \rightarrow t}^{\alpha_1}(V) \neq \tau_{s \rightarrow t}^{\alpha_2}(V)$$

para todo $V \in T_p M$. Como veremos depois, o transporte paralelo será o mesmo, independente do caminho utilizado para ir de p até q , se e somente se a curvatura riemanniana for nula.

11.4 Conexão Compatível com a Métrica

Como métricas e conexões definem cada uma uma estrutura geométrica particular, o caso mais relevante de variedade métrica dotada de uma conexão é quando as estruturas geométricas definidas por elas coincidem. Para isso a conexão deve satisfazer duas condições. A primeira delas é a chamada *compatibilidade da conexão com a métrica*, que pode ser definida de qualquer um dos três modos equivalentes a seguir.

11.12 Proposição. *Seja M uma variedade métrica com uma conexão ∇ . Então as seguintes afirmações são equivalentes:*

(i) *Para todos os campos paralelos V e W ao longo de qualquer curva diferenciável α em M vale*

$$\langle V, W \rangle \equiv \text{constante.} \quad (11.9)$$

(ii) *Para todos os campos vetoriais V e W ao longo de qualquer curva diferenciável α em M vale*

$$\frac{d}{dt} \langle V, W \rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle. \quad (11.10)$$

(iii) *Para todos os campos $X, Y, Z \in \mathfrak{T}(M)$ vale*

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle. \quad (11.11)$$

Prova: (ii) \Rightarrow (i) Se V e W são campos paralelos, então

$$\frac{d}{dt} \langle V, W \rangle = \langle 0, W \rangle + \langle V, 0 \rangle = 0$$

e portanto $\langle V, W \rangle$ é constante.

(i) \Rightarrow (ii) Seja $\alpha : I \rightarrow M$ uma curva diferenciável qualquer e para um ponto fixado $t_0 \in I$ escolha uma base ortonormal

$$\mathcal{B}_0 = \{E_1|_{t_0}, \dots, E_n|_{t_0}\}$$

para $T_{\alpha(t_0)}M$. Usando a Proposição 11.10, estenda cada um dos vetores $E_1|_{t_0}, \dots, E_n|_{t_0}$ paralelamente a campos E_1, \dots, E_n ao longo de α . Segue da definição de compatibilidade que

$$\mathcal{B}_t = \{E_1|_t, \dots, E_n|_t\}$$

é uma base ortonormal para $T_{\alpha(t)}M$ para cada $t \in I$. Dados campos diferenciáveis V e W ao longo de α , podemos então escrever

$$V = V^i(t) E_i|_t \quad \text{e} \quad W = W^j(t) E_j|_t$$

com as funções V^i, W^j diferenciáveis, pois

$$V^i(t) = \langle V, E_i \rangle \quad \text{e} \quad W^j(t) = \langle W, E_j \rangle.$$

Como os campos $E_1(t), \dots, E_n(t)$ são paralelos ao longo de α , temos

$$\frac{dE_1}{dt} = \dots = \frac{dE_n}{dt} = 0,$$

logo

$$\frac{dV}{dt} \Big|_t = \frac{dV^i}{dt}(t) E_i|_t + V^i(t) \frac{dE_i}{dt} \Big|_t = \frac{dV^i}{dt}(t) E_i|_t,$$

e, similarmente,

$$\frac{dW}{dt} \Big|_t = \frac{dW^j}{dt}(t) E_j|_t.$$

Daí,

$$\begin{aligned} \left\langle \frac{dV}{dt}, W \right\rangle + \left\langle V, \frac{dW}{dt} \right\rangle &= \left\langle \frac{dV^i}{dt} E_i, W^j E_j \right\rangle + \left\langle V^i E_i, \frac{dW^j}{dt} E_j \right\rangle \\ &= \frac{dV^i}{dt} W^j \langle E_i, E_j \rangle + V^i \frac{dW^j}{dt} \langle E_i, E_j \rangle \\ &= \left(\frac{dV^i}{dt} W^j + V^i \frac{dW^j}{dt} \right) \delta_{ij} \\ &= \sum_{i=1}^n \left(\frac{dV^i}{dt} W^i + V^i \frac{dW^i}{dt} \right) = \frac{d}{dt} \sum_{i=1}^n V^i W^i \\ &= \frac{d}{dt} \langle V, W \rangle. \end{aligned}$$

(ii) \Rightarrow (iii) Seja $p \in M$ e $\alpha : I \rightarrow M$ uma curva diferenciável com $\alpha(t_0) = p$ e $\alpha'(t_0) = X_p$. Então, por definição de vetor tangente,

$$\begin{aligned} X_p \langle Y, Z \rangle &= \frac{d}{dt} \langle Y_{\alpha(t)}, Z_{\alpha(t)} \rangle \Big|_{t=t_0} = \left\langle \frac{DY}{dt} \Big|_{t_0}, Z_{t_0} \right\rangle + \left\langle Y_{t_0}, \frac{DZ}{dt} \Big|_{t_0} \right\rangle \\ &= \left\langle (\nabla_{\alpha'(t_0)} Y)_{\alpha(t_0)}, Z_{t_0} \right\rangle + \left\langle Y_{t_0}, (\nabla_{\alpha'(t_0)} Z)_{\alpha(t_0)} \right\rangle \\ &= \langle (\nabla_X Y)_p, Z_p \rangle + \langle Y_p, (\nabla_X Z)_p \rangle. \end{aligned}$$

(iii) \Rightarrow (ii) Se V, W são campos ao longo de uma curva diferenciável α em M com $\alpha(t_0) = p$ e $\alpha'(t_0) = X_p$, então

$$\begin{aligned} \frac{d}{dt} \langle V_t, W_t \rangle \Big|_{t=t_0} &= X_p \langle V_t, W_t \rangle \\ &= \langle (\nabla_X V)_p, W_t \rangle + \langle V_t, (\nabla_X W)_p \rangle \\ &= \left\langle \left(\nabla_{\alpha'(t_0)} V \right)_{\alpha(t_0)}, W_{t_0} \right\rangle + \left\langle V_{t_0}, \left(\nabla_{\alpha'(t_0)} W \right)_{\alpha(t_0)} \right\rangle \\ &= \left\langle \frac{DV}{dt} \Big|_{t_0}, W_{t_0} \right\rangle + \left\langle V_{t_0}, \frac{DW}{dt} \Big|_{t_0} \right\rangle. \end{aligned}$$

■

A condição $\langle V, W \rangle \equiv \text{constante}$ justifica o nome campos *paralelos*.

11.13 Definição. Seja M uma variedade métrica com uma conexão ∇ . Dizemos que a conexão é **compatível com a métrica**, quando ela satisfaz qualquer uma das condições da proposição anterior. \square

11.5 O Tensor Torsão

11.5.1 Conexão Simétrica

A segunda condição para que a estrutura geométrica definida pela conexão seja a mesma definida pela métrica é que a conexão seja *simétrica*:

11.14 Definição. Seja M uma variedade diferenciável com uma conexão ∇ . O tensor **torção** da conexão ∇ é a aplicação

$$T : \mathcal{T}(M) \times \mathcal{T}(M) \longrightarrow \mathcal{T}(M)$$

definida por

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Dizemos que a conexão ∇ é **simétrica** se $T \equiv 0$, isto é, se para todos os campos $X, Y \in \mathcal{T}(M)$ vale

$$\nabla_X Y - \nabla_Y X = [X, Y]. \quad (11.12)$$

□

Apesar da conexão não ser um tensor, sua torção é de fato um $(2, 1)$ -tensor, pois ela só depende do valor no ponto. Em coordenadas, como

$$\begin{aligned} \nabla_X Y &= (X(Y^k) + X^i Y^j \Gamma_{ij}^k) \partial_k, \\ \nabla_Y X &= (Y(X^k) + Y^i X^j \Gamma_{ij}^k) \partial_k, \\ [X, Y] &= (X(Y^k) - Y(X^k)) \partial_k, \end{aligned}$$

o tensor torção é dado por

$$T(X, Y) = (X^i Y^j - Y^i X^j) \Gamma_{ij}^k \partial_k, \quad (11.13)$$

de onde vemos que $T(X, Y)$ é linear em relação a X e a Y , separadamente, e depende apenas de X_p e Y_p .

11.15 Proposição. Seja M uma variedade com uma conexão simétrica ∇ . Então

$$\nabla_{\partial_i} \partial_j = \nabla_{\partial_j} \partial_i \quad (11.14)$$

e

$$\Gamma_{ij}^k = \Gamma_{ji}^k. \quad (11.15)$$

Prova: Como

$$[\partial_i, \partial_j] = \nabla_{\partial_i} \partial_j - \nabla_{\partial_j} \partial_i = (\Gamma_{ij}^k - \Gamma_{ji}^k) \partial_k$$

e $[\partial_i, \partial_j] = 0$, seguem os resultados. ■

Em particular, para conexões simétricas o número de símbolos de Christoffel potencialmente diferentes cai para

$$\frac{n^2(n+1)}{2}.$$

Seja M uma variedade diferenciável com duas conexões ∇ e $\bar{\nabla}$. Defina o *tensor diferença*

$$D(X, Y) = \bar{\nabla}_Y X - \nabla_Y X.$$

Definindo os tensores S (simétrico) e A (alternado)

$$\begin{aligned} S(X, Y) &= \frac{1}{2} [D(X, Y) + D(Y, X)], \\ A(X, Y) &= \frac{1}{2} [D(X, Y) - D(Y, X)], \end{aligned}$$

segue que $D = S + A$.

- a) Mostre que as conexões ∇ e $\bar{\nabla}$ tem a mesma torsão se e somente se $A = 0$.
- b) Mostre que as conexões ∇ e $\bar{\nabla}$ tem as mesmas geodésicas (com as mesmas parametrizações) se e somente se $D(X, X) = 0$ para todo $X \in \mathfrak{X}(M)$.
- c) Mostre que as conexões ∇ e $\bar{\nabla}$ tem as mesmas geodésicas (com as mesmas parametrizações) se e somente se $S = 0$.
- d) Conclua que se as conexões ∇ e $\bar{\nabla}$ tem as mesmas geodésicas (com as mesmas parametrizações) e a mesma torção, então $\bar{\nabla} = \nabla$.
- e) Mostre que para toda conexão $\bar{\nabla}$, existe uma única conexão ∇ com as mesmas geodésicas e torção nula.

11.6 Conexões Métricas

11.16 Lema. *Seja M uma variedade métrica com uma conexão simétrica ∇ , compatível com a métrica de M . Então, para todos campos $X, Y, Z \in \mathcal{T}(M)$, vale*

$$\langle \nabla_X Y, Z \rangle = \frac{1}{2} (X \langle Y, Z \rangle + Y \langle X, Z \rangle - Z \langle X, Y \rangle - \langle X, [Y, Z] \rangle - \langle Y, [X, Z] \rangle + \langle Z, [X, Y] \rangle). \quad (11.16)$$

Em particular, uma conexão simétrica compatível com a métrica é unicamente determinada pela métrica.

Prova: Pela Proposição 11.12,

$$\begin{aligned} X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle, \\ Y \langle X, Z \rangle &= \langle \nabla_Y X, Z \rangle + \langle X, \nabla_Y Z \rangle, \\ Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle. \end{aligned}$$

Logo, por simetria,

$$\begin{aligned} X \langle Y, Z \rangle + Y \langle X, Z \rangle - Z \langle X, Y \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle + \langle \nabla_Y X, Z \rangle + \langle X, \nabla_Y Z \rangle \\ &\quad - \langle \nabla_Z X, Y \rangle - \langle X, \nabla_Z Y \rangle \\ &= \langle X, \nabla_Y Z - \nabla_Z Y \rangle + \langle Y, \nabla_X Z - \nabla_Z X \rangle + \langle \nabla_X Y, Z \rangle + \langle \nabla_Y X, Z \rangle \\ &= \langle X, [Y, Z] \rangle + \langle Y, [X, Z] \rangle - \langle \nabla_X Y - \nabla_Y X, Z \rangle + 2 \langle \nabla_X Y, Z \rangle \\ &= \langle X, [Y, Z] \rangle + \langle Y, [X, Z] \rangle - \langle [X, Y], Z \rangle + 2 \langle \nabla_X Y, Z \rangle, \end{aligned}$$

onde segue o resultado. ■

11.17 Teorema. Seja M uma variedade métrica. Então existe uma única conexão simétrica ∇ compatível com a métrica de M .

Prova: O lema anterior mostra como definir uma conexão simétrica compatível com a métrica através da expressão (11.16). Além disso, pelo lema, qualquer conexão simétrica compatível com a métrica satisfará a identidade (11.5), o que estabelece a unicidade. ■

11.18 Definição. Seja M uma variedade métrica. A única conexão simétrica ∇ compatível com a métrica de M é chamada a **conexão métrica** (ou **conexão de Levi-Civita**) de M .

Se M é uma variedade riemanniana, esta conexão é chamada **conexão riemanniana**, se M é uma variedade semi-riemanniana, ela é chamada **conexão semi-riemanniana** e se M é uma variedade de Lorentz, ela é chamada **conexão de Lorentz**. □

Isometrias preservam conexões riemannianas, como esperado:

11.19 Proposição. Sejam (M, g) e (\tilde{M}, \tilde{g}) variedades métricas isométricas com conexões métricas ∇ e $\tilde{\nabla}$, respectivamente. Se $F : M \rightarrow \tilde{M}$ é uma isometria, então

$$F_*(\nabla_X Y) = \tilde{\nabla}_{F_*X}(F_*Y)$$

Em particular, se $\alpha : I \rightarrow M$ é uma curva diferenciável e V é um campo vetorial ao longo de α , então

$$F_*\left(\frac{dV}{dt}\right) = \frac{\tilde{D}(F_*V)}{dt}.$$

Prova: Defina uma aplicação

$$\bar{\nabla} : \mathcal{T}(M) \times \mathcal{T}(M) \rightarrow \mathcal{T}(M)$$

por

$$\bar{\nabla}_X Y = F_*^{-1} [\tilde{\nabla}_{F_*X}(F_*Y)].$$

Mostraremos que $\bar{\nabla}$ é uma conexão métrica em M . A unicidade da conexão métrica garantirá então que

$$\bar{\nabla} = \nabla,$$

o que provará o resultado. De fato, temos

$$\begin{aligned} \bar{\nabla}_{fX+gY} Z &= F_*^{-1} [\tilde{\nabla}_{F_*(fX+gY)}(F_*Z)] \\ &= F_*^{-1} [\tilde{\nabla}_{F_*(fX+gY)}(F_*Z)] \\ &= F_*^{-1} [\tilde{\nabla}_{(f \circ F^{-1})F_*X + (g \circ F^{-1})F_*Y}(F_*Z)] \\ &= F_*^{-1} [(f \circ F^{-1}) \tilde{\nabla}_{F_*X}(F_*Z) + (g \circ F^{-1}) \tilde{\nabla}_{F_*Y}(F_*Z)] \\ &= F_*^{-1} [(f \circ F^{-1}) \tilde{\nabla}_{F_*X}(F_*Z)] + F_*^{-1} [(g \circ F^{-1}) \tilde{\nabla}_{F_*Y}(F_*Z)] \\ &= fF_*^{-1} [\tilde{\nabla}_{F_*X}(F_*Z)] + gF_*^{-1} [\tilde{\nabla}_{F_*Y}(F_*Z)] \\ &= f\bar{\nabla}_X Z + g\bar{\nabla}_Y Z, \end{aligned}$$

$$\begin{aligned} \bar{\nabla}_X(Y+Z) &= F_*^{-1} [\tilde{\nabla}_{F_*X}F_*(Y+Z)] \\ &= F_*^{-1} [\tilde{\nabla}_{F_*X}F_*Y + \tilde{\nabla}_{F_*X}F_*Z] \\ &= F_*^{-1} [\tilde{\nabla}_{F_*X}F_*Y] + F_*^{-1} [\tilde{\nabla}_{F_*X}F_*Z] \\ &= \bar{\nabla}_X Y + \bar{\nabla}_X Z \end{aligned}$$

e

$$\begin{aligned}
\bar{\nabla}_X(fY) &= F_*^{-1} \left[\tilde{\nabla}_{F_*X} F_*(fY) \right] \\
&= F_*^{-1} \left[\tilde{\nabla}_{F_*X} (fF_*Y) \right] \\
&= F_*^{-1} \left[(f \circ F^{-1}) \tilde{\nabla}_{F_*X} F_*Y + [(F_*X)(f \circ F^{-1})] F_*Y \right] \\
&= F_*^{-1} \left[(f \circ F^{-1}) \tilde{\nabla}_{F_*X} F_*Y \right] + F_*^{-1} \left[[(F_*X)(f \circ F^{-1})] F_*Y \right] \\
&= F_*^{-1} \left[(f \circ F^{-1}) \tilde{\nabla}_{F_*X} F_*Y \right] + [(F_*X)(f \circ F^{-1})] F_*^{-1}(F_*Y) \\
&= fF_*^{-1} \left[\tilde{\nabla}_{F_*X} F_*Y \right] + [X(f \circ F^{-1} \circ F)] F_*^{-1}(F_*Y) \\
&= f\bar{\nabla}_X Y + (Xf)Y,
\end{aligned}$$

o que prova que $\bar{\nabla}$ é uma conexão.

Agora verificaremos que $\bar{\nabla}$ é simétrica:

$$\begin{aligned}
\bar{\nabla}_X Y - \bar{\nabla}_Y X &= F_*^{-1} \left[\tilde{\nabla}_{F_*X} (F_*Y) \right] - F_*^{-1} \left[\tilde{\nabla}_{F_*Y} (F_*X) \right] \\
&= F_*^{-1} \left[\tilde{\nabla}_{F_*X} (F_*Y) - \tilde{\nabla}_{F_*Y} (F_*X) \right] \\
&= F_*^{-1} [F_*X, F_*Y] \\
&= F_*^{-1} F_* [X, Y] \\
&= [X, Y].
\end{aligned}$$

Observe que para provar que $\bar{\nabla}$ é uma conexão simétrica, foi suficiente usar o fato que F é um difeomorfismo.

Para estabelecer a compatibilidade de $\bar{\nabla}$ com a métrica de M é necessário usar o fato que F é uma isometria. Com efeito, dados $X, Y, Z \in \mathcal{T}(M)$, sejam $V = F_*Y$ e $W = F_*Z$. Então temos

$$\begin{aligned}
&\langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \\
&= \left\langle F_*^{-1} \left[\tilde{\nabla}_{F_*X} (F_*Y) \right], F_*^{-1} W \right\rangle + \left\langle F_*^{-1} V, F_*^{-1} \left[\tilde{\nabla}_{F_*X} (F_*Z) \right] \right\rangle \\
&= \left\langle \tilde{\nabla}_{F_*X} (F_*Y), W \right\rangle + \left\langle V, \tilde{\nabla}_{F_*X} (F_*Z) \right\rangle \\
&= \left\langle \tilde{\nabla}_{F_*X} (F_*Y), F_*Z \right\rangle + \left\langle F_*Y, \tilde{\nabla}_{F_*X} (F_*Z) \right\rangle \\
&= F_*X \langle F_*Y, F_*Z \rangle \\
&= X \langle Y, Z \rangle.
\end{aligned}$$

A última passagem merece ser mais detalhada: definindo $f : N \rightarrow \mathbb{R}$ por

$$f(q) = \left\langle (F_*Y)_q, (F_*Z)_q \right\rangle_q,$$

por isometria segue que se $p = F^{-1}(q)$ então

$$f(q) = \langle Y_p, Z_p \rangle_p = \langle Y_{F^{-1}(q)}, Z_{F^{-1}(q)} \rangle_{F^{-1}(q)},$$

isto é,

$$f = \langle Y, Z \rangle \circ F^{-1};$$

assim, usando a propriedade

$$(F_*X)f = X(f \circ F) \circ F^{-1},$$

temos

$$F_* X \langle F_* Y, F_* Z \rangle = X (f \circ F) \circ F^{-1} = X \langle Y, Z \rangle \circ F^{-1},$$

ou seja,

$$[\langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle] (p) = [X \langle Y, Z \rangle] (p).$$

■

11.20 Exemplo (Conexão Métrica em \mathbb{R}^n). A conexão euclidiana definida no Exemplo 11.4 é a conexão riemanniana de \mathbb{E}^n , é a conexão semi-riemanniana de \mathbb{M}_k^n e é a conexão de Lorentz do espaço de Minkowski \mathbb{M}^n . De fato, a conexão é compatível com a métrica, pois se $\alpha : I \rightarrow \mathbb{R}^n$ é uma curva diferenciável e V, W são campos ao longo de α induzidos pelos campos vetoriais X, Y , respectivamente, então segue da regra da cadeia que

$$\begin{aligned} \frac{d}{dt} \langle V_{\alpha(t)}, W_{\alpha(t)} \rangle &= \left\langle \frac{d}{dt} V_{\alpha(t)}, W_{\alpha(t)} \right\rangle + \left\langle V_{\alpha(t)}, \frac{d}{dt} W_{\alpha(t)} \right\rangle \\ &= \langle dX_{\alpha(t)}(\alpha'(t)), W_{\alpha(t)} \rangle + \langle V_{\alpha(t)}, dY_{\alpha(t)}(\alpha'(t)) \rangle \\ &= \left\langle (\nabla_{\alpha'(t)} X)_{\alpha(t)}, W_{\alpha(t)} \right\rangle + \left\langle V_{\alpha(t)}, (\nabla_{\alpha'(t)} Y)_{\alpha(t)} \right\rangle \\ &= \left\langle \frac{DV}{dt} \Big|_{\alpha(t)}, W_{\alpha(t)} \right\rangle + \left\langle V_{\alpha(t)}, \frac{DW}{dt} \Big|_{\alpha(t)} \right\rangle. \end{aligned}$$

e ela é simétrica porque, conforme (11.5),

$$\begin{aligned} [(\nabla_X Y)_p - (\nabla_Y X)_p](f) &= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x_j} - Y^i \frac{\partial X^j}{\partial x^i} \frac{\partial f}{\partial x_j} \\ &= \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial f}{\partial x_j} \\ &= [X, Y]_p(f). \end{aligned}$$

ou também, de forma mais sucinta, conforme (11.6),

$$\begin{aligned} \nabla_X Y - \nabla_Y X &= X(Y^j) \frac{\partial}{\partial x_j} - Y(X^j) \frac{\partial}{\partial x_j} \\ &= (X(Y^j) - Y(X^j)) \frac{\partial}{\partial x_j} \\ &= [X, Y]. \end{aligned}$$

□

11.6.1 Símbolos de Christoffel da Conexão Métrica

Vamos agora ver como os símbolos de Christoffel de uma conexão métrica podem ser calculados através dos componentes g_{ij} da métrica. Antes introduzimos a seguinte notação: a matriz $G = (g_{ij})$ é uma matriz positiva definida, logo admite uma inversa, que denotaremos por

$$G^{-1} = (g^{ij}). \quad (11.17)$$

A justificativa para isso será vista no Exemplo 13.11.

11.21 Lema. *Seja M uma variedade métrica com uma conexão ∇ . Então*

$$\langle \nabla_{\partial_i} \partial_j, \partial_k \rangle = \Gamma_{ij}^m g_{mk}. \quad (11.18)$$

Prova: Segue imediatamente da definição dos símbolos de Christoffel:

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^m \partial_m.$$

■

11.22 Lema. Seja M uma variedade métrica com uma conexão métrica ∇ . Então

$$\langle \nabla_{\partial_i} \partial_j, \partial_k \rangle = \frac{1}{2} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}). \quad (11.19)$$

Prova: Por (11.16) temos que

$$\begin{aligned} \langle \nabla_{\partial_i} \partial_j, \partial_k \rangle &= \frac{1}{2} (\partial_i \langle \partial_j, \partial_k \rangle + \partial_j \langle \partial_i, \partial_k \rangle - \partial_k \langle \partial_i, \partial_j \rangle - \langle \partial_i, [\partial_j, \partial_k] \rangle - \langle \partial_j, [\partial_i, \partial_k] \rangle + \langle \partial_k, [\partial_i, \partial_j] \rangle) \\ &= \frac{1}{2} (\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij}), \end{aligned}$$

já que $[\partial_m, \partial_l] = 0$. ■

11.23 Proposição. Seja M uma variedade métrica com uma conexão métrica ∇ . Então

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk}. \quad (11.20)$$

Prova: Pelos lemas temos

$$g_{lm} \Gamma_{ij}^l = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}).$$

Logo

$$g^{mk} g_{lm} \Gamma_{ij}^l = \frac{1}{2} \sum_{m=1}^n (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk}.$$

O lado esquerdo desta equação é

$$g^{km} g_{ml} \Gamma_{ij}^l = \delta_l^k \Gamma_{ij}^l = \Gamma_{ij}^k.$$

■

11.24 Corolário. Se ∇ é a conexão métrica de \mathbb{E}^n , \mathbb{M}_k^n e \mathbb{M}^n , então

$$\Gamma_{ij}^k = 0. \quad (11.21)$$

Consequentemente,

$$\nabla_X Y = \sum_{k=1}^n X(Y^k) \partial_k \quad (11.22)$$

e

$$\nabla_{\partial_i} \partial_j = 0. \quad (11.23)$$

Prova: Pois, como

$$g_{ij} = \eta_{ij}^k \equiv \text{constante},$$

segue que

$$\partial_k g_{ij} = 0$$

para todos os índices i, j, k . ■

11.25 Corolário. *Seja M uma variedade métrica com uma conexão métrica ∇ . Então*

$$\partial_k g_{ij} = g_{ip} \Gamma_{jk}^p + g_{pj} \Gamma_{ik}^p \quad (11.24)$$

e

$$\partial_k g^{ij} = -g^{ip} \Gamma_{pk}^j - g^{pj} \Gamma_{pk}^i. \quad (11.25)$$

Prova: Para provar a primeira identidade, usando a compatibilidade da métrica temos

$$\begin{aligned} \partial_k g_{ij} &= \frac{\partial}{\partial x_k} \langle \partial_i, \partial_j \rangle = \langle \nabla_{\partial_k} \partial_i, \partial_j \rangle + \langle \partial_i, \nabla_{\partial_k} \partial_j \rangle \\ &= \langle \Gamma_{ki}^p \partial_p, \partial_j \rangle + \langle \partial_i, \Gamma_{kj}^p \partial_p \rangle \\ &= \Gamma_{ki}^p \langle \partial_p, \partial_j \rangle + \Gamma_{kj}^p \langle \partial_i, \partial_p \rangle \\ &= \Gamma_{ki}^p g_{pj} + \Gamma_{kj}^p g_{ip}. \end{aligned}$$

Para provar a segunda identidade, primeiro diferenciamos a identidade

$$g_{lp} g^{pj} = \delta_l^j,$$

obtendo

$$g_{lp} \partial_k g^{pj} = -(\partial_k g_{lp}) g^{pj}.$$

Como

$$\begin{aligned} g^{il} g_{lp} \partial_k g^{pj} &= g^{il} g_{lp} \partial_k g^{pj} \\ &= \delta_p^i \partial_k g^{pj} \\ &= \partial_k g^{ij}, \end{aligned}$$

segue da primeira identidade que

$$\begin{aligned} \partial_k g^{ij} &= -g^{il} (\partial_k g_{lp}) g^{pj} = -g^{il} g^{pj} \partial_k g_{lp} \\ &= -g^{il} g^{pj} (g_{lm} \Gamma_{pk}^m + g_{mp} \Gamma_{lk}^m) \\ &= -g^{pj} g^{il} g_{lm} \Gamma_{pk}^m - g^{il} g^{pj} g_{mp} \Gamma_{lk}^m \\ &= -g^{pj} \delta_m^i \Gamma_{pk}^m - g^{il} \delta_m^j \Gamma_{lk}^p \\ &= -g^{pj} \Gamma_{pk}^i - g^{il} \Gamma_{lk}^j. \end{aligned}$$

■

11.6.2 Interpretação Geométrica da Derivada Covariante

Usando a conexão métrica, podemos dar uma interpretação geométrica da derivada covariante em termos do transporte paralelo.

11.26 Lema. *Se M é uma variedade métrica com uma conexão métrica ∇ , então a aplicação transporte paralelo é uma isometria que preserva orientação.*

Prova: Seja $\alpha : I \rightarrow M$ uma curva diferenciável passando por um ponto $p \in M$. Dados $V, W \in T_p M$, considere a função real $f : I \rightarrow \mathbb{R}$ definida por

$$f(t) = \langle \tau_t(V), \tau_t(W) \rangle.$$

Pela compatibilidade da métrica, segue que

$$f'(t) = \left\langle \frac{D}{dt} \tau_t(V), \tau_t(W) \right\rangle + \left\langle \tau_t(V), \frac{D}{dt} \tau_t(W) \right\rangle = 0,$$

já que os campos $\tau_t(V), \tau_t(W)$ são paralelos ao longo de α por definição. Portanto, $f(t) = f(0)$ para todo $t \in I$, ou seja

$$\langle \tau_t(V), \tau_t(W) \rangle = \langle V, W \rangle$$

o que prova que τ_t é uma isometria.

Para provar que τ_t preserva orientação, seja

$$\mathcal{B} = \{E_1, \dots, E_n\}$$

uma base ortonormal positivamente orientada para $T_p M$. Como τ_t é uma isometria,

$$\mathcal{B}_t = \{\tau_t(E_1), \dots, \tau_t(E_n)\}$$

é uma base ortonormal de $T_{\alpha(t)} M$ para todo $t \in I$. A orientação positiva de \mathcal{B}_t segue por continuidade da função determinante. ■

11.27 Proposição (Interpretação Geométrica da Derivada Covariante). *Seja M uma variedade métrica com uma conexão métrica ∇ . Dado um campo $X \in \mathcal{T}(M)$, seja $\alpha : I \rightarrow M$ uma curva integral do campo X passando por p , ou seja,*

$$\begin{aligned} \alpha(0) &= p, \\ \alpha'(t) &= X(\alpha(t)) \quad \text{para todo } t \in I. \end{aligned}$$

Se $Y \in \mathcal{T}(M)$, então

$$(\nabla_X Y)_p = \lim_{t \rightarrow 0} \frac{\tau_t^{-1}(Y_{\alpha(t)}) - Y_p}{t} = \left. \frac{d}{dt} \tau_t^{-1}(Y_{\alpha(t)}) \right|_{t=0}.$$

Prova: Seja $\mathcal{B} = \{E_1, \dots, E_n\}$ uma base ortonormal para $T_p M$. Pelo lema, $\mathcal{B}_t = \{\tau_t(E_1), \dots, \tau_t(E_n)\}$ é uma base ortonormal de $T_{\alpha(t)} M$ para todo $t \in I$. Como a aplicação transporte paralelo é linear, se

$$Y_{\alpha(t)} = Y^i(t) \tau_t(E_i),$$

segue que

$$\tau_t^{-1}(Y_{\alpha(t)}) = Y^i(t) E_i.$$

Logo,

$$\left. \frac{d}{dt} \tau_t^{-1}(Y_{\alpha(t)}) \right|_{t=0} = \left. \frac{dY^i}{dt}(0) E_i \right|_{t=0}.$$

Por outro lado, pela Proposição 11.7, temos também

$$\begin{aligned} (\nabla_X Y)_p &= \left. \frac{D}{dt} (Y^i(t) \tau_t(E_i)) \right|_{t=0} \\ &= \left. \frac{dY^i}{dt}(0) \tau_0(E_i) + Y^i(t) \frac{D}{dt}(\tau_t(E_i)) \right|_{t=0} \\ &= \left. \frac{dY^i}{dt}(0) E_i \right|_{t=0}, \end{aligned}$$

já que os campos $\tau_t(E_1), \dots, \tau_t(E_n)$ são paralelos ao longo de α por definição. ■

Este resultado também vale para conexões gerais em variedades diferenciáveis, mas o resultado requer um pouco mais de conhecimento de fibrados (veja [Dodson-Poston], Theorem 4.05, pp. 226–227).

Capítulo 12

Geodésicas

De agora em diante, sempre que nos referirmos a uma variedade métrica, estaremos supondo que ela está munida da sua conexão métrica.

12.1 Definição – A Equação Geodésica

12.1 Definição. Seja M uma variedade métrica. Dizemos que uma curva diferenciável $\gamma : I \rightarrow M$ é uma **geodésica** se

$$\frac{D\gamma'}{dt}(t) = 0$$

para todo $t \in I$. \square

Em outras palavras, uma geodésica é uma curva cujo campo velocidade é paralelo ao longo da curva (uma curva que transporta paralelamente o seu próprio vetor tangente). Ou seja, uma geodésica é uma curva que não muda de direção. Às vezes, por abuso de linguagem, a imagem $\gamma(I)$ de uma geodésica γ também é chamada geodésica.

Note que o conceito de geodésica pode ser definido para qualquer variedade diferenciável dotada de uma conexão. O resultado seguinte, no entanto, depende da compatibilidade da métrica com a conexão, ou seja, requer uma conexão métrica.

12.2 Proposição. Se $\gamma : I \rightarrow M$ é uma geodésica, então

$$\|\gamma'(t)\| \equiv \text{constante}. \quad (12.1)$$

Consequentemente, uma reparametrização $\gamma \circ h$ de uma geodésica γ é uma geodésica se e somente se

$$h(t) = at + b$$

para algumas constantes $a, b \in \mathbb{R}$.

Prova: Pois, como a conexão é compatível com a métrica e o campo velocidade γ' é paralelo ao longo de γ

$$\|\gamma'(t)\|^2 = \langle \gamma'(t), \gamma'(t) \rangle \equiv \text{constante}.$$

Como

$$\begin{aligned} (\gamma \circ h)'(t) &= h'(t) \gamma'(h(t)), \\ \|\gamma'(t)\| &\equiv \text{constante}, \end{aligned}$$

e, como acabamos de provar, uma condição necessária para que $\gamma \circ h$ seja uma geodésica é

$$\|(\gamma \circ h)'(t)\| \equiv \text{constante},$$

concluímos que uma condição necessária para que a reparametrização $\gamma \circ h$ seja uma geodésica é que

$$h'(t) \equiv \text{constante},$$

ou seja,

$$h(t) = at + b$$

para algumas constantes $a, b \in \mathbb{R}$. Além disso, escrevendo $\beta(t) = \gamma \circ h(t) = \gamma(at + b)$, segue que

$$\frac{D\beta'}{dt}(t) = a^2 \frac{D\gamma'}{dt}(at + b) = 0,$$

logo $\gamma \circ h$ é uma geodésica. ■

12.3 Definição. Uma geodésica $\gamma : I \rightarrow M$ é **normalizada** (ou **unitária**) se

$$\|\gamma'(t)\| \equiv 1.$$

□

Toda geodésica que não é um ponto (ou seja, $\|\gamma'(t)\| \neq 0$) pode ser normalizada através de uma *parametrização por comprimento de arco*: se $\gamma : I \rightarrow M$ é uma parametrização qualquer para uma geodésica, ela pode ser reparametrizada para se tornar uma geodésica normalizada escolhendo-se um ponto $t_0 \in I$ e definindo o parâmetro **comprimento de arco**

$$s(t) = \int_{t_0}^t \|\gamma'(t)\| dt. \quad (12.2)$$

De fato, pela regra da cadeia

$$\|\gamma'(s)\| = \|\gamma'(t)\| |t'(s)| = \|\gamma'(t)\| \frac{1}{|s'(t)|} = \|\gamma'(t)\| \frac{1}{\|\gamma'(t)\|} = 1.$$

12.4 Teorema (Teorema de Existência e Unicidade de Geodésicas). Seja M uma variedade métrica. Então para todos $p \in M$ e $v \in T_p M$, e para cada $t_0 \in \mathbb{R}$, existe um intervalo aberto $I \subset \mathbb{R}$ contendo t_0 e uma única geodésica $\gamma : I \rightarrow M$ tal que $\gamma(t_0) = p$ e $\gamma'(t_0) = v$.

Prova: Seja V uma vizinhança coordenada de p , com (x^1, \dots, x^n) suas coordenadas. Pela expressão em coordenadas locais da derivada covariante de um campo ao longo de uma curva obtida no capítulo anterior, uma curva $\gamma(t) = x(t) = (x^1(t), \dots, x^n(t))$ é uma geodésica se e somente se as suas componentes satisfazem o sistema de equações diferenciais ordinárias de segunda ordem não linear (quasilinear), chamado a **equação geodésica**,

$$\frac{d^2x^k}{dt^2} + \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} = 0, \quad k = 1, \dots, n. \quad (12.3)$$

Este sistema de segunda ordem pode ser transformado num sistema de primeira ordem introduzindo as n equações de primeira ordem

$$v^k = \frac{dx^k}{dt}, \quad k = 1, \dots, n,$$

de modo que estas equações juntamente com

$$\frac{dv^k}{dt} + \Gamma_{ij}^k v^i v^j = 0, \quad k = 1, \dots, n,$$

formam um sistema de primeira ordem equivalente ao primeiro. O resultado segue então do teorema de existência e unicidade para soluções de sistemas de equações diferenciais ordinárias de primeira ordem. ■
Note que este teorema permanece válido para geodésicas definidas em variedades diferenciáveis dotadas de uma conexão não necessariamente métrica.

Isometrias preservam geodésicas:

12.5 Proposição. *Sejam M, N variedades métricas isométricas e seja $F : M \rightarrow N$ uma isometria. Se $\gamma : I \rightarrow M$ é uma geodésica de M tal que*

$$\begin{aligned}\gamma(0) &= p, \\ \gamma'(0) &= v,\end{aligned}$$

então $\beta = F \circ \gamma : I \rightarrow N$ é uma geodésica de N tal que

$$\begin{aligned}\beta(0) &= F(p), \\ \beta'(0) &= dF_p(v).\end{aligned}$$

Prova: Segue imediatamente do fato de isometrias preservarem conexões métricas (Proposição 11.19) e da regra da cadeia. ■

12.2 Exemplos de Geodésicas

Vamos calcular as geodésicas para algumas variedades riemannianas. Lembramos que os símbolos de Christoffel da conexão riemanniana são dados por

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{km}.$$

12.6 Exemplo (Geodésicas de \mathbb{R}^n). Como o tensor métrico de \mathbb{R}^n é $g_{ij} = \delta_{ij}$, temos

$$\Gamma_{ij}^k \equiv 0,$$

onde a equação geodésica é simplesmente

$$\frac{d^2 x^k}{dt^2} = 0, \quad k = 1, \dots, n.$$

As soluções para esta equação diferencial são

$$x(t) = tv + x_0$$

onde $v, x_0 \in \mathbb{R}^n$ são vetores fixados. Em outras palavras, as geodésicas de \mathbb{R}^n são retas. □

12.7 Exemplo (Geodésicas de \mathbb{S}_r^2). Enxergando a esfera \mathbb{S}_r^2 de centro na origem e raio r como uma superfície de revolução com carta

$$(x, y, z)(\phi, \theta) = (r \sin \phi \cos \theta, r \sin \phi \sin \theta, r \cos \phi),$$

segue do Exemplo 10.11 que

$$\begin{aligned}G(\phi, \theta) &= \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \sin^2 \phi \end{bmatrix}, \\ G^{-1}(\phi, \theta) &= \begin{bmatrix} \frac{1}{r^2} & 0 \\ 0 & \frac{1}{r^2 \sin^2 \phi} \end{bmatrix}.\end{aligned}$$

Para variedades de dimensão 2 temos

$$\Gamma_{ij}^k = \frac{1}{2} [(\partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij}) g^{k1} + (\partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij}) g^{k2}].$$

Nesta carta,

$$\begin{aligned} g_{11} &\equiv \text{constante}, \\ g_{12} = g_{21} = g^{12} = g^{21} &= 0, \\ \partial_2 g_{ij} &= 0. \end{aligned}$$

Portanto,

$$\begin{aligned} \Gamma_{ij}^1 &= \frac{1}{2} (\partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij}) g^{11}, \\ \Gamma_{ij}^2 &= \frac{1}{2} (\partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij}) g^{22} = \frac{1}{2} (\partial_i g_{j2} + \partial_j g_{i2}) g^{22}. \end{aligned}$$

Logo,

$$\begin{aligned} \Gamma_{11}^1 &= \frac{1}{2} (\partial_1 g_{11} + \partial_1 g_{11} - \partial_1 g_{11}) g^{11} = 0, \\ \Gamma_{12}^1 = \Gamma_{21}^1 &= \frac{1}{2} (\partial_1 g_{21} + \partial_2 g_{11} - \partial_1 g_{12}) g^{11} = 0, \\ \Gamma_{22}^1 &= \frac{1}{2} (\partial_2 g_{21} + \partial_2 g_{21} - \partial_1 g_{22}) g^{11} = \frac{1}{2} (-\partial_1 g_{22}) g^{11} = \frac{1}{2} (-2r^2 \sen \phi \cos \phi) \frac{1}{r^2} = -\sen \phi \cos \phi, \\ \Gamma_{11}^2 &= \frac{1}{2} (\partial_1 g_{12} + \partial_1 g_{12}) g^{22} = 0, \\ \Gamma_{12}^2 = \Gamma_{21}^2 &= \frac{1}{2} (\partial_1 g_{22} + \partial_2 g_{12}) g^{22} = \frac{1}{2} (\partial_1 g_{22}) g^{22} = \frac{1}{2} (2r^2 \sen \phi \cos \phi) \frac{1}{r^2 \sen^2 \phi} = \frac{\cos \phi}{\sen \phi}, \\ \Gamma_{22}^2 &= \frac{1}{2} (\partial_2 g_{22} + \partial_2 g_{22}) g^{22} = 0. \end{aligned}$$

Apenas os símbolos $\Gamma_{12}^2 = \Gamma_{21}^2$ e Γ_{22}^1 são não nulos. Portanto a equação geodésica para a esfera é

$$\begin{cases} \frac{d^2 x^1}{dt^2} + \Gamma_{22}^1 \frac{dx^2}{dt} \frac{dx^2}{dt} = 0, \\ \frac{d^2 x^2}{dt^2} + 2\Gamma_{12}^2 \frac{dx^1}{dt} \frac{dx^2}{dt} = 0, \end{cases}$$

ou seja,

$$\begin{cases} \frac{d^2 \phi}{dt^2} - \sen \phi \cos \phi \left(\frac{d\theta}{dt} \right)^2 = 0, \\ \frac{d^2 \theta}{dt^2} + 2\frac{\cos \phi}{\sen \phi} \frac{d\phi}{dt} \frac{d\theta}{dt} = 0. \end{cases}$$

Resolver este sistema de equações diferenciais não lineares acopladas não é fácil. Por outro lado, é fácil verificar por substituição direta que o equador

$$\begin{aligned} \phi &= \pi/2, \\ \theta &= t, \end{aligned}$$

ou qualquer meridiano da forma

$$\begin{aligned}\phi &= t, \\ \theta &= \theta_0,\end{aligned}$$

é uma solução para este sistema. Observe agora que rotações são isometrias da esfera: rotações do espaço euclídeo \mathbb{R}^3 são transformações ortogonais e portanto isometrias de \mathbb{R}^3 ; como a restrição de uma rotação de \mathbb{R}^3 à esfera é uma rotação da esfera e esta tem a métrica induzida de \mathbb{R}^3 , ela é uma isometria da esfera. Uma vez que geodésicas são preservadas por isometrias, concluímos que os **grandes círculos** da esfera (isto é, círculos cujos centros são o centro da esfera, que podem também ser obtidos intersectando a esfera com planos passando pela origem) são suas geodésicas, já que dado um ponto p da esfera e um vetor tangente v à esfera neste ponto existe um grande círculo passando por p com a direção de v . \square

12.8 Exemplo (Geodésicas de \mathbb{S}^n). Um argumento envolvendo isometrias pode ser usado para obter diretamente que as geodésicas de $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ são os grandes círculos, sem passar pela equação geodésica.

Primeiramente, provamos que a geodésica $\gamma(t) = (x^1(t), \dots, x^n(t), x^{n+1}(t))$ que passa pelo polo norte $\gamma(0) = e_{n+1}$ com velocidade $\gamma'(0) = e_1$ é o meridiano $x^2 = \dots = x^n = 0$. De fato, suponha por absurdo que $x^i(t_0) \neq 0$ para algum t_0 e para algum índice $2 \leq i \leq n$. Considere a isometria $\phi : \mathbb{S}^n \rightarrow \mathbb{S}^n$ definida por

$$\phi(x^1, \dots, x^{i-1}, x^i, x^{i+1}, \dots, x^{n+1}) = (x^1, \dots, x^{i-1}, -x^i, x^{i+1}, \dots, x^{n+1})$$

(ϕ é a restrição da transformação ortogonal correspondente de \mathbb{R}^{n+1} , logo é uma simetria da esfera, como observado no final do exemplo anterior). Como isometrias preservam geodésicas e

$$\begin{aligned}\phi(\gamma(0)) &= \phi(e_{n+1}) = e_{n+1}, \\ d\phi_{\gamma(0)}\gamma'(0) &= \phi(e_1) = e_1,\end{aligned}$$

segue que ϕ leva γ em γ . Mas $\phi(\gamma(t_0)) \neq \gamma(t_0)$, uma contradição.

Como rotações da esfera que deixam o eixo x^{n+1} fixado transformam este meridiano em qualquer outro meridiano passando pelo polo norte, abrangendo todas as direções tangentes possíveis, obtemos que as geodésicas passando pelo polo norte são os meridianos. Como qualquer grande círculo da esfera pode ser transformado em um meridiano passando pelo polo norte através de uma rotação da esfera, concluímos o resultado. \square

12.9 Exemplo (Geodésicas de \mathbb{H}^2). Como vimos no Exemplo 10.12, temos

$$\begin{aligned}G(x, y) &= \begin{bmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{bmatrix}, \\ G^{-1}(x, y) &= \begin{bmatrix} y^2 & 0 \\ 0 & y^2 \end{bmatrix}.\end{aligned}$$

Para variedades de dimensão 2 temos

$$\Gamma_{ij}^k = \frac{1}{2} [(\partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij}) g^{k1} + (\partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij}) g^{k2}].$$

Nesta carta,

$$\begin{aligned}g_{12} &= g_{21} = g^{12} = g^{21} = 0, \\ \partial_1 g_{ij} &= 0.\end{aligned}$$

Portanto,

$$\begin{aligned}\Gamma_{ij}^1 &= \frac{1}{2} (\partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij}) g^{11} = \frac{1}{2} (\partial_i g_{j1} + \partial_j g_{i1}) g^{11}, \\ \Gamma_{ij}^2 &= \frac{1}{2} (\partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij}) g^{22}.\end{aligned}$$

Logo,

$$\begin{aligned}\Gamma_{11}^1 &= \frac{1}{2} (\partial_1 g_{11} + \partial_1 g_{11}) g^{11} = 0, \\ \Gamma_{12}^1 &= \Gamma_{21}^1 = \frac{1}{2} (\partial_1 g_{21} + \partial_2 g_{11}) g^{11} = \frac{1}{2} (\partial_2 g_{11}) g^{11} = \frac{1}{2} \left(-\frac{2}{y^3} \right) y^2 = -\frac{1}{y}, \\ \Gamma_{22}^1 &= \frac{1}{2} (\partial_2 g_{21} + \partial_2 g_{21}) g^{11} = 0,\end{aligned}$$

e

$$\begin{aligned}\Gamma_{11}^2 &= \frac{1}{2} (\partial_1 g_{12} + \partial_1 g_{12} - \partial_2 g_{11}) g^{22} = \frac{1}{2} (-\partial_2 g_{11}) g^{22} = \frac{1}{2} \left(\frac{2}{y^3} \right) y^2 = \frac{1}{y}, \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{1}{2} (\partial_1 g_{22} + \partial_2 g_{12} - \partial_2 g_{12}) g^{22} = 0, \\ \Gamma_{22}^2 &= \frac{1}{2} (\partial_2 g_{22} + \partial_2 g_{22} - \partial_2 g_{22}) g^{22} = \frac{1}{2} (\partial_2 g_{22}) g^{22} = \frac{1}{2} \left(-\frac{2}{y^3} \right) y^2 = -\frac{1}{y}.\end{aligned}$$

A equação geodésica para o plano hiperbólico é

$$\begin{cases} x'' - \frac{2}{y} x' y' = 0, \\ y'' + \frac{1}{y} (x')^2 - \frac{1}{y} (y')^2 = 0, \end{cases}$$

Para resolver este sistema, consideraremos dois casos.

(i) Caso $x' = 0$.

Neste caso, $x(t) \equiv x_0$, enquanto que $y(t)$ satisfaz a equação

$$y'' - \frac{1}{y} (y')^2 = 0$$

que escrevemos na forma

$$yy'' - (y')^2 = 0$$

que é equivalente a

$$\left(\frac{y'}{y} \right)' = 0,$$

ou seja,

$$y' = ky$$

para alguma constante $k \in \mathbb{R}$. A solução desta equação é

$$y(t) = y_0 e^{kt}$$

para alguma constante $y_0 > 0$. Portanto, neste caso as geodésicas são $\gamma(t) = (x_0, y_0 e^{kt})$, as semi-retas superiores do plano hiperbólico.

(ii) Caso $x' \neq 0$.

Neste caso podemos calcular

$$\begin{aligned}
 \left(\frac{yy'}{x'} \right)' &= \frac{x'yy'' + x'(y')^2 - x''yy'}{(x')^2} \\
 &= \frac{x'yy'' + x'(x')^2 - x'(y')^2 - x''yy' + 2x'(y')^2 - x'(x')^2}{(x')^2} \\
 &= \frac{x' [yy'' + (x')^2 - (y')^2] - y' [x''y - 2x'y'] - (x')^3}{(x')^2} \\
 &= \frac{x'0 - y'0 - (x')^3}{(x')^2} \\
 &= -x'.
 \end{aligned}$$

Isso implica que

$$\left(\frac{yy'}{x'} + x \right)' = 0$$

ou seja,

$$\frac{yy'}{x'} + x = a$$

para alguma constante $a \in \mathbb{R}$. A solução implícita desta equação, equivalente a

$$xx' + yy' = ax'$$

pode ser obtida integrando-se em relação a t :

$$\frac{x^2}{2} + \frac{y^2}{2} = ax + b,$$

ou

$$(x - a)^2 + y^2 = R^2$$

onde $R > 0$. Portanto, neste caso as geodésicas são semicírculos superiores centrados em pontos $(a, 0)$ do eixo x . \square

12.3 Fluxo Geodésico

O resultado obtido pode ser melhorado, no sentido de que em uma vizinhança de qualquer ponto da variedade podemos garantir que os intervalos de definição das geodésicas passando por pontos na vizinhança, independentemente da velocidade inicial, possuem um comprimento comum mínimo. Isto segue do seguinte resultado para o fluxo induzido por um campo vetorial:

12.10 Teorema (Fluxo de um Campo Vetorial). *Sejam M uma variedade diferenciável, V uma vizinhança de um ponto $p \in M$ e $X \in \mathcal{T}(V)$ um campo vetorial. Então existem um aberto $V_0 \subset V$ contendo p , $\delta > 0$ e uma aplicação diferenciável*

$$\varphi : (-\delta, \delta) \times V_0 \longrightarrow V$$

tais que $\varphi_q(t) = \varphi(t, q) : (-\delta, \delta) \longrightarrow V$ é a única trajetória de X tal que $\varphi(0, q) = q$ para todo $q \in V_0$.

Este resultado permite definir:

12.11 Definição. O **campo geodésico** no fibrado tangente TM é o único campo cujas trajetórias são da forma $t \mapsto (\gamma(t), \gamma'(t))$ onde γ é uma geodésica em M .

Seu fluxo φ é chamado o **fluxo geodésico** de TM . \square

12.12 Corolário. Seja M uma variedade riemanniana. Então, para cada $p \in M$ existem uma vizinhança \mathcal{V} de $(p, 0)$ em TM tal que $\pi(\mathcal{V}) = U$ é uma vizinhança coordenada de p em M , $\delta > 0$ e uma aplicação diferenciável

$$\varphi : (-\delta, \delta) \times \mathcal{V} \longrightarrow TU$$

tais que

$$\varphi_{(q,v)}(t) = \varphi(t, q, v) : (-\delta, \delta) \longrightarrow TU$$

é a única trajetória do campo geodésico que satisfaz a condição inicial $\varphi(0, q, v) = (q, v)$ para todo $(q, v) \in \mathcal{V}$.

Prova: Basta aplicar o Teorema 12.10 ao campo geodésico, elemento de $\mathcal{T}(TM)$, e ao ponto $P = (p, 0) \in TM$. ■

Uma reinterpretação do resultado acima que vai ser mais útil e fácil de usar na sequência:

12.13 Proposição. Seja M uma variedade riemanniana. Então, para cada $p \in M$ existem uma vizinhança V de p em M , $\delta, \varepsilon > 0$ e uma aplicação diferenciável

$$\gamma : (-\delta, \delta) \times \mathcal{V} \longrightarrow M$$

onde

$$\mathcal{V} = \{(q, v) : q \in V \text{ e } v \in T_q M, \|v\| < \varepsilon\}$$

tais que

$$\gamma_{(q,v)}(t) = \gamma(t, q, v) : (-\delta, \delta) \longrightarrow M$$

é a única geodésica de M que satisfaz as condições iniciais $\gamma(0) = q, \gamma'(0) = v$ para todo $q \in V$ e para todo $v \in T_q M$ tal que $\|v\| < \varepsilon$.

Este resultado afirma que se $\|v\| < \varepsilon$, a geodésica $\gamma(t, q, v)$ existe no intervalo $(-\delta, \delta)$. Podemos de fato aumentar a velocidade da geodésica, mas ao preço de diminuir o seu intervalo de definição (e vice-versa). Para isso basta reparametrizar a geodésica:

12.14 Proposição. Seja $\gamma : (-\delta, \delta) \longrightarrow M$ uma geodésica tal que $\gamma(0) = p$ e $\gamma'(0) = v$. Para qualquer $a > 0$ a curva

$$\beta(t) = \gamma(at)$$

é uma geodésica tal que $\beta(0) = p$ e $\beta'(0) = av$ definida no intervalo $\left(-\frac{\delta}{a}, \frac{\delta}{a}\right)$. Consequentemente,

$$\gamma(t, q, av) = \gamma(at, q, v). \quad (12.4)$$

Prova: O resultado segue da Proposição 12.2 e do fato que $\beta'(t) = a\gamma'(at)$. ■

12.15 Corolário. Seja M uma variedade riemanniana. Então, para cada $p \in M$ existem uma vizinhança V de p em M , $\varepsilon > 0$ e uma aplicação diferenciável

$$\gamma : (-2, 2) \times \mathcal{V} \longrightarrow M$$

onde

$$\mathcal{V} = \{(q, v) : q \in V \text{ e } v \in T_q M, \|v\| < \varepsilon\}$$

tais que

$$\gamma_{(q,v)}(t) = \gamma(t, q, v) : (-2, 2) \longrightarrow M$$

é a única geodésica de M que satisfaz as condições iniciais $\gamma(0) = q, \gamma'(0) = v$ para todo $q \in V$ e para todo $v \in T_q M$ tal que $\|v\| < \varepsilon$.

12.4 A Aplicação Exponencial

12.16 Definição. Se $v \in T_p M$ é tal que existe uma geodésica $\gamma : [0, 1] \rightarrow M$ satisfazendo $\gamma(0) = p$ e $\gamma'(0) = v$, definimos a **exponencial** de v por

$$\exp_p(v) = \gamma(1).$$

□

Assim,

$$\exp_p(v) = \gamma(1, p, v) = \gamma\left(\|v\|, p, \frac{v}{\|v\|}\right)$$

e geometricamente $\exp_p(v)$ é o ponto de M obtido percorrendo a partir de p a geodésica com velocidade inicial v durante um intervalo de tempo unitário ou, equivalentemente, percorrendo a geodésica que parte de p com velocidade unitária igual a $v/\|v\|$ um comprimento igual a $\|v\|$. Em particular, se $\gamma : (-2, 2) \times \mathcal{V} \rightarrow M$ é a aplicação diferenciável dada pelo Corolário 12.15, a aplicação exponencial \exp está definida em todo \mathcal{V} e para todo $q \in V$ a aplicação exponencial

$$\exp_q : B_\varepsilon(0) \subset T_q M \rightarrow M$$

está definida em uma bola $B_\varepsilon(0) \subset T_q M$.

12.17 Proposição. Para cada $p \in M$ existe $\varepsilon > 0$ tal que \exp_p é um difeomorfismo de $B_\varepsilon(0)$ sobre um aberto de M .

Além disso,

$$d(\exp_p)_0 = \text{id}.$$

Prova: Temos

$$d(\exp_p)_0 v = \frac{\partial \exp_p}{\partial v}(0) = \frac{d}{dt} \exp_p(tv) \Big|_{t=0} = \frac{d}{dt} \gamma(1, p, tv) \Big|_{t=0} = \frac{d}{dt} \gamma(t, p, v) \Big|_{t=0} = v,$$

o que mostra que $d(\exp_p)_0 : T_p M \rightarrow T_p M$ é a aplicação identidade, onde identificamos o espaço tangente a $T_p M$ na origem com o próprio $T_p M$, isto é, $T_0(T_p M) = T_p M$. Pelo teorema da função inversa, \exp_p é um difeomorfismo local em uma vizinhança de 0. ■

A aplicação exponencial recebe este nome porque no caso de variedades riemannianas que são grupos de Lie compactos ela é a aplicação exponencial entre grupos de Lie, a qual recebe este nome lá porque ela é a exponencial de matrizes no caso de grupos de Lie de matrizes.

12.18 Definição. Se \exp_p é um difeomorfismo de uma vizinhança U da origem em $T_p M$ sobre uma vizinhança V de p em M , dizemos que V é uma **vizinhança normal** de p .

Nestas condições, se $B_\varepsilon(0)$ é tal que $\overline{B_\varepsilon(0)} \subset V$, então nós chamamos

$$B_\varepsilon(p) := \exp_p(B_\varepsilon(0)) \tag{12.5}$$

a **bola geodésica** de centro em p e raio ε . □

Em uma vizinhança normal, podemos definir um sistema de coordenadas com propriedades especiais.

12.19 Definição. Seja $V = \exp_p(U)$ uma vizinhança normal de p em M . Seja $\{E_1, \dots, E_n\}$ uma base ortonormal para $T_p M$ e $E : \mathbb{R}^n \rightarrow T_p M$ o isomorfismo $Ex = x^i E_i$. Considere a parametrização $\varphi : E^{-1}(U) \rightarrow V$ definida por

$$\varphi(x) = (\exp_p \circ E)(x) = \exp_p(x^i E_i).$$

As coordenadas desta parametrização são chamadas **coordenadas normais** em p . □

Devido às propriedades especiais de coordenadas normais enumeradas na proposição a seguir, elas constituem uma ferramenta vital em geometria riemanniana, como veremos em vários resultados nestas notas.

12.20 Proposição. *Seja V uma vizinhança normal de p e x^1, \dots, x^n coordenadas normais em p . Nestas coordenadas vale*

$$g_{ij}(p) = \delta_{ij},$$

$$\frac{\partial g_{ij}}{\partial x^k}(p) = 0$$

e

$$\Gamma_{ij}^k(p) = 0$$

para todos os índices i, j, k .

Além disso, para cada $v \in T_p M$, a geodésica $\gamma_v(t) = \gamma(t, p, v)$ partindo de p com velocidade inicial v é representada em coordenadas normais por um segmento radial, isto é,

$$\gamma_v(t) = \exp_p(tv)$$

enquanto ela estiver dentro de V .

Prova: A condição

$$g_{ij}(p) = \delta_{ij}$$

é consequência imediata da definição de coordenadas normais, pois (usando a notação da Definição 12.19)

$$\begin{aligned} \left. \frac{\partial}{\partial x^i} \right|_q &= d\varphi_x(e_i) = d(\exp_p \circ E)_x(e_i) = d(\exp_p)_{Ex} \circ dE_x(e_i) = d(\exp_p)_{Ex} \circ E(e_i) \\ &= d(\exp_p)_{Ex} \circ E_i, \end{aligned}$$

de modo que

$$\left. \frac{\partial}{\partial x^i} \right|_p = d(\exp_p)_0 E_i = E_i,$$

logo

$$g_{ij}(p) = \left\langle \left. \frac{\partial}{\partial x^i} \right|_p, \left. \frac{\partial}{\partial x^j} \right|_p \right\rangle = \langle E_i, E_j \rangle = \delta_{ij}.$$

O resultado sobre a geodésica γ segue do fato de que por definição

$$\exp_p(tv) = \gamma(1, p, tv) = \gamma(t, p, v).$$

Isso implica por sua vez que

$$\Gamma_{ij}^k(p) = 0,$$

pois as geodésicas que passam por p são todas dadas por equações lineares. Mais especificamente, a equação geodésica para γ é

$$\frac{d^2 x^k}{dt^2} + \Gamma_{ij}^k \frac{dx^i}{dt} \frac{dx^j}{dt} = 0$$

e como a solução é

$$x^k(t) = tv^k,$$

a equação geodésica se reduz a

$$\Gamma_{ij}^k(\gamma_v(t)) v^i v^j = 0$$

Em $t = 0$, temos $\gamma_v(0) = p$, de modo que a equação

$$\Gamma_{ij}^k(p) v^i v^j = 0$$

vale para todo $v \in T_p M$. Isso só é possível se $\Gamma_{ij}^k(p) = 0$.

Para provar a condição sobre as derivadas das componentes da métrica lembramos do Corolário 11.25 a fórmula

$$\frac{\partial g_{ij}}{\partial x^k} = g_{ip}\Gamma_{jk}^p + g_{pj}\Gamma_{ik}^p.$$

Logo

$$\frac{\partial g_{ij}}{\partial x^k}(p) = g_{ip}(p)\Gamma_{jk}^p(p) + g_{pj}(p)\Gamma_{ik}^p(p) = 0.$$

■

Por causa deste resultado, as geodésicas dentro de uma vizinhança normal que partem de p são chamadas **raios geodésicos** ou **geodésicas radiais**. Observe que em coordenadas normais em p a derivada covariante em p é simplesmente

$$(\nabla_X Y)_p = X_p(Y^i) \left. \frac{\partial}{\partial x^i} \right|_p = X^j(p) \left. \frac{\partial Y^i}{\partial x^j}(p) \right. \left. \frac{\partial}{\partial x^i} \right|_p,$$

e a derivada covariante ao longo de uma curva $\alpha : I \rightarrow M$ com $\alpha(t_0) = p$ é dada em t_0 por

$$\left. \frac{DV}{dt} \right|_{t_0} = \left. \frac{dV^i}{dt}(t_0) \right. \left. \frac{\partial}{\partial x^i} \right|_p.$$

Como coordenadas normais em p simplificam bastante as expressões locais de vários objetos diferenciais e geométricos no ponto p , elas são muito usadas para obter resultados em geometria riemanniana que dependem apenas da situação em p , facilitando os cálculos necessários.

Vamos agora mostrar que uma isometria é definida pela diferencial em um único ponto:

12.21 Proposição. *Sejam M, N variedades riemannianas, com M conexa. Se $F, G : M \rightarrow N$ são isometrias tais que*

$$\begin{aligned} F(p) &= G(p), \\ dF_p &= dG_p, \end{aligned}$$

então $F = G$.

Prova: Seja

$$A = \{q \in M : F(q) = G(q)\}.$$

Então $A \neq \emptyset$, pois $p \in A$, e por continuidade A é fechado. Como M é conexa, para provar o resultado basta mostrar que A é aberto. Para isso, provaremos que se $q_0 \in A$ então qualquer vizinhança normal V de q_0 está contida em A . De fato, se $q \in V$, seja $v \in T_{q_0} M$ tal que

$$\gamma_v(1) = \exp_{q_0}(v) = q.$$

Da Proposição 12.5 segue que

$$F(q) = F(\gamma_v(1)) = \gamma_{dF_p(v)}(1) = \gamma_{dG_p(v)}(1) = G(\gamma_v(1)) = G(q).$$

■

12.5 Lema de Gauss

Para estudar as propriedades minimizantes das geodésicas, precisamos considerar o conceito de *variação* de uma curva, que tem sua origem no *cálculo de variações*.

12.22 Definição. Uma **curva diferenciável por partes** em M é uma aplicação contínua $\gamma : [a, b] \rightarrow M$ tal que existe uma partição do intervalo $[a, b]$

$$a = t_0 < t_1 < \dots < t_k < t_{k+1} = b$$

tal que as restrições $\gamma|_{[t_i, t_{i+1}]}$ são curvas diferenciáveis para $i = 0, \dots, k$.

Cada ponto $\gamma(t_i)$ é chamado um **vértice** da curva γ e o **ângulo** da curva no vértice $\gamma(t_i)$ é definido como sendo o ângulo entre os vetores

$$\lim_{t \rightarrow t_i^-} \gamma'(t) \quad \text{e} \quad \lim_{t \rightarrow t_i^+} \gamma'(t)$$

□

12.23 Definição. Seja $\gamma : [a, b] \rightarrow M$ uma curva diferenciável por partes. Uma **variação** de γ é uma aplicação

$$F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$$

tal que

- (i) $F(0, t) = \gamma(t)$;
- (ii) existe uma partição de $[a, b]$ por pontos

$$a = t_0 < t_1 < \dots < t_k < t_{k+1} = b$$

tal que $F|_{(-\varepsilon, \varepsilon) \times [t_i, t_{i+1}]}$ é diferenciável para $i = 0, \dots, k$.

Para cada s fixado, a curva $F_s(t) = F(s, t) : [a, b] \rightarrow M$ é chamada uma **curva principal da variação**. Para cada t fixado, a curva $F_t(s) = F(s, t) : (-\varepsilon, \varepsilon) \rightarrow M$ é chamada uma **curva transversal da variação**.

Um **campo vetorial ao longo de F** é uma aplicação $V : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow TM$ tal que $V(s, t) \in T_{F(s,t)}M$ para cada (s, t) e $V|_{(-\varepsilon, \varepsilon) \times [t_i, t_{i+1}]}$ é diferenciável para $i = 0, \dots, k$. Se F é uma variação de γ , então o **campo variacional de F** é o campo vetorial ao longo de γ

$$V(t) = \frac{\partial F}{\partial s}(0, t).$$

□

Dada uma variação $F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ de uma curva diferenciável por partes γ , consideraremos as derivadas covariantes de um campo vetorial qualquer, definido ao longo de F , ao longo das curvas principais e ao longo das curvas transversais da variação, onde estas derivadas estiverem definidas:

12.24 Notação. Denotaremos as **derivadas covariantes** de um campo vetorial V ao longo de uma variação F ao longo das curvas principais e transversais de F (onde estiverem definidas) respectivamente por

$$\frac{DV}{dt} \quad \text{e} \quad \frac{DV}{ds}.$$

□

12.25 Lema (Lema de Simetria). Seja $F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ uma variação. Então

$$\frac{D}{ds} \frac{\partial F}{\partial t} = \frac{D}{dt} \frac{\partial F}{\partial s}.$$

Prova: Escrevendo F localmente em coordenadas na forma

$$F(s, t) = (x^1(s, t), \dots, x^n(s, t)),$$

segue da regra da cadeia que

$$\begin{aligned}\frac{\partial F}{\partial t} &= \frac{\partial x^i}{\partial t} \partial_i, \\ \frac{\partial F}{\partial s} &= \frac{\partial x^j}{\partial s} \partial_i.\end{aligned}$$

Usando a comutatividade das derivadas parciais e a simetria da conexão, temos

$$\begin{aligned}\frac{D}{ds} \frac{\partial F}{\partial t} &= \frac{D}{ds} \left(\frac{\partial x^i}{\partial t} \partial_i \right) \\ &= \frac{\partial^2 x^i}{\partial t \partial s} \partial_i + \frac{\partial x^i}{\partial t} \frac{D}{ds} \partial_i \\ &= \frac{\partial^2 x^i}{\partial t \partial s} \partial_i + \frac{\partial x^i}{\partial t} \nabla_{\frac{\partial x^j}{\partial s}} \partial_j \partial_i \\ &= \frac{\partial^2 x^i}{\partial t \partial s} \partial_i + \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial s} \nabla_{\partial_i} \partial_j \\ &= \frac{\partial^2 x^i}{\partial s \partial t} \partial_i + \frac{\partial x^j}{\partial t} \frac{\partial x^i}{\partial s} \nabla_{\partial_j} \partial_i \\ &= \frac{D}{dt} \left(\frac{\partial x^i}{\partial s} \partial_i \right) \\ &= \frac{D}{dt} \frac{\partial F}{\partial s}.\end{aligned}$$

■

No que se segue, identificaremos o espaço tangente a $T_p M$ no vetor v com o próprio $T_p M$, isto é, $T_v(T_p M) = T_p M$. A fronteira $\partial B_\varepsilon(p)$ de uma bola geodésica é chamada uma **esfera geodésica** e denotada $S_\varepsilon(p)$.

12.26 Lema (Lema de Gauss). *Seja M uma variedade métrica, $p \in M$ e $v \in T_p M$ tal que $\exp_p(v)$ está definido. Então*

$$\langle d(\exp_p)_v v, d(\exp_p)_v w \rangle = \langle v, w \rangle$$

para todo $w \in T_p M$.

Em particular, as geodésicas radiais que partem de p são ortogonais às esferas geodésicas centradas em p .

Prova: Decompondo $T_p M = \langle v \rangle + \langle v \rangle^\perp$, se $w \in \langle v \rangle$ então $w = av$ para algum escalar $a \neq 0$. Temos

$$\begin{aligned}d(\exp_p)_v v &= \frac{\partial \exp_p}{\partial v}(v) = \frac{d}{dt} \exp_p((t+1)v) \Big|_{t=0} = \frac{d}{dt} \gamma(1, p, (t+1)v) \Big|_{t=0} \\ &= \frac{d}{dt} \gamma(t+1, p, v) \Big|_{t=0} = \gamma'_{(p,v)}(1),\end{aligned}$$

e como a norma do vetor tangente a uma geodésica é constante, segue que

$$\begin{aligned}\langle d(\exp_p)_v v, d(\exp_p)_v w \rangle &= a \langle d(\exp_p)_v v, d(\exp_p)_v v \rangle \\ &= a \left\| \gamma'_{(p,v)}(1) \right\|^2 \\ &= a \left\| \gamma'_{(p,v)}(0) \right\|^2 \\ &= a \langle v, v \rangle \\ &= \langle v, w \rangle.\end{aligned}$$

Se $w \in \langle v \rangle^\perp$, então $\langle v, w \rangle = 0$ e para terminar a demonstração do lema temos que provar que

$$\left\langle d(\exp_p)_v v, d(\exp_p)_v w \right\rangle = 0.$$

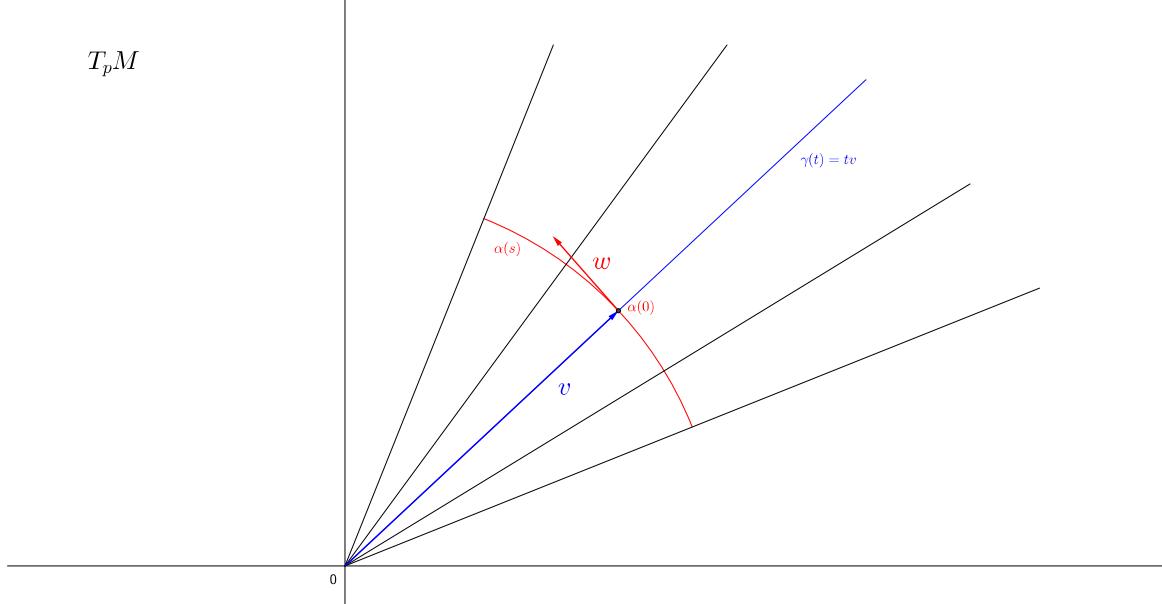
Considere o raio geodésico $\gamma : [0, 1] \rightarrow M$

$$\gamma(t) = \exp_p(tv).$$

Para algum $\varepsilon > 0$, considere a curva $\alpha : (-\varepsilon, \varepsilon) \rightarrow T_p M$ com $\alpha(0) = v$, $\alpha'(0) = w$ e $\|\alpha(s)\| \equiv \text{constante} = \|w\|$, ou seja, α é um pequeno arco de círculo centrado na origem de $T_p M$ e raio $\|v\|$, começando em v e com velocidade inicial w . Definimos uma variação $F : (-\varepsilon, \varepsilon) \times [0, 1] \rightarrow M$ de γ por

$$F(s, t) = \exp_p(t\alpha(s)).$$

Em particular, as curvas principais da variação $F_s(t)$ são geodésicas, mais especificamente as geodésicas radiais $\gamma(1, p, t\alpha(s)) = \gamma(t, p, \alpha(s))$ que passam por p com velocidade $\alpha(s)$. Uma variação F cujas curvas principais são geodésicas é chamada uma **variação geodésica**.



Como

$$\begin{aligned} \frac{\partial F}{\partial t}(s, t) &= d(\exp_p)_{t\alpha(s)}(\alpha(s)), \\ \frac{\partial F}{\partial s}(s, t) &= d(\exp_p)_{t\alpha(s)}(t\alpha'(s)), \end{aligned}$$

segue que

$$\begin{aligned} \frac{\partial F}{\partial t}(0, 1) &= d(\exp_p)_v v, \\ \frac{\partial F}{\partial s}(0, 1) &= d(\exp_p)_v w, \end{aligned}$$

de modo que

$$\left\langle d(\exp_p)_v v, d(\exp_p)_v w \right\rangle_p = \left\langle \frac{\partial F}{\partial t}(0,1), \frac{\partial F}{\partial s}(0,1) \right\rangle_p.$$

Mostraremos que

$$\left\langle \frac{\partial F}{\partial t}(0,1), \frac{\partial F}{\partial s}(0,1) \right\rangle_p = 0$$

analizando o produto interno

$$\left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle(s,t).$$

Afirmamos que ele independe de t . De fato, pela compatibilidade da conexão riemanniana com a métrica, temos para todo (s,t) ,

$$\begin{aligned} \frac{\partial}{\partial t} \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle &= \left\langle \frac{D}{dt} \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle + \left\langle \frac{\partial F}{\partial t}, \frac{D}{dt} \frac{\partial F}{\partial s} \right\rangle \\ &= \left\langle \frac{\partial F}{\partial t}, \frac{D}{dt} \frac{\partial F}{\partial s} \right\rangle, \end{aligned}$$

pois

$$\frac{D}{dt} \frac{\partial F}{\partial t} = 0$$

já que $F_s(t)$ são geodésicas e $\frac{\partial F}{\partial t}$ é seu campo velocidade. Pelo Lema de Simetria,

$$\left\langle \frac{\partial F}{\partial t}, \frac{D}{dt} \frac{\partial F}{\partial s} \right\rangle = \left\langle \frac{\partial F}{\partial t}, \frac{D}{ds} \frac{\partial F}{\partial t} \right\rangle = \frac{1}{2} \frac{\partial}{\partial s} \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial t} \right\rangle.$$

Mas

$$\left\| \frac{\partial F}{\partial t}(s,t) \right\| = \left\| \frac{\partial F}{\partial t}(s,0) \right\| = \left\| d(\exp_p)_0(\alpha(s)) \right\| = \|\alpha(s)\| \equiv \text{constante},$$

e assim

$$\frac{\partial}{\partial t} \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle = 0.$$

Em particular, como

$$\frac{\partial F}{\partial s}(0,0) = 0,$$

concluímos que

$$\left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle(0,1) = \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right\rangle(0,0) = 0.$$

■

12.6 Geodésicas em Variedades Riemannianas

12.6.1 Geodésicas em Variedades Riemannianas minimizam distâncias localmente

Veremos agora que as geodésicas minimizam distâncias, localmente. A distância entre dois pontos de uma variedade riemanniana é definida como sendo o ínfimo dos comprimentos das curvas que ligam estes pontos, como veremos em detalhes na Seção 12.8.

12.27 Definição. Se $\gamma : I \rightarrow M$ é uma geodésica e $[a, b] \subset I$, a restrição $\gamma|_{[a,b]}$ é chamada o **segmento de geodésica** ligando $\gamma(a)$ a $\gamma(b)$.

Denote o comprimento de uma curva α em M ligando os pontos p e q por $\ell(\alpha)$. Dizemos que o segmento de geodésica $\gamma : [a, b] \rightarrow M$ é **minimizante** se $\ell(\gamma) \leq \ell(\alpha)$ para toda curva α ligando $\gamma(a)$ a $\gamma(b)$. \square

12.28 Proposição (Geodésicas minimizam distâncias localmente). Sejam M uma variedade riemanniana, $p \in M$ e $B(p)$ uma bola normal centrada em p . Seja $\gamma : [0, 1] \rightarrow B(p)$ um segmento de geodésica com $\gamma(0) = p$ e denote $q = \gamma(1)$.

Se $\alpha : [0, 1] \rightarrow M$ é qualquer curva diferenciável por partes ligando p a q , então

$$\ell(\gamma) \leq \ell(\alpha).$$

Se $\ell(\alpha) = \ell(\gamma)$, então necessariamente $\alpha([0, 1]) = \gamma([0, 1])$.

Prova: Suponha em primeiro lugar que $\alpha([0, 1]) \subset B(p)$. Como \exp_p é um difeomorfismo em $B(p)$, podemos escrever $\alpha(t)$ de forma única em “coordenadas polares” na forma

$$\alpha(t) = \exp_p(r(t)c(t)) \quad (12.6)$$

para $t \in (0, 1]$, onde c é uma curva diferenciável por partes em $T_p M$ com $\|c(t)\| = 1$ para todo t e $r : (0, 1] \rightarrow \mathbb{R}$ é uma função diferenciável por partes; de fato, basta considerar $\beta(t) = \exp_p^{-1}(\alpha(t))$ e definir

$$c(t) = \frac{\beta(t)}{\|\beta(t)\|} \quad \text{e} \quad r(t) = \|\beta(t)\|.$$

Se α retorna ao ponto p , de modo que $\beta(t_1) = 0$ para algum $t_1 > 0$, redefinimos c eliminando todo o segmento $\alpha([0, t_1])$, o que não prejudicará o resultado, pois mostraremos que a curva resultante ainda assim possui comprimento maior que a geodésica. Denote

$$F(r, t) = \exp_p(rc(t)), \quad (12.7)$$

de modo que, exceto possivelmente para um número finito de pontos,

$$\alpha'(t) = \frac{\partial F}{\partial r} r'(t) + \frac{\partial F}{\partial t}.$$

Como

$$\begin{aligned} \frac{\partial F}{\partial r} &= d(\exp_p)_{\beta(t)} c(t), \\ \frac{\partial F}{\partial t} &= d(\exp_p)_{\beta(t)} (rc'(t)), \end{aligned}$$

segue do lema de Gauss que

$$\left\langle \frac{\partial F}{\partial r}, \frac{\partial F}{\partial t} \right\rangle = \langle c(t), rc'(t) \rangle = r \langle c(t), c'(t) \rangle = 0, \quad (12.8)$$

e

$$\left\| \frac{\partial F}{\partial r} \right\|^2 = \left\langle \frac{\partial F}{\partial r}, \frac{\partial F}{\partial r} \right\rangle = \langle c(t), c(t) \rangle = \|c(t)\|^2 = 1. \quad (12.9)$$

Logo,

$$\|\alpha'(t)\|^2 = \left\| \frac{\partial F}{\partial r} \right\|^2 |r'(t)|^2 + \left\| \frac{\partial F}{\partial t} \right\|^2 = |r'(t)|^2 + \left\| \frac{\partial F}{\partial t} \right\|^2 \geq |r'(t)|^2.$$

Daí,

$$\ell(\alpha) \geq \int_{\varepsilon}^1 \|\alpha'(t)\| dt \geq \int_{\varepsilon}^1 |r'(t)| dt \geq \int_{\varepsilon}^1 r'(t) dt = r(1) - r(\varepsilon) = \ell(\gamma) - r(\varepsilon).$$

Fazendo $\varepsilon \rightarrow 0$, como $r(\varepsilon) \rightarrow 0$, segue o primeiro resultado.

Para que tenhamos $\ell(\gamma) = \ell(\alpha)$ é necessário que

$$\frac{\partial F}{\partial t} = 0,$$

ou seja,

$$c'(t) = 0$$

e portanto $c \equiv$ constante. Como $r'(t) > 0$, α é então uma reparametrização monótona de γ , donde $\alpha([0, 1]) = \gamma([0, 1])$.

Se $\alpha([0, 1]) \not\subset B(p) = B_\rho(p)$, consideramos o primeiro instante $t_1 \in (0, 1]$ tal que $\alpha(t_1) \in \partial B(p)$. Então,

$$\ell(\alpha) \geq \ell(\alpha|_{[0, t_1]}) \geq \rho > \ell(\gamma).$$

■

O resultado vale apenas localmente: um segmento de geodésica suficientemente grande pode não ser minimizante. Por exemplo, as geodésicas de uma esfera (círculos maximais) que partem de um ponto p deixam de ser minimizantes depois que passam pelo antípoda de p .

Vale a recíproca da Proposição 12.28 e ela é até mais forte. Antes uma definição e um resultado preliminar.

12.29 Definição. Seja M uma variedade riemanniana.

Se V é uma vizinhança em M que é uma vizinhança normal para todos os seus pontos, dizemos que V é uma **vizinhança totalmente normal**.

Dizemos que uma vizinhança V em M é uma **vizinhança uniformemente normal** se existe $\delta > 0$ tal que para todo $q \in V$ a bola geodésica $B_\delta(q) = \exp_q(B_\delta(0))$ contém V . □

Observe que em uma vizinhança uniformemente normal, dois pontos quaisquer da vizinhança são ligados por uma geodésica radial. Toda vizinhança uniformemente normal é em particular totalmente normal.

12.30 Lema. *Seja M uma variedade riemanniana. Para cada $p \in M$ existe uma vizinhança W de p que é uma vizinhança uniformemente normal.*

Prova: Sejam $\varepsilon > 0$, V e \mathcal{V} como na Proposição 12.13. Defina $F : \mathcal{V} \rightarrow M \times M$ por

$$F(q, v) = (q, \exp_q v).$$

Temos $F(p, 0) = (p, p)$ e

$$dF(p, 0) = \begin{bmatrix} I & * \\ 0 & I \end{bmatrix}$$

pois $d(\exp_p)_0 = I$. Portanto, F é um difeomorfismo local em uma vizinhança de $(p, 0)$, logo existe uma vizinhança $\mathcal{V}' \subset \mathcal{V}$ de $(p, 0)$ em TM tal que F aplica \mathcal{V}' difeomorficamente sobre uma vizinhança W' de $(p, 0)$ em $M \times M$. Para algum $\delta > 0$, escolha \mathcal{V}' da forma

$$\mathcal{V}' = \{(q, v) : q \in V' \text{ e } v \in B_\delta(0) \subset T_q M\}$$

onde V' é uma vizinhança de p em M . Escolha a vizinhança W de p em M de tal forma que $W \times W \subset W'$. Então, para todo $q \in W$ e $B_\delta(0) \subset T_q M$ temos

$$F(\{q\} \times B_\delta(0)) \supset \{q\} \times W$$

o que implica, pela definição de F ,

$$B_\delta(q) = \exp_q(B_\delta(0)) \supset W.$$

■

12.31 Corolário (Curvas que minimizam distâncias são geodésicas). Seja M uma variedade riemanniana. Se $\alpha : [a, b] \rightarrow M$ é uma curva diferenciável por partes com parâmetro proporcional ao comprimento de arco que tem comprimento menor ou igual a qualquer outra curva diferenciável por partes ligando $\alpha(a)$ a $\alpha(b)$, então α é uma geodésica.

Em particular, α é diferenciável em todo o intervalo $[a, b]$.

Prova: Sejam $t_0 \in [a, b]$ e V uma vizinhança uniformemente normal de $\alpha(t_0)$. Seja $I = [t_0 - \delta, t_0 + \delta] \subset [a, b]$ um intervalo fechado tal que $\alpha(I) \subset V$. Então a restrição $\alpha|_I$ é uma curva diferenciável por partes ligando dois pontos de uma bola normal. Pela Proposição 12.28, $\ell(\alpha|_I)$ é igual ao comprimento da geodésica radial ligando $\alpha(t_0 - \delta)$ a $\alpha(t_0 + \delta)$ (caso contrário obteríamos outra curva diferenciável por partes com comprimento menor que α ligando $\alpha(a)$ a $\alpha(b)$, contrariando a hipótese). Segue da mesma proposição que a imagem de $\alpha|_I$ é a imagem desta geodésica. Como $\alpha|_I$ está parametrizada por um parâmetro proporcional ao comprimento de arco, segue que $\alpha|_I$ é ela própria uma geodésica. ■

Observe que este resultado é global, não apenas local: dados dois pontos quaisquer de uma variedade riemanniana, se existir uma curva ligando eles que minimiza distâncias, então ela é uma geodésica. Por outro lado, dados dois pontos arbitrários em uma variedade riemanniana pode não existir nenhuma geodésica que os ligue: considere \mathbb{R}^n com a bola centrada em zero e raio 1 removida; para $\delta > 0$ suficientemente pequeno, não existe nenhuma geodésica que une os pontos

$$(-1 - \delta, 0, \dots, 0) \quad \text{e} \quad (1 + \delta, 0, \dots, 0),$$

por exemplo, e esta afirmação vale para quaisquer dois pontos antípodas muito próximos ao bordo da bola removida.

12.32 Exemplo (Geodésicas de H^2 usando isometrias). Como no Exemplo 12.8, para determinar as geodésicas do plano hiperbólico usaremos isometrias.

Primeiramente, usando o Corolário 12.31 mostraremos que o semieixo y superior é uma geodésica. De fato, seja $\gamma(t) = (0, t)$, $t > 0$, uma parametrização por comprimento de arco deste semieixo e considere um segmento $\gamma|_{[a,b]}$. Temos

$$\ell(\gamma) = \int_a^b \|\gamma'(t)\|_{\gamma(t)} dt = \int_a^b \frac{1}{t} dt = \log b - \log a.$$

Se $c : [a, b] \rightarrow \mathbb{H}^2$, $c(t) = (x(t), y(t))$, é qualquer curva diferenciável por partes com $c(a) = \gamma(a) = (0, a)$ e $c(b) = \gamma(b) = (0, b)$, então

$$\ell(c) = \int_a^b \|c'(t)\|_{c(t)} dt = \int_a^b \frac{\sqrt{[x'(t)]^2 + [y'(t)]^2}}{y(t)} dt \geq \int_a^b \frac{\sqrt{[y'(t)]^2}}{y(t)} dt = \int_a^b \frac{y'(t)}{y(t)} dt = \ell(\gamma).$$

Observe que, diferentemente da esfera, variedade compacta, esta geodésica de \mathbb{H}^2 tem comprimento infinito, como as retas do plano \mathbb{R}^2 .

As isometrias do plano hiperbólico são dadas pelas transformações de Möbius

$$T(z) = \frac{az + b}{cz + d}, \quad ad - bc = 1,$$

onde $z = x + iy \in \mathbb{C}$. Estas isometrias transformam o semieixo y em semicírculos superiores ou em semi-retas superiores $x = x_0$, $y > 0$. Estas são as geodésicas de \mathbb{H}^2 , pois por todo ponto $p \in \mathbb{H}^2$ passa um tal semicírculo ou uma tal semi-reta, em todas as direções possíveis. A demonstração de todos estes detalhes é deixada ao leitor (Exercício 12.54). □

12.6.2 Vizinhanças Convexas

Em uma vizinhança uniformemente normal, como a considerada no Lema 12.30, dados dois pontos $q_1, q_2 \in W$ existe uma única geodésica minimizante de comprimento menor que 2δ ligando q_1 a q_2 . Por outro lado, nada

garante que esta geodésica esteja contida em W (as duas geodésicas radiais, uma partindo de q_1 e outra partindo de q_2 podem se ligar através de um ponto p fora de $W \subset B_\delta(q_1) \cup B_\delta(q_2)$).

12.33 Definição. Seja M uma variedade riemanniana. Dizemos que uma vizinhança V em M é uma **vizinhança convexa** se para todos os pontos $p, q \in V$ existe uma única geodésica minimizante ligando p a q cuja imagem está inteiramente contida em V . \square

O raio de uma bola geodésica pode ser escolhido de tal forma que ela se torne convexa.

12.34 Lema. *Seja M uma variedade riemanniana. Para todo $p \in M$ existe $R > 0$ tal que qualquer geodésica de M que é tangente à esfera geodésica $S_r(p)$ de raio $r < R$ em $q \in M$ está fora da bola geodésica $B_r(p)$ numa vizinhança de q .*

Prova: Seja V uma vizinhança totalmente normal de p . Restringindo convenientemente o intervalo de definição, podemos assumir que todas as geodésicas de V tem velocidade 1. Podemos portanto nos restringir ao fibrado tangente unitário

$$T_1V = \{(q, v) \in TM : q \in V, v \in T_qM \text{ e } \|v\| = 1\}.$$

Seja $\gamma : (-\varepsilon, \varepsilon) \times T_1V \rightarrow M$ a aplicação diferenciável tal que $\gamma(t, q, v)$ é a geodésica que no instante $t = 0$ passa por q com velocidade v , $\|v\| = 1$. Definimos

$$u(t, q, v) = \exp_p^{-1}(\gamma(t, q, v))$$

e $F : (-\varepsilon, \varepsilon) \times T_1V \rightarrow \mathbb{R}$ por

$$F(t, q, v) = \|u(t, q, v)\|^2,$$

ou seja, F mede o quadrado da distância (medida em T_pM , não na variedade M) à origem de um ponto que se desloca sobre a geodésica γ . Temos

$$\begin{aligned} \frac{\partial F}{\partial t} &= 2 \left\langle u, \frac{\partial u}{\partial t} \right\rangle, \\ \frac{\partial^2 F}{\partial t^2} &= 2 \left\langle u, \frac{\partial^2 u}{\partial t^2} \right\rangle + 2 \left\| \frac{\partial u}{\partial t} \right\|^2. \end{aligned}$$

Seja agora $r > 0$ tal que $B_r(p) = \exp_p B_r(0) \subset V$ e γ uma geodésica tangente à esfera geodésica no ponto $q = \gamma(0, q, v)$. Pelo lema de Gauss,

$$\left\langle u(0, q, v), \frac{\partial u}{\partial t}(0, q, v) \right\rangle = 0$$

já que $\gamma(t, q, v)$ é uma geodésica radial que parte de q . Em particular,

$$\frac{\partial F}{\partial t}(0, q, v) = 0.$$

Se mostrarmos que para r suficientemente pequeno o ponto crítico $(0, q, v)$ de F é um mínimo estrito, de modo que os pontos de γ estão a uma distância de p maior que r , a demonstração estará concluída.

Para provar isso, basta observar que para $q = p$ temos $u(t, p, v) = tv$ e portanto

$$\frac{\partial^2 F}{\partial t^2}(0, p, v) = 2\|v\|^2 = 2,$$

logo existe uma vizinhança $W \subset V$ de p tal que

$$\frac{\partial^2 F}{\partial t^2}(0, q, v) > 0$$

para todo $q \in W$. Escolha $R > 0$ tal que $B_r(p) = \exp_p B_R(0) \subset W$. Então, se $r < R$, F tem um mínimo estrito em $(0, q, v)$. ■

12.35 Proposição. Seja M uma variedade riemanniana. Para todo $p \in M$ existe $\varepsilon > 0$ tal que a bola geodésica $B_\varepsilon(p)$ é convexa.

Prova: Seja $R > 0$ como no lema. Escolha W e $\delta < 0$ como no Lema 12.30 com $\delta < R/2$. Tome $\varepsilon < \delta$ tal que $B_\varepsilon(p) \subset W$. Afirmamos que $B_\varepsilon(p)$ é convexa.

De fato, sejam $q_1, q_2 \in \overline{B}_\varepsilon(p)$ e γ a única geodésica minimizante de comprimento menor que $2\varepsilon < 2\delta < R$ ligando q_1 a q_2 ; em particular, γ está contida na bola $B_R(p)$. Suponha por absurdo que o interior de γ não está contido em $B_\varepsilon(p)$. Então existe um ponto q no interior de γ onde a distância máxima $r < R$ de p a γ é atingida. Logo γ é tangente à esfera geodésica $S_r(p)$ em q , mas os pontos de γ numa vizinhança de q estarão no fecho de $B_r(p)$, contradizendo o lema. ■

12.7 Geodésicas em Variedades de Lorentz

12.36 Definição. Seja M uma variedade de Lorentz.

Uma curva diferenciável $\alpha : I \rightarrow M$ é do tipo **tempo** se $\alpha'(t)$ é do tipo tempo em $T_{\alpha(t)}M$ para todo $t \in I$. Ela é do tipo **luz** se $\alpha'(t)$ é do tipo luz em $T_{\alpha(t)}M$ para todo $t \in I$. Ela é do tipo **espaço** se $\alpha'(t)$ é do tipo espaço em $T_{\alpha(t)}M$ para todo $t \in I$.

Uma curva diferenciável por partes $\alpha : I \rightarrow M$ é do tipo **tempo** se nos pontos t onde ela for diferenciável $\alpha'(t)$ é do tipo tempo e se nos pontos t_i onde ela não for diferenciável valer

$$\langle \alpha'(t_i^-), \alpha'(t_i^+) \rangle < 0.$$

Ela é do tipo **luz** se nos pontos t onde ela for diferenciável $\alpha'(t)$ é do tipo luz e se nos pontos t_i onde ela não for diferenciável valer

$$\langle \alpha'(t_i^-), \alpha'(t_i^+) \rangle = 0.$$

Ela é do tipo **espaço** se nos pontos t onde ela for diferenciável $\alpha'(t)$ é do tipo espaço e se nos pontos t_i onde ela não for diferenciável valer

$$\langle \alpha'(t_i^-), \alpha'(t_i^+) \rangle > 0.$$

□

Em geral, uma curva arbitrária em uma variedade de Lorentz não tem nenhuma destas *características causais*. Por outro lado, a trajetória de uma partícula material no espaço-tempo de Minkowski é sempre uma curva do tipo tempo, como vimos no Capítulo 3 (em Relatividade Especial, a velocidade da luz não pode ser ultrapassada). Para uma partícula material que segue uma trajetória $\alpha : I \rightarrow \mathbb{M}$ do tipo tempo no espaço-tempo de Minkowski, com velocidade unitária, isto é,

$$\langle \alpha'(\tau), \alpha'(\tau) \rangle = -1$$

para todo $\tau \in I$, então o parâmetro τ é o precisamente o tempo próprio da partícula. Daí segue que

$$\ell(\alpha) = \int_0^{\tau_0} \|\alpha'(\tau)\| d\tau = \tau_0,$$

ou seja, o seu comprimento é o seu tempo próprio. O mesmo vale para curvas em variedades de Lorentz, isto é, curvas no espaço-tempo curvo.

Da Proposição 12.2 segue no entanto que *geodésicas* sempre são ou do tipo luz, ou do tipo espaço ou do tipo tempo.

12.37 Lema. Sejam M uma variedade de Lorentz e $p \in M$. Seja $\alpha : [0, T] \rightarrow M$ uma curva diferenciável por partes do tipo tempo com $\alpha(0) = p$, tal que

$$\beta = \exp_p^{-1} \circ \alpha : I \rightarrow T_p M$$

está definida. Então β é uma curva do tipo tempo que está contida em um único cone temporal de $T_p M$.

Em particular, β não tem autointerseções.

Prova: Provaremos o resultado apenas no caso em que β é diferenciável, o caso mais geral ficando como exercício (ou veja [ONeill], Lema 5.33, p. 146). Note que $\beta(0) = 0$.

O Lema de Gauss garante que β é uma curva do tipo tempo em $T_p M$: de fato, como

$$\alpha'(t) = (\exp_p \circ \beta)'(t) = d(\exp_p)_{\beta(t)} \beta'(t),$$

segue do Lema de Gauss que

$$\langle \alpha'(t), \alpha'(t) \rangle = \left\langle d(\exp_p)_{\beta(t)} \beta'(t), d(\exp_p)_{\beta(t)} \beta'(t) \right\rangle = \langle \beta'(t), \beta'(t) \rangle,$$

e portanto β é uma curva do tipo tempo em $T_p M$ se e somente se α é uma curva do tipo tempo em M .

Seja \mathcal{C} o cone temporal de $T_p M$ tal que $\beta'(0) \in \mathcal{C}$. Como a função

$$f(t) = \langle \beta'(0), \beta'(t) \rangle$$

é contínua e $f(0) < 0$, segue que $f(t) < 0$ para todo $t \in [0, \varepsilon)$ se $\varepsilon > 0$ é suficientemente pequeno, de modo que $\beta'([0, \varepsilon)) \subset \mathcal{C}$ pela Proposição 2.25. Além disso, pela fórmula de Taylor,

$$\beta(t) = \beta'(0)t + O(t^2),$$

de modo que

$$\langle \beta(t), \beta(t) \rangle = \langle \beta'(0), \beta'(0) \rangle t^2 + O(t^3),$$

onde $\beta(t)$ é um vetor do tipo tempo para todo $t \in [0, \varepsilon)$ (diminuindo ε , se necessário); analogamente, temos

$$\langle \beta(t), \beta'(0) \rangle = \langle \beta'(0), \beta'(0) \rangle t + O(t^2),$$

de modo que

$$\langle \beta(t), \beta'(0) \rangle < 0$$

para todo $t \in [0, \varepsilon)$ (diminuindo ε , se necessário) e portanto $\beta([0, \varepsilon)) \subset \mathcal{C}$. Portanto, $\beta(t)$ e $\beta'(t)$ estão no mesmo cone temporal para todo $t \in [0, \varepsilon)$, isto é,

$$\langle \beta(t), \beta'(t) \rangle < 0 \tag{12.10}$$

para todo $t \in [0, \varepsilon)$, pela Proposição 2.25. Logo, a função

$$g(t) = \langle \beta(t), \beta(t) \rangle$$

é negativa e estritamente decrescente para $t \in (0, \varepsilon)$, pois $g'(t) < 0$. Mas o mesmo argumento acima permite concluir que enquanto a curva β permanecer no cone \mathcal{C} , $\beta(t)$ será do tipo tempo, $\beta(t)$ e $\beta'(t)$ estarão no mesmo cone temporal e portanto (12.10) valerá. Segue que nunca teremos $g(t) = 0$, isto é, a curva β nunca cruzará a fronteira do cone temporal \mathcal{C} . ■

Observe que o resultado acima não impede uma curva do tipo tempo em uma variedade de Lorentz de possuir autointerseções, pois se a imagem de uma tal curva não está inteiramente contida na imagem da aplicação exponencial \exp_p para algum ponto p , ela pode ter autointerseções fora de $\exp_p(T_p M)$. Fisicamente, uma partícula que percorresse tal trajetória estaria viajando no tempo para o passado. Existem variedades, mesmo riemannianas que não são cobertas por $\exp_p(T_p M)$ para nenhum $p \in M$; tais variedades são chamadas variedades não completas (veja a Seção 12.9 para mais detalhes). E existem exemplos de modelos de espaço-tempo para o universo que admitem curvas do tipo tempo com autointerseções (os primeiros tais modelos foram obtidos por Gödel).

12.38 Proposição (Geodésicas do tipo tempo em Variedades de Lorentz maximizam distâncias localmente). *Sejam M uma variedade de Lorentz, $p \in M$, V uma vizinhança normal de p e $q \in V$. Se existir uma curva do tipo tempo em V ligando p a q , então existe uma geodésica radial do tipo tempo em V ligando p a q e além disso ela é a única curva do tipo tempo com comprimento maximal ligando p a q .*

Prova: Seja α uma curva do tipo tempo ligando p a q . Pelo lema, $\beta = \exp_p^{-1} \circ \alpha$ está contida em um único cone temporal de $T_p M$. Portanto, o segmento radial $L(t) = tx_0$, $t \in [0, 1]$, que liga 0 a $v = \exp_p^{-1} q$ está neste mesmo cone temporal. Segue que a geodésica radial $\gamma(t) = \exp_p(L(t))$ que liga p a q é do tipo tempo pelo Lema de Gauss.

O argumento para provar o restante do teorema é uma modificação da demonstração da Proposição 12.28, a seguir.

Como na demonstração daquela proposição, escrevemos $\alpha(t)$ de forma única em coordenadas polares na forma

$$\alpha(t) = \exp_p(r(t)c(t))$$

para $t \in (0, 1]$, onde c é uma curva diferenciável por partes em $T_p M$ com $\langle c(t), c(t) \rangle = -1$ para todo t , já que β é do tipo tempo, e $r : (0, 1] \rightarrow \mathbb{R}$ é uma função diferenciável por partes, e definimos

$$F(r, t) = \exp_p(rc(t)).$$

Exceto possivelmente para um número finito de pontos vale

$$\alpha'(t) = \frac{\partial F}{\partial r} r'(t) + \frac{\partial F}{\partial t}.$$

Como

$$\begin{aligned} \frac{\partial F}{\partial r} &= d(\exp_p)_{\beta(t)} c(t), \\ \frac{\partial F}{\partial t} &= d(\exp_p)_{\beta(t)}(rc'(t)), \end{aligned}$$

segue do lema de Gauss que

$$\left\langle \frac{\partial F}{\partial r}, \frac{\partial F}{\partial t} \right\rangle = \langle c(t), rc'(t) \rangle = r \langle c(t), c'(t) \rangle = 0.$$

Em particular, como $\frac{\partial F}{\partial r}$ é do tipo tempo, $\frac{\partial F}{\partial t}$ é do tipo espaço, isto é,

$$\left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial t} \right\rangle \geq 0.$$

Também pelo lema de Gauss temos

$$\left\langle \frac{\partial F}{\partial r}, \frac{\partial F}{\partial r} \right\rangle = \left\langle d(\exp_p)_{\beta(t)} c(t), d(\exp_p)_{\beta(t)} c(t) \right\rangle = \langle c(t), c(t) \rangle = -1.$$

Logo,

$$\begin{aligned} \langle \alpha'(t), \alpha'(t) \rangle &= \left\langle \frac{\partial F}{\partial r}, \frac{\partial F}{\partial r} \right\rangle |r'(t)|^2 + \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial t} \right\rangle \\ &= -|r'(t)|^2 + \left\langle \frac{\partial F}{\partial t}, \frac{\partial F}{\partial t} \right\rangle \\ &\geq -|r'(t)|^2. \end{aligned}$$

Como

$$-|r'(t)|^2 \leq \langle \alpha'(t), \alpha'(t) \rangle < 0,$$

segue que

$$\|\alpha'(t)\| = |\langle \alpha'(t), \alpha'(t) \rangle| \leq |r'(t)|.$$

Pelo lema, α está contida em um único cone temporal, logo

$$r'(t) > 0;$$

de fato, se a curva voltasse sobre si mesma, seu vetor velocidade seria do tipo luz em algum momento. Daí,

$$\ell(\alpha) = \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} \|\alpha'(t)\| dt \leq \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} |r'(t)| dt = \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} r'(t) dt = \ell(\gamma).$$

■ Como em variedades riemannianas, todo ponto de uma variedade de Lorentz possui uma vizinhança convexa (de fato, este resultado é verdadeiro para qualquer variedade métrica). A demonstração deste resultado é exatamente igual à do Lema 12.34 e Proposição 12.35. Note que na demonstração daquele lema, a função distância pode ser tomada como a função distância euclideana nas coordenadas normais de $T_p M$, não a função distância da variedade riemanniana (isto é, não consideramos a métrica de Lorentz de $T_p M$; se desejar, veja a demonstração detalhada em [ONeill], Cap. 5, Proposição 7, p. 103). Enunciamos o resultado explicitamente para futura referência:

12.39 Proposição. *Seja M uma variedade de Lorentz. Para todo $p \in M$ existe $\varepsilon > 0$ tal que a bola geodésica $B_\varepsilon(p)$ é convexa.*

12.8 Função Distância em Variedades Riemannianas

12.40 Definição. Seja M uma variedade riemanniana conexa. Dados $p, q \in M$, a **distância** entre p e q é definida por

$$\text{dist}(p, q) = \inf \{ \ell(c) : c \text{ é uma curva diferenciável por partes ligando } p \text{ e } q \}.$$

□

12.41 Proposição. *Se existe uma geodésica minimizante ligando p e q , então $\text{dist}(p, q) = \ell(\gamma)$.*

Prova: Segue do Corolário 12.31. ■

12.42 Proposição. *Seja M uma variedade riemanniana conexa. Com a função distância definida acima, M é um espaço métrico.*

Além disso, a topologia de M como espaço métrico coincide com a topologia inicial de M como variedade diferenciável.

Prova: De fato, a função distância satisfaz as três propriedades da função distância de um espaço métrico:

(i) Simetria:

$$\text{dist}(p, q) = \text{dist}(q, p).$$

(ii) Desigualdade triangular:

$$\text{dist}(p, q) \leq \text{dist}(p, r) + \text{dist}(r, q).$$

(iii) Positividade:

$$\text{dist}(p, q) \geq 0$$

e

$$\text{dist}(p, q) = 0 \quad \text{se e somente se } p = q.$$

Todas as propriedades seguem imediatamente da definição (a desigualdade triangular segue da definição de ínfimo) exceto a afirmação que $\text{dist}(p, q) = 0$ implica $p = q$. Suponha o contrário e considere uma bola normal $B_\varepsilon(p)$ que não contém q . Como $\text{dist}(p, q) = 0$, existe uma curva c ligando p a q com comprimento menor que ε , contradizendo a Proposição 12.28.

Pela Proposição 12.41, se existir uma geodésica minimizante γ ligando p e q então $\text{dist}(p, q) = \ell(\gamma)$. Em particular, dado $p \in M$, se $\varepsilon > 0$ é suficientemente pequeno, a bola geodésica $B_\varepsilon(p)$ de raio ε coincide com a bola métrica de centro em p e raio ε definida pela função distância:

$$B(p; \varepsilon) = \{q \in M : \text{dist}(p, q) < \varepsilon\}.$$

Logo, bolas geodésicas contêm bolas métricas e vice-versa, portanto as topologias são as mesmas. ■
De agora em diante, consideraremos variedades riemannianas também como espaços métricos com a noção de distância definida acima.

12.9 Variedades Riemannianas Completas e Teorema de Hopf-Rinow para Variedades Riemannianas

12.43 Definição. Seja M uma variedade riemanniana. Dizemos que M é **geodesicamente completa** se para todo $p \in M$ as geodésicas radiais $\gamma(t)$ partindo de p estão definidas para todo $t \in \mathbb{R}$. \square

Equivalentemente, M é geodesicamente completa se para todo $p \in M$ a aplicação exponencial está definida em todo $T_p M$ (Exercício 12.55).

Lembramos que um espaço métrico é completo se toda sequência de Cauchy é convergente.

12.44 Lema. *Seja M uma variedade riemanniana conexa. Se \exp_p está definida em todo $T_p M$, então qualquer ponto $q \in M$ pode ser ligado a p por um segmento geodésico γ tal que*

$$\ell(\gamma) = \text{dist}(p, q).$$

Prova: Seja $r = \text{dist}(p, q)$. Tome uma bola geodésica fechada $\overline{B}_\delta(p)$. Se $q \in \overline{B}_\delta(p)$, então existe uma geodésica radial minimizante ligando p a q e não há nada a provar. Se $q \notin \overline{B}_\delta(p)$, lembrando que a função distância é contínua e esferas geodésicas são conjuntos compactos, seja $x_0 \in S_\delta(p) = \partial \overline{B}_\delta(p)$ onde a função

$$f(x) = \text{dist}(x, q)$$

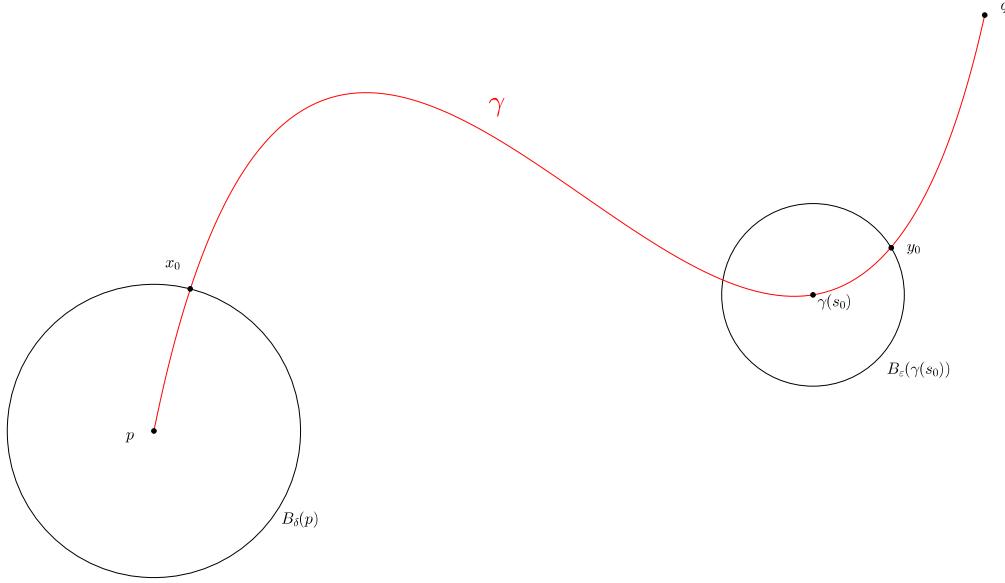
atinge um mínimo em $S_\delta(p)$. Seja $\gamma(s) = \exp_p(sV)$ a geodésica radial unitária ligando p a x_0 . Por hipótese, γ está definida para todo $t \in \mathbb{R}$. Para provar o lema, basta mostrar que

$$\gamma(r) = q.$$

Para provar isso, considere o conjunto não vazio (pois $0 \in A$)

$$A = \{s \in [0, r] : \text{dist}(\gamma(s), q) = r - s\}.$$

A é fechado em $[0, r]$ pela continuidade da função distância e de γ . Se provarmos que para todo $s_0 \in A$ vale $s_0 + \varepsilon \in A$ para todo $\varepsilon > 0$ suficientemente pequeno, isso implicará que A também é aberto no conjunto conexo $[0, r]$ e portanto $A = [0, r]$; em particular, $r \in A$, o que implica $\text{dist}(\gamma(r), q) = r - r = 0$, o que é equivalente a $\gamma(r) = q$.



Seja então $s_0 \in A$ e considere uma bola geodésica fechada $\overline{B}_\varepsilon(\gamma(s_0))$. Temos

$$r - s_0 = \text{dist}(\gamma(s_0), q) = \varepsilon + \min_{y \in \overline{B}_\varepsilon(\gamma(s_0))} \text{dist}(y, q) = \varepsilon + \text{dist}(y_0, q)$$

onde $y_0 \in S_\varepsilon(\gamma(s_0)) = \partial \overline{B}_\varepsilon(\gamma(s_0))$ é o ponto onde a função

$$g(y) = \text{dist}(y, q)$$

atinge um mínimo em $S_\varepsilon(\gamma(s_0))$. Para provar a afirmação, basta então mostrar que $y_0 = \gamma(s_0 + \varepsilon)$, pois neste caso

$$r - s_0 = \varepsilon + \text{dist}(\gamma(s_0 + \varepsilon), q),$$

onde

$$\text{dist}(\gamma(s_0 + \varepsilon), q) = r - (s_0 + \varepsilon).$$

De fato, temos

$$\text{dist}(p, y_0) \geq \text{dist}(p, q) - \text{dist}(q, y_0) = r - [r - (s_0 + \varepsilon)] = s_0 + \varepsilon.$$

Por outro lado, a curva quebrada que liga p a y_0 constituída do segmento geodésico γ que vai de p a $\gamma(s_0)$ e do raio geodésico que vai de $\gamma(s_0)$ a y_0 tem comprimento $s_0 + \varepsilon$. Portanto,

$$\text{dist}(p, y_0) = s_0 + \varepsilon$$

e esta curva quebrada é uma geodésica (logo, não é quebrada), donde $y_0 = \gamma(s_0 + \varepsilon)$, o que termina a demonstração. ■

12.45 Teorema (Teorema de Hopf-Rinow). *Uma variedade riemanniana conexa é geodesicamente completa se e somente se ela é completa como espaço métrico.*

Prova:

1. M variedade riemanniana completa como espaço métrico $\implies M$ variedade riemanniana geodesicamente completa.

Suponha por absurdo que exista uma geodésica unitária $\gamma : [0, a) \rightarrow M$ que não se estende a um intervalo $[0, a + \varepsilon)$ para nenhum $\varepsilon > 0$. Seja $\{t_i\}$ uma sequência crescente tal que $t_i \rightarrow a$; em particular, $\{t_i\}$ é uma sequência de Cauchy. Seja $q_i = \gamma(t_i)$. Como γ é parametrizada por comprimento de arco, segue que

$$\text{dist}(q_i, q_j) \leq |t_i - t_j|$$

e (q_i) é uma sequência de Cauchy em M . Logo $q_i \rightarrow q \in M$. Seja V uma vizinhança uniformemente normal de q e $\delta > 0$ tal que V está contido em qualquer bola geodésica de raio δ centrada em um ponto de V . Para j suficientemente grande temos $q_j \in V$ e $t_j > a - \delta$. O fato que $B_\delta(q_j)$ é uma bola geodésica implica que toda geodésica partindo de q_j existe por um intervalo de tempo pelo menos igual a δ . Em particular isso vale para a geodésica σ satisfazendo $\sigma(0) = q_j$ e $\sigma'(0) = \gamma'(t_j)$. Por unicidade de geodésica, esta é uma reparametrização de γ , isto é, $\sigma(t) = \gamma(t + t_j)$, logo $\tilde{\gamma}(t) = \sigma(t - t_j)$ é uma extensão de γ além de a , contradizendo a hipótese inicial.

2. M variedade riemanniana geodesicamente completa $\implies M$ espaço métrico completo.

Para provar a recíproca, demonstraremos um resultado mais forte:

Se existe $p \in M$ tal que \exp_p está definida em todo $T_p M$, então M é um espaço métrico completo.

Seja $\{q_i\} \subset M$ uma sequência de Cauchy. Para cada i , seja $\gamma_i(s) = \exp_p(sV_i)$ a geodésica radial unitária que liga p a q_i , e seja

$$d_i = \text{dist}(p, q_i),$$

de modo que pelo lema

$$q_i = \exp_p(d_i V_i).$$

Além disso, $\{d_i\}$ é uma sequência de Cauchy em \mathbb{R} , pois

$$|d_i - d_j| = |\text{dist}(p, q_i) - \text{dist}(p, q_j)| \leq \text{dist}(q_i, q_j).$$

Como sequências de Cauchy são limitadas, $\{d_i\}$ é limitada; além disso $\|V_i\| = 1$ para todo i , logo $\{d_i V_i\}$ é limitada em $T_p M$. Portanto, uma subsequência $\{d_{i_k} V_{i_k}\}$ converge para $V \in T_p M$. Por continuidade da aplicação exponencial,

$$q_{i_k} = \exp_p(d_{i_k} V_{i_k}) \rightarrow \exp_p(V).$$

Como a sequência original $\{q_i\}$ é de Cauchy, ela converge para o mesmo ponto para o qual sua subsequência converge. ■

Em particular, o conceito de variedade riemanniana geodesicamente completa é equivalente ao conceito de variedade riemanniana completa como espaço métrico para variedades conexas e para estas podemos nos referir simplesmente a uma *variedade riemanniana completa*, implicando ambos os conceitos.

12.46 Corolário. *Toda variedade riemanniana compacta é geodesicamente completa.*

Prova: Pois todo espaço métrico compacto é completo. ■

12.47 Corolário. *Uma subvariedade fechada de uma variedade riemanniana completa é geodesicamente completa na métrica induzida.*

Prova: Pois todo subconjunto fechado de um espaço métrico é completo. ■

12.48 Corolário. *Os conjuntos fechados e limitados de uma variedade riemanniana conexa são compactos se e somente se ela é geodesicamente completa.*

Prova: Seja $A \subset M$ fechado e limitado. Como A é limitado, A está contido em uma bola métrica $B(p; r)$. Pelo Lema 12.44, existe uma bola $B_R(0) \subset T_p M$ tal que

$$B(p; r) \subset \exp_p \left(\overline{B_R(0)} \right) = \overline{B_R(p)}.$$

A bola geodésica fechada $\overline{B_R(p)}$ é compacta, pois é a imagem de um compacto pela aplicação contínua \exp_p . Logo, A é um subconjunto fechado de um compacto, portanto é compacto.

Reciprocamente, a condição de conjuntos fechados limitados serem compactos em um espaço métrico implica que o espaço métrico é completo: sequências de Cauchy são limitadas, logo seu fecho é compacto e portanto possuem uma subsequência convergente, o que implica para sequências de Cauchy que a sequência toda é convergente. ■

12.49 Corolário. *Se M é uma variedade riemanniana conexa e existe $p \in M$ tal que \exp_p está definida em todo $T_p M$ então M é geodesicamente completa.*

Prova: Segue da demonstração do teorema de Hopf-Rinow. ■

12.50 Corolário. *Se M é uma variedade riemanniana conexa completa, então quaisquer dois pontos $p, q \in M$ podem ser ligados por um segmento geodésico γ tal que*

$$\ell(\gamma) = \text{dist}(p, q).$$

Prova: Segue do teorema de Hopf-Rinow e do Lema 12.44. ■

A recíproca não é válida: todos os pontos de um aberto convexo U de \mathbb{R}^n (uma variedade riemanniana na métrica euclideana) podem ser ligados por segmentos de retas, mas se U é um subconjunto próprio de \mathbb{R}^n ele não é uma variedade completa.

12.10 Apêndice: Geodésicas através do Método Variacional

Excepcionalmente denotaremos a derivada de uma curva $\alpha'(t)$ nesta seção por $\dot{\alpha}(t)$, para adotar uma notação consistente com a do Cálculo de Variações.

12.51 Definição. Seja M uma variedade diferenciável. Considere o espaço \mathcal{C} das curvas diferenciáveis em M que satisfazem a condição de fronteira $\alpha(a) = p$ e $\alpha(b) = q$:

$$\mathcal{C} = \{\alpha : [a, b] \rightarrow M : \alpha \text{ é diferenciável}, \alpha(a) = p \text{ e } \alpha(b) = q\}.$$

O lagrangiano L de um problema variacional para \mathcal{C} é uma função diferenciável

$$L : TM \rightarrow \mathbb{R}.$$

Associado ao lagrangiano L é o funcional ação $S : \mathcal{C} \rightarrow \mathbb{R}$ definido por

$$S(\alpha) = \int_a^b L(\alpha(t), \dot{\alpha}(t)) dt.$$

□

No Cálculo de Variações, procura-se encontrar um método para determinar curvas α que minimizam o funcional ação; isto é chamado o princípio da ação mínima.

12.52 Proposição (Equações de Euler-Lagrange). *Se uma curva $\gamma \in \mathcal{C}$ é um valor extremo do funcional ação, então a equação de Euler-Lagrange*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right) = \frac{\partial L}{\partial x^i}$$

$i = 1, \dots, n$, são satisfeitas ao longo de γ em coordenadas locais $(x^1, \dots, x^n, v^1, \dots, v^n)$ para TM .

Prova: Vamos considerar primeiro o caso $M = \mathbb{R}^n$ para compreender melhor a idéia da demonstração. Seja $\beta : [a, b] \rightarrow \mathbb{R}^n$ uma curva diferenciável arbitrária satisfazendo $\beta(a) = \beta(b) = 0$ e considere a seguinte variação de γ

$$F(s, t) = \gamma(t) + s\beta(t).$$

Denotando

$$\alpha_s(t) = F(s, t)$$

considere a função diferenciável real $f : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ definida por

$$f(s) = S(\alpha_s) = \int_a^b L(\alpha_s(t), \dot{\alpha}_s(t)) dt.$$

Como γ é um extremo para S , segue que 0 é um extremo para f , de modo que

$$f'(0) = 0.$$

Temos, derivando sob o sinal de integração e pela regra da cadeia,

$$f'(s) = \int_a^b \left[\frac{\partial L}{\partial x^i}(\alpha_s, \dot{\alpha}_s) \frac{d\alpha_s^i}{ds} + \frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \frac{d\dot{\alpha}_s^i}{ds} \right] dt.$$

Já que

$$\begin{aligned} \frac{d\alpha_s}{ds} &= \beta(t), \\ \frac{d\dot{\alpha}_s}{ds} &= \dot{\beta}(t), \end{aligned}$$

segue que

$$f'(s) = \int_a^b \left[\frac{\partial L}{\partial x^i}(\alpha_s, \dot{\alpha}_s) \beta^i + \frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \dot{\beta}^i \right] dt.$$

Por outro lado, integrando por partes, temos

$$\int_a^b \frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \dot{\beta}^i dt = \frac{\partial L}{\partial v^i} \beta^i \Big|_a^b - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \right) \beta^i dt = - \int_a^b \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \right) \beta^i dt,$$

de modo que

$$f'(s) = \int_a^b \left[\frac{\partial L}{\partial x^i}(\alpha_s, \dot{\alpha}_s) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\alpha_s, \dot{\alpha}_s) \right) \right] \beta^i dt.$$

Agora, $f'(0) = 0$ implica

$$\int_a^b \left[\frac{\partial L}{\partial x^i}(\gamma, \dot{\gamma}) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\gamma, \dot{\gamma}) \right) \right] \beta^i dt = 0$$

Como β é arbitrária, segue que

$$\frac{\partial L}{\partial x^i}(\gamma, \dot{\gamma}) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\gamma, \dot{\gamma}) \right) = 0$$

para todo i . ■

Em uma variedade riemanniana (M, g) existem dois lagrangianos associados de forma natural:

$$L(p, v) = \|v\|_p,$$

$$E(p, v) = \|v\|_p^2.$$

As ações destes lagrangianos são respectivamente o funcional **comprimento** e o funcional **energia**:

$$\begin{aligned}\mathcal{L}(\alpha) &= \int_a^b \|\alpha'(t)\|_p dt, \\ \mathcal{E}(\alpha) &= \int_a^b \|\alpha'(t)\|_p^2 dt.\end{aligned}$$

É mais fácil trabalhar com o funcional energia, já que o lagrangiano da energia não apresenta problemas de diferenciabilidade na origem. Além disso, um mínimo para o funcional energia é um mínimo para o funcional comprimento, pois a desigualdade de Cauchy-Schwartz implica que

$$\mathcal{L}(\alpha) = \int_a^b \|\alpha'(t)\|_p dt \leq \left(\int_a^b \|\alpha'(t)\|_p^2 dt \right)^{1/2} \left(\int_a^b 1 dt \right)^{1/2} = \sqrt{b-a} \mathcal{E}(\alpha)^{1/2}.$$

12.53 Proposição. *Geodésicas são extremos do funcional energia.*

Prova: Em coordenadas locais o lagrangiano da energia é

$$E(x^1, \dots, x^n, v^1, \dots, v^n) = g_{ij}(x^1, \dots, x^n) v^i v^j,$$

e ao longo de uma curva $\alpha(t) = (x^1(t), \dots, x^n(t))$ ele é dado por

$$E\left(x^1, \dots, x^n, \frac{dx^1}{dt}, \dots, \frac{dx^n}{dt}\right) = g_{ij}(x^1, \dots, x^n) \frac{dx^i}{dt} \frac{dx^j}{dt}.$$

Logo,

$$\frac{\partial E}{\partial x^k} = \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt},$$

enquanto que

$$\frac{\partial E}{\partial v^k} = 2g_{ik} \frac{dx^i}{dt}.$$

de modo que, pela regra da cadeia, segue que

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial E}{\partial v^k} \right) &= 2 \left(\frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + g_{ik} \frac{d^2 x^i}{dt^2} \right) \\ &= \frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + \frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + 2g_{ik} \frac{d^2 x^i}{dt^2} \\ &= \frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + \frac{\partial g_{jk}}{\partial x^i} \frac{dx^i}{dt} \frac{dx^j}{dt} + 2g_{ik} \frac{d^2 x^i}{dt^2}\end{aligned}$$

trocando i, j no segundo termo. Portanto, a equação de Euler-Lagrange do funcional energia

$$\frac{d}{dt} \left(\frac{\partial E}{\partial v^k} \right) = \frac{\partial E}{\partial x^k}$$

transforma-se na equação

$$2g_{ik} \frac{d^2 x^i}{dt^2} + \frac{\partial g_{ik}}{\partial x^j} \frac{dx^j}{dt} \frac{dx^i}{dt} + \frac{\partial g_{jk}}{\partial x^i} \frac{dx^i}{dt} \frac{dx^j}{dt} - \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0,$$

ou seja,

$$g_{ik} \frac{d^2 x^i}{dt^2} + \left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right) \frac{dx^i}{dt} \frac{dx^j}{dt} = 0.$$

Multiplicando por g^{mk} e somando em k , obtemos

$$\delta_i^m \frac{d^2x^i}{dt^2} + \left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right) g^{mk} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0,$$

que é exatamente a equação geodésica

$$\frac{d^2x^m}{dt^2} + \Gamma_{ij}^m \frac{dx^i}{dt} \frac{dx^j}{dt} = 0.$$

■

12.11 Exercícios

12.54 Exercício. Mostre que as transformações de Möbius introduzidas no Exemplo 12.32 são isometrias do plano hiperbólico e complete os outros detalhes.

12.55 Exercício. Mostre que M é geodesicamente completa se e somente se para todo $p \in M$ a aplicação exponencial está definida em todo $T_p M$.

Capítulo 13

Curvatura

13.1 Mais sobre Tensores

Observe que o espaço vetorial real $\mathcal{T}(M) = \mathcal{T}_1(M) = \mathcal{T}_1^0(M)$ dos campos diferenciáveis de uma variedade diferenciável M também pode ser considerado como um módulo sobre o anel $C^\infty(M)$. Pensando desta forma, estamos considerando a estrutura linear de $\mathcal{T}(M)$ quando tomamos as funções reais diferenciáveis em M como escalares. Da mesma forma, o espaço vetorial real $\mathcal{T}^1(M) = \mathcal{T}_0^1(M)$ dos campos covetoriais é um módulo sobre $C^\infty(M)$.

Um campo tensorial $T \in \mathcal{T}_l^k(M)$ então induz naturalmente uma aplicação multilinear sobre $C^\infty(M)$

$$T : \mathcal{T}_1(M) \times \dots \times \mathcal{T}_1(M) \times \mathcal{T}^1(M) \times \dots \times \mathcal{T}^1(M) \longrightarrow C^\infty(M) \quad (13.1)$$

definindo

$$T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)(p) = T_p \left(X_1|_p, \dots, X_k|_p, \omega^1|_p, \dots, \omega^l|_p \right). \quad (13.2)$$

De fato, a função $T(X_1, \dots, X_k, \omega_1, \dots, \omega_l) \in C^\infty(M)$ e T é claramente multilinear. Além disso, pode-se provar que toda aplicação multilinear (13.1) sobre $C^\infty(M)$ é induzida por um campo tensorial em $\mathcal{T}_l^k(M)$ (Exercício 13.38). Em particular, observe que se a base coordenada associada para o espaço tensorial $\mathcal{T}_l^k(T_p M)$ é

$$(\mathcal{B}_l^k)_p = \left\{ dx^{i_1}|_p \otimes \dots \otimes dx^{i_k}|_p \otimes \partial_{j_1}|_p \otimes \dots \otimes \partial_{j_l}|_p \right\}_{\substack{1 \leq i_1, \dots, i_k \leq n \\ 1 \leq j_1, \dots, j_l \leq n}}$$

para todo $p \in V$, de modo que nesta base coordenada o campo tensorial T se escreve na forma

$$T_p = T_{i_1, \dots, i_k}^{j_1, \dots, j_l}(p) dx^{i_1}|_p \otimes \dots \otimes dx^{i_k}|_p \otimes \partial_{j_1}|_p \otimes \dots \otimes \partial_{j_l}|_p,$$

então se os campos vetoriais e covetoriais se escrevem localmente na forma

$$\begin{aligned} X_r|_p &= X^{i_r}(p) \partial_{i_r}|_p, \quad r = 1, \dots, k, \\ \omega^s|_p &= \omega_{j_s}(p) dx^{j_s}|_p, \quad s = 1, \dots, l, \end{aligned}$$

temos

$$T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)(p) = X^{i_1} \dots X^{i_k} \omega_{j_1} \dots \omega_{j_l} T_{i_1, \dots, i_k}^{j_1, \dots, j_l}(p). \quad (13.3)$$

Ou seja, o valor de $T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)$ em p depende apenas das componentes do campo tensorial T em p e dos valores dos campos vetoriais e de 1-formas em p . Neste sentido, dizemos que tensores são *puntuais*. Contraste isso com uma conexão, que não é um tensor porque ao invés da linearidade satisfaz a regra do produto.

13.2 O Tensor Curvatura

Quando Riemann introduziu a noção de métrica em variedades, ele precisava mostrar que elas não eram localmente isométricas ao espaço euclidiano. Para provar isso, ele introduziu a noção de *curvatura*.

13.2.1 O Endomorfismo Curvatura

Para motivar a definição de curvatura, considere \mathbb{R}^n com a métrica η_{ij}^k , $k = 0, \dots, n$, e campos vetoriais $X, Y, Z \in \mathcal{T}(\mathbb{R}^n)$. Observe que

$$\begin{aligned}\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z &= X (\nabla_Y Z)^j \frac{\partial}{\partial x^j} - Y (\nabla_X Z)^j \frac{\partial}{\partial x^j} \\ &= X [Y (Z^j)] \frac{\partial}{\partial x^j} - Y [X (Z^j)] \frac{\partial}{\partial x^j} \\ &= (X [Y (Z^j)] - Y [X (Z^j)]) \frac{\partial}{\partial x^j} \\ &= (XY - YX) (Z^j) \frac{\partial}{\partial x^j} \\ &= \nabla_{[X,Y]} Z.\end{aligned}$$

Esta relação é válida para qualquer variedade métrica localmente isométrica a \mathbb{R}^n (Exercício 13.39).

13.1 Definição. Seja M uma variedade métrica. O **endomorfismo curvatura** de M é o campo tensorial de ordem $(3, 1)$

$$R : \mathcal{T}(M) \times \mathcal{T}(M) \times \mathcal{T}(M) \longrightarrow \mathcal{T}(M)$$

definido por

$$R(X, Y) Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \quad (13.4)$$

□

Observamos que alguns autores, tais como [Carmo], definem a curvatura com o sinal oposto, mas a maioria define como acima. O fato de termos identificado R com um tensor do tipo $(3, 1)$ segue da Proposição 9.19 e pelo fato de R ser multilinear sobre $C^\infty(M)$, conforme é demonstrado na proposição a seguir. Em outras palavras, R é identificado com o campo tensorial

$$\bar{R} : \mathcal{T}_1(M) \times \mathcal{T}_1(M) \times \mathcal{T}_1(M) \times \mathcal{T}^1(M) \longrightarrow \mathbb{R}$$

definido por

$$\bar{R}(X, Y, Z, \omega) = \omega(R(X, Y) Z).$$

Uma motivação extra para a definição do tensor curvatura pode ser vista no apêndice no final deste capítulo.

13.2 Proposição. R é multilinear sobre $C^\infty(M)$.

Prova: Temos, em relação à segunda variável,

$$\begin{aligned}R(X, Y_1 + Y_2) Z &= \nabla_X \nabla_{Y_1 + Y_2} Z - \nabla_{Y_1 + Y_2} \nabla_X Z - \nabla_{[X, Y_1 + Y_2]} Z \\ &= \nabla_X \nabla_{Y_1} Z + \nabla_X \nabla_{Y_2} Z - \nabla_{Y_1} \nabla_X Z - \nabla_{Y_2} \nabla_X Z \\ &\quad - \nabla_{[X, Y_1]} Z - \nabla_{[X, Y_2]} Z \\ &= R(X, Y_1) Z + R(X, Y_2) Z\end{aligned}$$

e

$$\begin{aligned}
R(X, fY)Z &= \nabla_X \nabla_{fY} Z - \nabla_{fY} \nabla_X Z - \nabla_{[X, fY]} Z \\
&= \nabla_X(f\nabla_Y Z) - f\nabla_Y \nabla_X Z - \nabla_{f[X, Y] + (Xf)Y} Z \\
&= f\nabla_X \nabla_Y Z + (Xf) \nabla_Y Z - f\nabla_Y \nabla_X Z - f\nabla_{[X, Y]} Z - (Xf) \nabla_Y Z \\
&= f(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z) \\
&= fR(X, Y)Z.
\end{aligned}$$

A linearidade em relação à primeira variável é imediatamente estabelecida uma vez que se observa que

$$R(X, Y)Z = -R(Y, X)Z. \quad (13.5)$$

De fato,

$$\begin{aligned}
R(Y, X)Z &= \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z - \nabla_{[Y, X]} Z \\
&= -\nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z - \nabla_{-[X, Y]} Z \\
&= -\nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z + \nabla_{[X, Y]} Z \\
&= -R(X, Y)Z.
\end{aligned}$$

Para estabelecer a linearidade em relação à terceira variável, temos

$$\begin{aligned}
R(X, Y)(Z_1 + Z_2) &= \nabla_X \nabla_Y (Z_1 + Z_2) - \nabla_Y \nabla_X (Z_1 + Z_2) - \nabla_{[X, Y]} (Z_1 + Z_2) \\
&= \nabla_X \nabla_Y Z_1 - \nabla_Y \nabla_X Z_1 - \nabla_{[X, Y]} Z_1 \\
&\quad + \nabla_X \nabla_Y Z_2 - \nabla_Y \nabla_X Z_2 - \nabla_{[X, Y]} Z_2 \\
&= R(X, Y)Z_1 + R(X, Y)Z_2
\end{aligned}$$

e

$$\begin{aligned}
R(X, Y)(fZ) &= \nabla_X \nabla_Y (fZ) - \nabla_Y \nabla_X (fZ) - \nabla_{[X, Y]} (fZ) \\
&= \nabla_X(f\nabla_Y Z) + \nabla_X((Yf)Z) - \nabla_Y(f\nabla_X Z) - \nabla_Y((Xf)Z) \\
&\quad - f\nabla_{[X, Y]} Z - [X, Y](f)Z \\
&= f\nabla_X \nabla_Y Z + (Xf) \nabla_Y Z + (Yf) \nabla_X Z + X(Yf)Z \\
&\quad - f\nabla_Y \nabla_X Z - (Yf) \nabla_X Z - (Xf) \nabla_Y Z - Y(Xf)Z \\
&\quad - f\nabla_{[X, Y]} Z - [X, Y](f)Z \\
&= f\nabla_X \nabla_Y Z - f\nabla_Y \nabla_X Z + [X(Yf) - Y(Xf)]Z - f\nabla_{[X, Y]} Z - [X, Y](f)Z \\
&= f\nabla_X \nabla_Y Z - f\nabla_Y \nabla_X Z - f\nabla_{[X, Y]} Z + [Y, X](f)Z - [X, Y](f)Z \\
&= f(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z) \\
&= fR(X, Y)Z.
\end{aligned}$$

■

13.2.2 O Significado da Curvatura Nula

13.3 Proposição. Em \mathbb{R}^n ,

$$R = 0. \quad (13.6)$$

Em particular, a derivada covariante comuta

$$\nabla_{\partial_i} \nabla_{\partial_j} = \nabla_{\partial_j} \nabla_{\partial_i}. \quad (13.7)$$

Prova: Na introdução a esta seção vimos que $R \equiv 0$ para \mathbb{R}^n (de fato, foi o que nos levou a definir R em primeiro lugar; este resultado é trazido aqui apenas para referência e ênfase). Como $[\partial_i, \partial_j] = 0$, segue que

$$\nabla_{\partial_i} \nabla_{\partial_j} Z = \nabla_{\partial_j} \nabla_{\partial_i} Z$$

para todo campo vetorial $Z \in \mathcal{T}(\mathbb{R}^n)$. ■

Assim, a curvatura mede também o quanto a derivada covariante deixa de comutar. De fato, o principal papel do termo $\nabla_{[\partial_i, \partial_j]}$ é fazer com que a curvatura seja um tensor (veja o apêndice deste capítulo para mais detalhes).

13.4 Teorema. *Uma variedade métrica M^n é localmente isométrica a \mathbb{R}^n se e somente se $R = 0$.*

Prova: Vamos mostrar que se $R = 0$ então M é localmente isométrica a \mathbb{R}^n , já que a recíproca é óbvia (a curvatura é um invariante isométrico local; Exercício 13.39).

Dado $p \in M$, escolha uma base ortonormal

$$\left\{ E_1|_p, \dots, E_n|_p \right\}$$

para $T_p M$. Seja $\{x^1, \dots, x^n\}$ um sistema de coordenadas para uma vizinhança de p tal que $\partial_i|_p = E_i|_p$ para todo i (por exemplo, um sistema de coordenadas normal).

Diminuindo a vizinhança se necessário, podemos assumir que o domínio da parametrização é um pequeno cubo

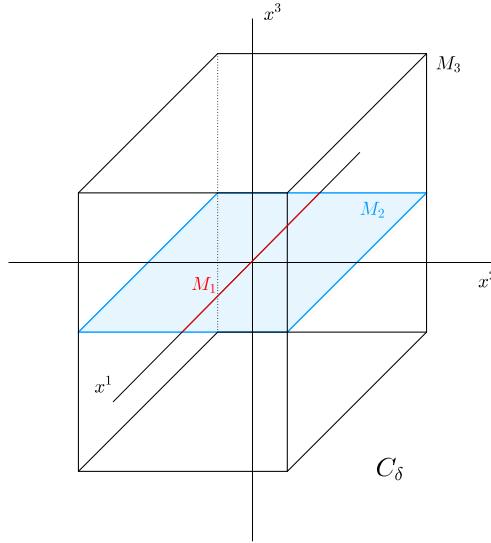
$$C_\delta = \{x \in \mathbb{R}^n : |x^i| < \delta, i = 1, \dots, n\}.$$

Para cada i faça um transporte paralelo do vetor $E_i|_p$ ao longo de todas as imagens parametrizadas dos eixos do cubo, cobrindo todos os pontos da vizinhança parametrizada pelo cubo: por exemplo, comece transportando paralelamente o vetor $E_i|_p$ ao longo da imagem parametrizada do eixo x^1 ; então, a partir de cada ponto da imagem parametrizada do eixo x^1 , transporte paralelamente ao longo da imagem parametrizada do segmento paralelo ao eixo x^2 ; então, sucessivamente, transporte paralelamente ao longo das imagens parametrizadas dos segmentos paralelos aos eixos x^3, \dots, x^n . Deste modo, obtemos n campos vetoriais ortonormais

$$E_1, \dots, E_n$$

definidos na vizinhança parametrizada pelo cubo C_δ . Estes campos vetoriais são suaves pela aplicação sucessiva do teorema da dependência suave das soluções de EDOs nas condições iniciais, pois os campos E_i são as soluções do sistema de EDOs

$$\frac{dE_i}{dt} = 0.$$



Afirmamos que os campos E_1, \dots, E_n são paralelos em C_δ . Para provar isso, basta provar que

$$\nabla_{\partial_i} E_j = 0$$

para todos i, j .

Fixe j . Seja

$$M^k = \{x \in C_\delta : x^{k+1} = \dots = x^n = 0\}$$

a fatia k -dimensional do cubo correspondente à interseção do plano k -dimensional $(x^1, \dots, x^k, \dots, 0)$ com o cubo C_δ . Por construção, $\nabla_{\partial_i} E_j = 0$ em M^1 , $\nabla_{\partial_2} E_j = 0$ em M^2 e, em geral, $\nabla_{\partial_k} E_j = 0$ em M^k .

Provaremos por indução que

$$\nabla_{\partial_1} E_j = \dots = \nabla_{\partial_k} E_j = 0 \quad (13.8)$$

em M^k para todo k ; em particular, para $k = n$, obtemos a afirmação.

Já vimos que o caso $k = 1$ é válido. Assuma a validade de (13.8) para k . Em M^{k+1} temos $\nabla_{\partial_{k+1}} E_j = 0$ por construção e

$$\nabla_{\partial_1} E_j = \dots = \nabla_{\partial_k} E_j = 0$$

em M^k por hipótese de indução. Resta apenas mostrar que $\nabla_{\partial_i} E_j = 0$ para $i = 1, \dots, k$ no resto de M^{k+1} . Para isso, como já temos $\nabla_{\partial_i} E_j = 0$ na fatia k -dimensional M^k (a fatia que ocupa o centro de M^{k+1}), basta provar que

$$\nabla_{\partial_{k+1}} (\nabla_{\partial_i} E_j) = 0,$$

pois isso implicará que o campo $\nabla_{\partial_i} E_j$ continuará sendo nulo ao ser paralelamente transportado para o resto de M^{k+1} . Mas, como $R = 0$, temos

$$\nabla_{\partial_i} (\nabla_{\partial_{k+1}} E_j) - \nabla_{\partial_{k+1}} (\nabla_{\partial_i} E_j) - \nabla_{[\partial_i, \partial_{k+1}]} E_j = 0$$

ou seja, usando o fato que $[\partial_i, \partial_{k+1}] = 0$ e a hipótese de indução,

$$\nabla_{\partial_{k+1}} (\nabla_{\partial_i} E_j) = \nabla_{\partial_i} (\nabla_{\partial_{k+1}} E_j) = 0.$$

Isso conclui a demonstração da afirmação que os campos E_1, \dots, E_n são paralelos em C_δ .

Agora, pela simetria da conexão riemanniana temos

$$[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_j} E_i.$$

Como os campos E_1, \dots, E_n são paralelos em C_δ , como acabamos de provar, segue que

$$\nabla_{E_i} E_j = \nabla_{E_j} E_i = 0,$$

onde

$$[E_i, E_j] = 0$$

para todos i, j em C_δ . Portanto, pelo Teorema 8.43, existem coordenadas $\{y^1, \dots, y^n\}$ para uma vizinhança possivelmente menor de p , tais que

$$E_i = \frac{\partial}{\partial y^i}.$$

Nestas coordenadas,

$$g_{ij} = \left\langle \frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j} \right\rangle = \langle E_i, E_j \rangle = \eta_{ij}^k,$$

logo a parametrização das coordenadas y^i define uma isometria entre esta vizinhança parametrizada de p e um aberto de \mathbb{R}^n . ■

Para a demonstração original deste resultado para variedades riemannianas por Riemann, comentada e em notação moderna, veja [Spivak], Vol. II, pp. 179-204. Observe que a demonstração acima mostra que se pudermos encontrar campos E_1, \dots, E_n ortonormais e paralelos na variedade M , então ela é localmente isométrica a \mathbb{R}^n . Consequentemente, tais campos em geral não existem. Isso significa que o transporte paralelo de vetores ao longo de curvas com extremidades coincidentes em geral dá resultados diferentes (como exemplo, considere o transporte paralelo de um vetor na esfera do polo norte ao polo sul, ao longo de dois meridianos fazendo um ângulo reto). Assim, o tensor curvatura mede a independência de caminhos do transporte paralelo (veja Teorema 13.37 no apêndice a este capítulo). Outra coisa importante que a curvatura mede é o assim chamado *desvio geodésico* (veja Proposição 13.35 no apêndice a este capítulo). Este e a influência da curvatura no transporte paralelo são discutidos detalhadamente no apêndice deste capítulo.

13.5 Proposição. Vale

$$R(\partial_i, \partial_j) \partial_k = (\nabla_{\partial_i} \nabla_{\partial_j} - \nabla_{\partial_j} \nabla_{\partial_i}) \partial_k. \quad (13.9)$$

Em particular, os componentes do tensor endomorfismo curvatura são

$$R_{ijk}^l = \Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l + \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j}. \quad (13.10)$$

Prova: A primeira afirmação segue de $[\partial_i, \partial_j] = 0$.

A segunda afirmação segue da primeira por cálculo direto, observando primeiramente que os componentes do tensor R são definidos da forma usual, usando a multilinearidade do tensor, por

$$R(X, Y) Z = R_{ijk}^l X^i Y^j Z^k \frac{\partial}{\partial x^l}, \quad (13.11)$$

de modo que

$$R(\partial_i, \partial_j) \partial_k = R_{ijk}^l \frac{\partial}{\partial x^l}. \quad (13.12)$$

Logo, lembrando que

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^m \partial_m,$$

segue que

$$\begin{aligned}
R(\partial_i, \partial_j) \partial_k &= \nabla_{\partial_i} (\nabla_{\partial_j} \partial_k) - \nabla_{\partial_j} (\nabla_{\partial_i} \partial_k) \\
&= \nabla_{\partial_i} (\Gamma_{jk}^m \partial_m) - \nabla_{\partial_j} (\Gamma_{ik}^m \partial_m) \\
&= \nabla_{\partial_i} (\Gamma_{jk}^m \partial_m) - \nabla_{\partial_j} (\Gamma_{ik}^m \partial_m) \\
&= \Gamma_{jk}^m \nabla_{\partial_i} \partial_m + \frac{\partial \Gamma_{jk}^m}{\partial x^i} \partial_m - \Gamma_{ik}^m \nabla_{\partial_j} \partial_m - \frac{\partial \Gamma_{ik}^m}{\partial x^j} \partial_m \\
&= \Gamma_{jk}^m \Gamma_{im}^l \partial_l + \frac{\partial \Gamma_{jk}^l}{\partial x^i} \partial_l - \Gamma_{ik}^m \Gamma_{jm}^l \partial_l - \frac{\partial \Gamma_{ik}^l}{\partial x^j} \partial_l \\
&= \left(\Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l + \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j} \right) \partial_l.
\end{aligned}$$

■

13.6 Proposição (Simetrias do Tensor Endomorfismo Curvatura). *O tensor endomorfismo curvatura satisfaçõas seguintes propriedades:*

(a) $R(X, Y)Z = -R(Y, X)Z$.

(b) (*identidade de Bianchi*) $R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$.

Em termos dos componentes do tensor endomorfismo curvatura, estas identidades são respectivamente equivalentes a

(a) $R_{ijk}^l = -R_{jik}^l$.

(b) $R_{ijk}^l + R_{jki}^l + R_{kij}^l = 0$.

Consequentemente,

$$R_{iik}^l = 0.$$

Prova: A primeira propriedade segue diretamente da definição, como já vimos. A segunda propriedade segue da simetria da conexão, aplicada duas vezes, e da identidade de Jacobi para o colchete de Lie:

$$\begin{aligned}
R(X, Y)Z + R(Y, Z)X + R(Z, X)Y &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z \\
&\quad + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X - \nabla_{[Y, Z]} X \\
&\quad + \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y - \nabla_{[Z, X]} Y \\
&= \nabla_Y (\nabla_Z X - \nabla_X Z) + \nabla_Z (\nabla_X Y - \nabla_Y X) + \nabla_X (\nabla_Y Z - \nabla_Z Y) \\
&\quad - \nabla_{[X, Y]} Z - \nabla_{[Y, Z]} X - \nabla_{[Z, X]} Y \\
&= \nabla_Y [Z, X] + \nabla_Z [X, Y] + \nabla_X [Y, Z] - \nabla_{[X, Y]} Z - \nabla_{[Y, Z]} X - \nabla_{[Z, X]} Y \\
&= (\nabla_Y [Z, X] - \nabla_{[Z, X]} Y) + (\nabla_Z [X, Y] - \nabla_{[X, Y]} Z) + (\nabla_X [Y, Z] - \nabla_{[Y, Z]} X) \\
&= [Y, [Z, X]] + [Z, [X, Y]] + [X, [Y, Z]] \\
&= 0,
\end{aligned}$$

A última propriedade é consequência de (a). ■

13.2.3 Operação de Subir ou Descer um Índice

Uma propriedade importante do produto interno é que ele permite converter vetores em covetores e vice-versa. De fato, ela permite definir um isomorfismo entre os espaços $\mathcal{T}_1(M)$ de campos vetoriais de M e $\mathcal{T}^1(M)$ de campos covetoriais de M , chamado o *isomorfismo musical*. A escolha deste nome se refere aos símbolos escolhidos para denotar o isomorfismo e seu inverso, conforme veremos na definição a seguir.

13.7 Definição. Seja M uma variedade métrica. Definimos o isomorfismo **bemol**

$$\begin{array}{ccc} \flat : & \mathcal{T}_1(M) & \longrightarrow \mathcal{T}^1(M) \\ & X & \mapsto X^\flat \end{array}$$

que leva X no covetor X^\flat definido por

$$X^\flat(Y) = g(X, Y) = \langle X, Y \rangle \quad (13.13)$$

para todo Y . Seu inverso é o isomorfismo **sustenido**

$$\begin{array}{ccc} \sharp : & \mathcal{T}^1(M) & \longrightarrow \mathcal{T}_1(M) \\ & \omega & \mapsto \omega^\sharp \end{array}$$

que leva o covetor ω no vetor ω^\sharp definido por

$$g(\omega^\sharp, Y) = \langle \omega^\sharp, Y \rangle = \omega(Y) \quad (13.14)$$

para todo Y . \square

Em coordenadas,

$$X^\flat(Y) = \langle X^i \partial_i, Y^j \partial_j \rangle = g_{ij} X^i Y^j = g_{ij} X^i dx^j(Y),$$

ou seja,

$$X^\flat = g_{ij} X^i dx^j. \quad (13.15)$$

Escrevendo os componentes do covetor X^\flat em coordenadas na forma

$$X^\flat = X_j dx^j, \quad (13.16)$$

segue que

$$X_j = g_{ij} X^i. \quad (13.17)$$

Diz-se que o covetor X^\flat é obtido a partir do vetor X **descendo um índice**. Por este motivo esta operação é denotada pelo símbolo *bemol*, porque em partituras musicais o símbolo bemol é usado para abaixar a altura da nota musical que lhe segue. Já no caso do isomorfismo inverso, escrevendo em coordenadas o vetor ω^\sharp na forma

$$\omega^\sharp = \omega^i \partial_i, \quad (13.18)$$

segue que

$$\langle \omega^k \partial_k, \partial_j \rangle = \omega(\partial_j),$$

onde

$$\omega_j = g_{kj} \omega^k.$$

Multiplicando pela matriz inversa g^{ij} , como

$$g^{ij} g_{jk} \omega^k = g^{ij} g_{jk} \omega^k = \delta_k^i \omega^k = \omega^i,$$

segue que

$$\omega^i = \sum_{j=1}^n g^{ij} \omega_j. \quad (13.19)$$

Diz-se que o vetor ω^\sharp é obtido a partir do covetor ω **subindo um índice**. Por este motivo esta operação é denotada pelo símbolo *sustenido*, porque em partituras musicais o símbolo sustenido é usado para subir a altura da nota musical que lhe segue.

O vetor gradiente é definido a partir da operação de subir um índice:

13.8 Definição. Seja M uma variedade métrica. Dada uma função $f \in C^\infty(M)$, definimos o **campo gradiente** de f por

$$\text{grad } f = df^\sharp.$$

Em outras palavras, o vetor gradiente é definido por

$$\langle \text{grad } f, Y \rangle = df(Y)$$

para todo Y . \square

Em coordenadas, como

$$df = \frac{\partial f}{\partial x^i} dx^i,$$

temos

$$\text{grad } f = g^{ij} \frac{\partial f}{\partial x^i} \partial_j. \quad (13.20)$$

A operação de subir ou descer um índice pode ser aplicada a qualquer tensor:

13.9 Definição. Seja M uma variedade métrica. Se $T \in \mathcal{T}_l^k(M)$, então $T^\flat \in \mathcal{T}_{l-1}^{k+1}(M)$ é o tensor definido por

$$T^\flat(X_1, \dots, X_k, X_{k+1}, \omega^1, \dots, \omega^{q-1}, \omega^{q+1}, \dots, \omega^l) = T(X_1, \dots, X_p, \dots, X_k, \omega^1, \dots, \omega^{q-1}, X_{k+1}^\flat, \omega^{q+1}, \dots, \omega^l)$$

e $T^\sharp \in \mathcal{T}_{l+1}^{k-1}(M)$ é o tensor definido por

$$T^\sharp(X_1, \dots, X_{p-1}, X_{p+1}, \dots, X_k, \omega^1, \dots, \omega^l, \omega^{l+1}) = T(X_1, \dots, X_{p-1}, (\omega^{l+1})^\sharp, X_{p+1}, \dots, X_k, \omega^1, \dots, \omega^l).$$

\square

Note que ao aplicar a operação de subir ou descer um índice de um tensor teremos que explicitar qual índice estamos subindo ou descendo. Em geral isto é feito em palavras, sem o uso de um símbolo especial. Em coordenadas, descendo o último índice, as componentes do tensor resultante são dadas por

$$\begin{aligned} T_{i_1 \dots i_k i_{k+1}}^{j_1 \dots j_{l-1}} &= T^\flat(\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, dx^{j_1}, \dots, dx^{j_{l-1}}) \\ &= T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_{l-1}}, (\partial_{i_{k+1}})^\flat) \\ &= T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_{l-1}}, g_{i_{k+1} p} dx^p) \\ &= g_{i_{k+1} p} T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_{l-1}}, dx^p) \\ &= g_{i_{k+1} p} T_{i_1 \dots i_k}^{j_1 \dots j_{l-1} p}, \end{aligned}$$

que escrevemos para futura referência

$$T_{i_1 \dots i_k i_{k+1}}^{j_1 \dots j_{l-1}} = g_{i_{k+1} p} T_{i_1 \dots i_k}^{j_1 \dots j_{l-1} p}. \quad (13.21)$$

Subindo o último índice, as componentes do tensor resultante são dadas por

$$\begin{aligned} T_{i_1 \dots i_{k-1}}^{j_1 \dots j_l j_{l+1}} &= T^\sharp(\partial_{i_1}, \dots, \partial_{i_{k-1}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}) \\ &= T(\partial_{i_1}, \dots, \partial_{i_{k-1}}, (dx^{j_{l+1}})^\sharp, dx^{j_1}, \dots, dx^{j_l}) \\ &= T(\partial_{i_1}, \dots, \partial_{i_{k-1}}, g^{j_{l+1} q} \partial_q, dx^1, \dots, dx^l) \\ &= g^{j_{l+1} q} T(\partial_{i_1}, \dots, \partial_{i_{k-1}}, \partial_q, dx^1, \dots, dx^l) \\ &= g^{j_{l+1} q} T_{i_1 \dots i_{k-1} q}^{j_1 \dots j_l}, \end{aligned}$$

que escrevemos para futura referência

$$T_{i_1 \dots i_{k-1}}^{j_1 \dots j_l j_{l+1}} = g^{j_{l+1} q} T_{i_1 \dots i_{k-1} q}^{j_1 \dots j_l}. \quad (13.22)$$

Outra aplicação importante do isomorfismo musical é estender a operação de tomar o traço de tensores. Enquanto que o traço de tensores elimina um índice covariante e um índice contravariante, o *traço em relação à métrica* definido a seguir elimina dois índices covariantes.

13.10 Definição. Seja M uma variedade métrica. Se $T \in \mathcal{T}_l^k(M)$, então o **traço de T em relação à métrica g** é o tensor $\text{tr}_g T \in \mathcal{T}_l^{k-2}(M)$ definido por

$$\text{tr}_g T = \text{tr}(T^\sharp).$$

□

Em coordenadas, o traço de T em relação à métrica g em relação aos dois últimos índices, subindo o último índice antes de tomar o traço, é dado por

$$(\text{tr}_g T)_{i_1 \dots i_{k-2}}^{j_1 \dots j_l} = (T^\sharp)_{i_1 \dots i_{k-2} i}^{j_1 \dots j_l i} = g^{ij} T_{i_1 \dots i_{k-2} ij}^{j_1 \dots j_l}.$$

Por exemplo, se $T \in \mathcal{T}^2(M)$ é um tensor simétrico, de forma que não importa qual índice escolhemos subir, então $\text{tr}_g T \in \mathcal{T}^0(M) = C^\infty(M)$ é a função definida por

$$\text{tr}_g T = (T^\sharp)_i^i = g^{ij} T_{ij}.$$

13.11 Exemplo (Contração do Tensor Métrica). Estamos agora em condição de entender porque denotamos a inversa da matriz métrica $G = (g_{ij})$ por $G^{-1} = (g^{ij})$, ou seja, porque

$$g^{ik} g_{kj} = \delta_j^i.$$

De fato, como vimos acima,

$$g^{ik} g_{kj} = g_j^i$$

e

$$g_j^i = g^\sharp(e_i, e^j) = g\left(e_i, (e^j)^\sharp\right) = e^j(e_i) = \delta_j^i.$$

□

13.2.4 O Tensor Curvatura

13.12 Definição. Seja M uma variedade métrica. O **tensor curvatura** de M é o campo tensorial covariante de ordem 4

$$R : \mathcal{T}(M) \times \mathcal{T}(M) \times \mathcal{T}(M) \times \mathcal{T}(M) \longrightarrow C^\infty(M)$$

definido por

$$R(X, Y, Z, W) = \langle R(X, Y)Z, W \rangle.$$

□

Em outras palavras, o tensor curvatura é obtido através de descer o índice do tensor endomorfismo curvatura. Usaremos a mesma letra R para denotar os dois tensores; ficará claro dentro do contexto a qual tensor estamos nos referindo. Enquanto que o tensor endomorfismo curvatura é dado localmente por

$$R = R_{ijk}^l dx^i \otimes dx^j \otimes dx^k \otimes \partial_l, \quad (13.23)$$

o tensor curvatura é dado localmente por

$$R = R_{ijkl} dx^i \otimes dx^j \otimes dx^k \otimes dx^l. \quad (13.24)$$

13.13 Proposição. *As componentes do tensor curvatura são*

$$R_{ijkl} = g_{ml} R_{ijk}^m. \quad (13.25)$$

Em particular,

$$R_{ijk}^l = g^{ml} R_{ijkm}. \quad (13.26)$$

Prova: Pois o tensor curvatura é obtido através do tensor endomorfismo curvatura pela operação de descer o seu único índice contravariante, enquanto que o tensor endomorfismo curvatura é obtido através do tensor curvatura pela operação inversa, isto é, subir o seu último índice. Também podemos provar diretamente: a primeira identidade segue de

$$\begin{aligned} R_{ijkl} &= R(\partial_i, \partial_j, \partial_k, \partial_l) = \langle R(\partial_i, \partial_j) \partial_k, \partial_l \rangle \\ &= \langle R_{ijk}^m \partial_m, \partial_l \rangle = R_{ijk}^m \langle \partial_m, \partial_l \rangle \\ &= g_{ml} R_{ijk}^m, \end{aligned}$$

enquanto que a segunda identidade segue de

$$g^{ml} R_{ijkm} = g^{ml} g_{pm} R_{ijk}^p = g_{pm} g^{ml} R_{ijk}^p = \delta_p^l R_{ijk}^p = R_{ijk}^l.$$

■

Potencialmente, o tensor curvatura tem n^4 componentes. A presença de simetrias reduz bastante o número de componentes não nulas e de componentes independentes:

13.14 Proposição (Simetrias do Tensor Curvatura). *O tensor curvatura satisfaz as seguintes propriedades:*

- (a) $R(X, Y, Z, W) = -R(Y, X, Z, W).$
- (b) $R(X, Y, Z, W) = -R(X, Y, W, Z).$
- (c) $R(X, Y, Z, W) = R(Z, W, X, Y).$
- (d) (*identidade de Bianchi*) $R(X, Y, Z, W) + R(Y, Z, X, W) + R(Z, X, Y, W) = 0.$

Em termos dos componentes do tensor curvatura, estas identidades são respectivamente equivalentes a

- (a) $R_{ijkl} = -R_{jikl}.$
- (b) $R_{ijkl} = -R_{ijlk}.$
- (c) $R_{ijkl} = R_{klji}.$
- (d) $R_{ijkl} + R_{jkil} + R_{kijl} = 0.$

Prova: (a) Segue imediatamente de

$$R(X, Y) Z = -R(Y, X) Z.$$

(b) Como

$$R(X, Y, Z + W, Z + W) = R(X, Y, Z, Z) + R(X, Y, Z, W) + R(X, Y, W, Z) + R(X, Y, W, W),$$

a identidade segue se provarmos que

$$R(X, Y, Z, Z) = 0. \quad (13.27)$$

A identidade (13.27) segue da compatibilidade da conexão com a métrica. De fato, temos

$$\begin{aligned} R(X, Y, Z, Z) &= \langle R(X, Y) Z, Z \rangle \\ &= \langle \nabla_X \nabla_Y Z, Z \rangle - \langle \nabla_Y \nabla_X Z, Z \rangle - \langle \nabla_{[X, Y]} Z, Z \rangle. \end{aligned}$$

Daí, pela compatibilidade da métrica,

$$\begin{aligned}\langle \nabla_X \nabla_Y Z, Z \rangle &= X \langle \nabla_Y Z, Z \rangle - \langle \nabla_Y Z, \nabla_X Z \rangle, \\ \langle \nabla_Y \nabla_X Z, Z \rangle &= Y \langle \nabla_X Z, Z \rangle - \langle \nabla_X Z, \nabla_Y Z \rangle, \\ \langle \nabla_{[X,Y]} Z, Z \rangle &= \frac{1}{2} [X, Y] \langle Z, Z \rangle.\end{aligned}$$

Logo, novamente pela compatibilidade da métrica,

$$\begin{aligned}R(X, Y, Z, Z) &= X \langle \nabla_Y Z, Z \rangle - Y \langle \nabla_X Z, Z \rangle - \frac{1}{2} [X, Y] \langle Z, Z \rangle \\ &= \frac{1}{2} X (Y \langle Z, Z \rangle) - \frac{1}{2} Y (X \langle Z, Z \rangle) - \frac{1}{2} [X, Y] \langle Z, Z \rangle \\ &= \frac{1}{2} [(XY - YX) \langle Z, Z \rangle - [X, Y] \langle Z, Z \rangle] \\ &= \frac{1}{2} [[X, Y] \langle Z, Z \rangle - [X, Y] \langle Z, Z \rangle] \\ &= 0.\end{aligned}$$

(d) A identidade de Bianchi para o tensor curvatura segue da identidade de Bianchi para o tensor endomorfismo curvatura.

(c) Segue da identidade de Bianchi aplicadas às quatro permutações cíclicas dos quatro vetores:

$$\begin{aligned}R(X, Y, Z, W) + R(Y, Z, X, W) + R(Z, X, Y, W) &= 0, \\ R(Y, Z, W, X) + R(Z, W, Y, X) + R(W, Y, Z, X) &= 0, \\ R(Z, W, X, Y) + R(W, X, Z, Y) + R(X, Z, W, Y) &= 0, \\ R(W, X, Y, Z) + R(X, Y, W, Z) + R(Y, W, X, Z) &= 0.\end{aligned}$$

Somamos estas 4 identidades. Aplicando (a) cancelamos todos os termos das primeiras duas colunas ((1, 1) cancela com (4, 2), (2, 1) cancela com (1, 2), (3, 1) cancela com (2, 2) e (4, 1) cancela com (3, 2)). Quanto aos termos da última coluna, segue de (a) e (b) que

$$\begin{aligned}R(Z, X, Y, W) &= R(X, Z, W, Y), \\ R(W, Y, Z, X) &= R(Y, W, X, Z),\end{aligned}$$

logo a soma das 4 identidades é

$$2R(Z, X, Y, W) + 2R(W, Y, Z, X) = 0.$$

Daí,

$$R(Z, X, Y, W) = -R(W, Y, Z, X) = R(Y, W, Z, X).$$

■

13.15 Corolário. Vale

$$R(X, X, Z, W) = R(X, Y, Z, Z) = 0.$$

Em termos dos componentes do tensor curvatura,

$$R_{iikl} = R_{ijkk} = 0. \quad (13.28)$$

Além disso, qualquer soma de três componentes do tensor curvatura obtidos através da permutação cíclica de três índices é igual a zero.

Prova: A validade das duas primeiras identidades foram vistas na demonstração da proposição; de qualquer modo, $R_{iikl} = 0$ segue da propriedade (a) e $R_{ijkk} = 0$ segue da propriedade (b). A prova da última afirmação é deixada como exercício (Exercício 13.40). ■

13.16 Corolário. *Se M é uma variedade métrica de dimensão 2, então das $2^4 = 16$ componentes do tensor curvatura, existem 12 componentes nulas e apenas uma componente independente.*

Prova: Pelo corolário temos

$$\begin{aligned} R_{1111} &= R_{1112} = R_{1121} = R_{1122} = 0, \\ R_{2211} &= R_{2212} = R_{2221} = R_{2222} = 0, \\ R_{1211} &= R_{2111} = 0, \\ R_{1222} &= R_{2122} = 0. \end{aligned}$$

As componentes potencialmente não nulas são apenas $R_{1212}, R_{1221}, R_{2112}, R_{2121}$. É possível escolher uma dentre estas quatro e escrever as três outras em função dela usando as relações de simetria. Por exemplo, escolhendo R_{1212} , temos

$$\begin{aligned} R_{1221} &= -R_{1212}, \\ R_{2112} &= -R_{1212}, \\ R_{2121} &= -R_{2112} = R_{1212}. \end{aligned}$$

Observe que a simetria da identidade de Bianchi não desempenha nenhum papel, porque no caso $n = 2$, como também no caso $n = 3$, ela é consequência de (a)-(c). De fato, como no máximo 3 índices são diferentes nestes casos, pelo menos um dos coeficientes na soma cíclica (d) será da forma R_{ijkk} e portanto nulo, logo ela se reduzirá a uma ou duas das propriedades (a)-(c). Basta ver isso no caso $k = l$:

$$R_{ijkk} + R_{jkik} + R_{kijk} = 0$$

é equivalente a ($R_{ijkk} = 0$)

$$R_{jkik} + R_{kijk} = 0$$

ou seja,

$$R_{jkik} = -R_{kijk},$$

que corresponde a aplicar (c) e depois (a). ■

As propriedades de simetria do endomorfismo curvatura reduzem um pouco o número de componentes a calcular, mas não tanto quanto o tensor curvatura, já que não tem tantas simetrias. Pela Proposição 13.6 temos imediatamente 8 componentes nulas

$$\begin{aligned} R_{111}^1 &= R_{111}^2 = R_{112}^1 = R_{112}^2 = 0, \\ R_{221}^1 &= R_{221}^2 = R_{222}^1 = R_{222}^2 = 0, \end{aligned}$$

e da propriedade (a) apenas 4 componentes independentes:

$$R_{kii}^l = -R_{iki}^l.$$

Basta então calcular $R_{121}^1, R_{121}^2, R_{122}^1, R_{122}^2$, já que

$$\begin{aligned} R_{211}^1 &= -R_{121}^1, \\ R_{122}^1 &= -R_{212}^1, \\ R_{211}^2 &= -R_{121}^2, \\ R_{122}^2 &= -R_{212}^2. \end{aligned}$$

13.17 Exemplo (Curvatura da Esfera). Considerando a esfera como uma superfície de revolução com parametrização

$$(x, y, z)(\phi, \theta) = (r \sen \phi \cos \theta, r \sen \phi \sen \theta, r \cos \phi),$$

vimos no Exemplo 12.7 que

$$\begin{aligned} g_{11} &= r^2, \\ g_{12} &= g_{21} = 0, \\ g_{22} &= r^2 \sen^2 \phi, \end{aligned}$$

e

$$\begin{aligned} \Gamma_{11}^1 &= \Gamma_{11}^2 = \Gamma_{12}^1 = \Gamma_{21}^1 = \Gamma_{22}^2 = 0, \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{\cos \phi}{\sen \phi}, \\ \Gamma_{22}^1 &= -\sen \phi \cos \phi. \end{aligned}$$

Primeiro calcularemos as componentes do tensor endomorfismo curvatura. Como observado antes, basta calcular as 4 componentes independentes $R_{121}^1, R_{212}^1, R_{121}^2, R_{212}^2$. Para isso usamos a fórmula

$$\begin{aligned} R_{ijk}^l &= \Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l + \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j} \\ &= \Gamma_{jk}^1 \Gamma_{i1}^l - \Gamma_{ik}^1 \Gamma_{j1}^l + \Gamma_{jk}^2 \Gamma_{i2}^l - \Gamma_{ik}^2 \Gamma_{j2}^l + \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j}. \end{aligned}$$

Fazendo $k = i$, segue que

$$R_{iji}^l = \Gamma_{ji}^1 \Gamma_{i1}^l - \Gamma_{ii}^1 \Gamma_{j1}^l + \Gamma_{ji}^2 \Gamma_{i2}^l - \Gamma_{ii}^2 \Gamma_{j2}^l + \frac{\partial \Gamma_{ji}^l}{\partial x^i} - \frac{\partial \Gamma_{ii}^l}{\partial x^j}.$$

Veremos que há apenas 2 componentes independentes não nulas.

No caso $l = 1$, temos que calcular R_{121}^1 e R_{212}^1 . O único símbolo de Christoffel Γ_{ij}^1 não nulo é Γ_{22}^1 e $\Gamma_{ii}^2 = 0$ para todo i , logo para $i \neq j$ temos

$$\begin{aligned} R_{iji}^1 &= \Gamma_{ji}^1 \Gamma_{i1}^1 - \Gamma_{ii}^1 \Gamma_{j1}^1 + \Gamma_{ji}^2 \Gamma_{i2}^1 - \Gamma_{ii}^2 \Gamma_{j2}^1 + \frac{\partial \Gamma_{ji}^1}{\partial x^i} - \frac{\partial \Gamma_{ii}^1}{\partial x^j} \\ &= \Gamma_{ji}^2 \Gamma_{i2}^1 - \frac{\partial \Gamma_{ii}^1}{\partial x^j}. \end{aligned}$$

Daí,

$$R_{121}^1 = \Gamma_{21}^2 \Gamma_{12}^1 - \frac{\partial \Gamma_{11}^1}{\partial \theta} = 0$$

e

$$\begin{aligned} R_{212}^1 &= \Gamma_{12}^2 \Gamma_{22}^1 - \frac{\partial \Gamma_{22}^1}{\partial \phi} \\ &= \frac{\cos \phi}{\sen \phi} (-\sen \phi \cos \phi) - \frac{\partial (-\sen \phi \cos \phi)}{\partial \phi} \\ &= -\cos^2 \phi + \cos^2 \phi - \sen^2 \phi \\ &= -\sen^2 \phi. \end{aligned}$$

No caso $l = 2$, temos que calcular R_{121}^2 e R_{212}^2 . Para $i \neq j$ temos

$$\begin{aligned} R_{iji}^l &= \Gamma_{ji}^1 \Gamma_{i1}^2 - \Gamma_{ii}^1 \Gamma_{j1}^2 + \Gamma_{ji}^2 \Gamma_{i2}^2 - \Gamma_{ii}^2 \Gamma_{j2}^2 + \frac{\partial \Gamma_{ji}^2}{\partial x^i} - \frac{\partial \Gamma_{ii}^2}{\partial x^j} \\ &= \Gamma_{ji}^2 \Gamma_{i2}^2 - \Gamma_{ii}^1 \Gamma_{j1}^2 + \frac{\partial \Gamma_{ji}^2}{\partial x^i}. \end{aligned}$$

Daí,

$$\begin{aligned} R_{121}^2 &= \Gamma_{21}^2 \Gamma_{12}^2 - \Gamma_{11}^1 \Gamma_{21}^2 + \frac{\partial \Gamma_{21}^2}{\partial \phi} \\ &= \frac{\cos \phi \cos \phi}{\sin \phi \sin \phi} - 0 + \frac{\partial}{\partial \phi} \left(\frac{\cos \phi}{\sin \phi} \right) \\ &= \frac{\cos^2 \phi}{\sin^2 \phi} + \frac{-\sin^2 \phi - \cos^2 \phi}{\sin^2 \phi} \\ &= -1 \end{aligned}$$

e

$$R_{212}^2 = \Gamma_{12}^2 \Gamma_{22}^2 - \Gamma_{22}^1 \Gamma_{11}^2 + \frac{\partial \Gamma_{12}^2}{\partial \theta} = 0.$$

Portanto,

$$\begin{aligned} R_{212}^1 &= -R_{122}^1 = -\sin^2 \phi, \\ R_{121}^2 &= -R_{211}^2 = -1, \end{aligned}$$

enquanto que as demais 12 componentes são todas nulas.

Para calcular as $2^4 = 16$ componentes do tensor curvatura da esfera, pelo Corolário 13.16 basta calcular uma componente não nula. Como

$$R_{ijkl} = g_{ml} R_{ijk}^m = g_{1l} R_{ijk}^1 + g_{2l} R_{ijk}^2$$

segue que

$$R_{1212} = g_{12} R_{121}^1 + g_{22} R_{121}^2 = -r^2 \sin^2 \phi.$$

Portanto, as 4 componentes não nulas do tensor curvatura da esfera são

$$\begin{aligned} R_{1212} &= -r^2 \sin^2 \phi, \\ R_{1221} &= -R_{1212} = r^2 \sin^2 \phi, \\ R_{2112} &= -R_{1212} = r^2 \sin^2 \phi, \\ R_{2121} &= R_{1212} = -r^2 \sin^2 \phi. \end{aligned}$$

□

13.18 Exemplo (Curvatura do Plano Hiperbólico). Como vimos no Exemplo 12.9, temos

$$\begin{aligned} g_{11} &= g_{22} = \frac{1}{y^2}, \\ g_{12} &= g_{21} = 0, \end{aligned}$$

e

$$\begin{aligned} \Gamma_{11}^1 &= \Gamma_{22}^1 = \Gamma_{12}^2 = \Gamma_{21}^2 = 0, \\ \Gamma_{11}^2 &= \frac{1}{y}, \\ \Gamma_{12}^1 &= \Gamma_{21}^1 = \Gamma_{22}^2 = -\frac{1}{y}. \end{aligned}$$

Primeiro calcularemos as componentes do tensor endomorfismo curvatura:

$$R_{iji}^l = \Gamma_{ji}^1 \Gamma_{i1}^l - \Gamma_{ii}^1 \Gamma_{j1}^l + \Gamma_{ji}^2 \Gamma_{i2}^l - \Gamma_{ii}^2 \Gamma_{j2}^l + \frac{\partial \Gamma_{ji}^l}{\partial x^i} - \frac{\partial \Gamma_{ii}^l}{\partial x^j}.$$

No caso $l = 1$, para $i \neq j$,

$$\begin{aligned} R_{iji}^1 &= \Gamma_{ji}^1 \Gamma_{i1}^1 - \Gamma_{ii}^1 \Gamma_{j1}^1 + \Gamma_{ji}^2 \Gamma_{i2}^1 - \Gamma_{ii}^2 \Gamma_{j2}^1 + \frac{\partial \Gamma_{ji}^1}{\partial x^i} - \frac{\partial \Gamma_{ii}^1}{\partial x^j} \\ &= \Gamma_{ji}^1 \Gamma_{i1}^1 - \Gamma_{ii}^2 \Gamma_{j2}^1 + \frac{\partial \Gamma_{ji}^1}{\partial x^i}. \end{aligned}$$

Daí

$$R_{121}^1 = \Gamma_{21}^1 \Gamma_{11}^1 - \Gamma_{11}^2 \Gamma_{22}^1 + \frac{\partial \Gamma_{21}^1}{\partial x} = 0$$

e

$$\begin{aligned} R_{212}^1 &= \Gamma_{12}^1 \Gamma_{21}^1 - \Gamma_{22}^2 \Gamma_{12}^1 + \frac{\partial \Gamma_{12}^1}{\partial y} \\ &= \frac{1}{y^2} - \frac{1}{y^2} + \frac{1}{y^2} = \frac{1}{y^2}. \end{aligned}$$

No caso $l = 2$, para $i \neq j$,

$$\begin{aligned} R_{iji}^2 &= \Gamma_{ji}^1 \Gamma_{i1}^2 - \Gamma_{ii}^1 \Gamma_{j1}^2 + \Gamma_{ji}^2 \Gamma_{i2}^2 - \Gamma_{ii}^2 \Gamma_{j2}^2 + \frac{\partial \Gamma_{ji}^2}{\partial x^i} - \frac{\partial \Gamma_{ii}^2}{\partial x^j} \\ &= \Gamma_{ji}^1 \Gamma_{i1}^2 - \Gamma_{ii}^1 \Gamma_{j1}^2 + \Gamma_{ji}^2 \Gamma_{i2}^2 - \Gamma_{ii}^2 \Gamma_{j2}^2 - \frac{\partial \Gamma_{ii}^2}{\partial x^j}. \end{aligned}$$

Daí,

$$\begin{aligned} R_{121}^2 &= \Gamma_{21}^1 \Gamma_{11}^2 - \Gamma_{11}^1 \Gamma_{21}^2 + \Gamma_{21}^2 \Gamma_{12}^2 - \Gamma_{11}^2 \Gamma_{22}^2 - \frac{\partial \Gamma_{11}^2}{\partial y} \\ &= -\frac{1}{y^2} + 0 - 0 + \frac{1}{y^2} + \frac{1}{y^2} \\ &= \frac{1}{y^2} \end{aligned}$$

e

$$R_{212}^2 = \Gamma_{12}^1 \Gamma_{21}^2 - \Gamma_{22}^1 \Gamma_{11}^2 + \Gamma_{12}^2 \Gamma_{22}^2 - \Gamma_{22}^2 \Gamma_{12}^2 - \frac{\partial \Gamma_{22}^2}{\partial x} = 0.$$

Portanto,

$$R_{122}^1 = -R_{212}^1 = -R_{121}^2 = R_{211}^2 = -\frac{1}{y^2}$$

e as demais 12 componentes são todas nulas.

Para calcular as componentes do tensor curvatura do plano hiperbólico, basta calcular a componente não nula

$$R_{1212} = g_{12} R_{121}^1 + g_{22} R_{121}^2 = \frac{1}{y^4}.$$

Portanto, as 4 componentes não nulas do tensor curvatura do plano hiperbólico são

$$\begin{aligned} R_{1212} &= \frac{1}{y^4}, \\ R_{1221} &= -R_{1212} = -\frac{1}{y^4}, \\ R_{2112} &= -R_{1212} = -\frac{1}{y^4}, \\ R_{2121} &= R_{1212} = \frac{1}{y^4}. \end{aligned}$$

□

13.19 Corolário. *Se M é uma variedade métrica de dimensão n , então das n^4 componentes do tensor curvatura, existem*

$$\frac{n^2(n^2-1)}{12} \quad (13.29)$$

componentes independentes.

Prova: Pela Proposição 13.14 (c), temos

$$R_{ijkl} = R_{klji}.$$

Portanto, denotando

$$p = ij, q = kl,$$

podemos considerar R como uma matriz simétrica

$$R_{pq} = R_{qp},$$

cada índice p, q compreendendo uma matriz antisimétrica, pois pela Proposição 13.14 (a), (b), temos

$$\begin{aligned} R_{jikl} &= -R_{ijkl}, \\ R_{ijlk} &= -R_{ijkl}. \end{aligned}$$

A dimensão do subespaço das matrizes simétricas $n \times n$, ou seja, o número de componentes independentes de uma matriz simétrica $n \times n$, é

$$\frac{n(n+1)}{2},$$

enquanto que a dimensão do subespaço das matrizes antisimétricas $n \times n$, isto é, o número de componentes independentes de uma matriz antisimétrica $n \times n$, é

$$\frac{n(n-1)}{2}.$$

Portanto, R_{pq} vista como uma matriz simétrica tem $m(m+1)/2$ componentes independentes com $m = n(n-1)/2$ componentes independentes. Logo, sem levar em conta a simetria da identidade de Bianchi temos

$$\frac{\frac{n(n-1)}{2} \left(\frac{n(n-1)}{2} + 1 \right)}{2} = \frac{n^2(n-1)^2 + 2n(n-1)}{8} = \frac{n^4 - 2n^3 + 3n^2 - 2n}{8}.$$

componentes independentes. A identidade de Bianchi não desempenha um papel para $n = 2$ e $n = 3$, como vimos no fim da demonstração do Corolário 13.16, e note que

$$\begin{aligned} \frac{n^4 - 2n^3 + 3n^2 - 2n}{8} \Big|_{n=2} &= \frac{n^2(n^2-1)}{12} \Big|_{n=2} = 1, \\ \frac{n^4 - 2n^3 + 3n^2 - 2n}{8} \Big|_{n=3} &= \frac{n^2(n^2-1)}{12} \Big|_{n=3} = 6. \end{aligned}$$

A partir de $n \geq 4$, identidade de Bianchi reduz o número de componentes independentes. Para ver isso de forma concreta no caso $n = 4$, primeiro observe que seguindo o raciocínio acima restam apenas

$$\left. \frac{n^2(n^2 - 1)}{12} \right|_{n=3} = 21$$

componentes independentes, que são exatamente as componentes da parte triangular superior ou inferior da matriz

$\begin{matrix} kl \\ ij \downarrow \end{matrix}$	12	13	14	23	24	34
12	R_{1212}	R_{1213}	R_{1214}	R_{1223}	R_{1224}	R_{1234}
13	R_{1312}	R_{1313}	R_{1314}	R_{1323}	R_{1324}	R_{1334}
14	R_{1412}	R_{1413}	R_{1414}	R_{1423}	R_{1424}	R_{1434}
23	R_{2312}	R_{2313}	R_{2314}	R_{2323}	R_{2324}	R_{2334}
24	R_{2412}	R_{2413}	R_{2414}	R_{2423}	R_{2424}	R_{2434}
34	R_{3412}	R_{3413}	R_{3414}	R_{3423}	R_{3424}	R_{3434}

No caso $n = 4$ a identidade de Bianchi desempenha um papel, reduzindo de 21 para 20 o número de componentes independentes. De fato, como visto no final da demonstração do Corolário 13.16, qualquer nova condição deve envolver 4 índices distintos. As componentes com 4 índices diferentes ocupam a antidiagonal da matriz acima, logo só há 3 componentes independentes com 4 índices diferentes; por exemplo, na parte triangular superior da matriz acima estas componentes são: R_{1234} , $R_{1423} = R_{2314}$ e $R_{1324} = -R_{3124}$. A identidade de Bianchi

$$R_{1234} + R_{2314} + R_{3124} = 0$$

permite escrever

$$R_{1234} = -R_{2314} - R_{3124},$$

o que elimina R_{1234} como componente independente.

No caso geral, a identidade de Bianchi elimina $\binom{n}{4}$ componentes independentes, já que este é o número de modos que podemos escolher 4 índices diferentes dentre n índices e a identidade de Bianchi elimina um destes (no caso $n = 5$, a identidade de Bianchi eliminará R_{1234} , R_{1345} , R_{1245} , R_{1235} , R_{2345} como componentes independentes). Portanto, o número total final de componentes independentes é

$$\begin{aligned} \frac{n^4 - 2n^3 + 3n^2 - 2n}{8} - \binom{n}{4} &= \frac{n^4 - 2n^3 + 3n^2 - 2n}{8} - \frac{n!}{(n-4)!4!} \\ &= \frac{n^4 - 2n^3 + 3n^2 - 2n}{8} - \frac{n(n-1)(n-2)(n-3)}{24} \\ &= \frac{3n^4 - 6n^3 + 9n^2 - 6n - (n^4 - 6n^3 + 11n^2 - 6n)}{24} \\ &= \frac{2n^4 - 2n^2}{24} \\ &= \frac{n^2(n^2 - 1)}{12}. \end{aligned}$$

■ Assim, existem 6 componentes independentes para o tensor curvatura de variedades de dimensão 3 dentre $3^4 = 81$ componentes, e 20 componentes independentes para o tensor curvatura de variedades de dimensão 4 dentre $4^4 = 256$ componentes, como as que são estudadas em relatividade geral.

13.3 Curvatura de Ricci e Curvatura Escalar

Porque 4-tensores são tão complexos, como vimos na seção anterior, é frequentemente útil construir tensores mais simples que resumem *alguma* informação contida no tensor curvatura. Obviamente, ao simplificar o

tensor curvatura estamos perdendo informação: os tensores mais simples contêm *menos* informação que o tensor original. O mais importante dentre estes tensores é o *tensor de Ricci*.

13.20 Definição. Seja M uma variedade métrica. O **tensor curvatura de Ricci** de M (ou simplesmente **tensor de Ricci**) denotado Ric , é o campo tensorial covariante de ordem 2 definido como o *traço* do tensor endomorfismo curvatura em relação ao seu primeiro índice covariante e seu único índice contravariante ou, equivalentemente, como o *traço em relação à métrica* do tensor curvatura no seu primeiro e último índices. Portanto, os componentes da curvatura de Ricci são dados por

$$R_{ij} = R_{kij}^k = g^{km} R_{kijm}. \quad (13.30)$$

□

Pelas simetrias do tensor endomorfismo curvatura, usar traços diferentes não faria diferença ou apenas implicaria em uma troca de sinal:

$$R_{ij} = \sum_{k=1}^n R_{kij}^k = - \sum_{k=1}^n R_{ikj}^k = \sum_{k=1}^n R_{ijk}^k.$$

13.21 Proposição. A curvatura de Ricci é um campo tensorial simétrico, isto é,

$$R_{ij} = R_{ji}. \quad (13.31)$$

Prova: Temos

$$R_{ij} = g^{km} R_{kijm} = g^{km} R_{jmki} = g^{mk} R_{mjik} = R_{ji}.$$

■

O segundo tensor é a curvatura escalar, que é uma função real.

13.22 Definição. Seja M uma variedade métrica. A **curvatura escalar** de M , denotada S , é a função real $S : M \rightarrow \mathbb{R}$ definida como o traço em relação à métrica do tensor de Ricci:

$$S = \text{tr}_g \text{Ric} = g^{ij} R_{ij}. \quad (13.32)$$

□

13.23 Exemplo (Curvatura de Ricci e Curvatura Escalar da Esfera). Para variedades métricas de dimensão 2 a curvatura de Ricci é

$$R_{ij} = R_{kij}^k = R_{1ij}^1 + R_{2ij}^2.$$

Referindo aos cálculos efetuados no Exemplo 13.17, segue que a curvatura de Ricci da esfera é

$$\begin{aligned} R_{11} &= R_{111}^1 + R_{211}^2 = 1, \\ R_{12} &= R_{21} = R_{112}^1 + R_{212}^2 = 0, \\ R_{22} &= R_{122}^1 + R_{222}^2 = \sin^2 \phi. \end{aligned}$$

Para variedades riemannianas de dimensão 2 a curvatura escalar é

$$S = g^{ij} R_{ij} = g^{11} R_{11} + 2g^{12} R_{12} + g^{22} R_{22}.$$

Logo, a curvatura escalar da esfera de raio r é

$$S = \frac{1}{r^2} + \frac{\sin^2 \phi}{r^2 \sin^2 \phi} = \frac{2}{r^2}.$$

Ou seja, a curvatura escalar da esfera tende a 0 (curvatura escalar do plano euclideano) quando $r \rightarrow \infty$. □

13.24 Exemplo (Curvatura de Ricci e Curvatura Escalar do Plano Hiperbólico). Referindo aos cálculos efetuados no Exemplo 13.18, segue que

$$\begin{aligned} R_{11} &= R_{111}^1 + R_{211}^2 = -\frac{1}{y^2}, \\ R_{12} &= R_{21} = R_{112}^1 + R_{212}^2 = 0, \\ R_{22} &= R_{122}^1 + R_{222}^2 = -\frac{1}{y^2}. \end{aligned}$$

Logo, a curvatura escalar do plano hiperbólico é

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

□

Como o tensor métrica e a curvatura de Ricci são ambos tensores 2-covariantes simétricos, é natural se perguntar se existe uma relação mais direta entre eles.

13.25 Definição. Uma variedade métrica M é chamada uma **variedade de Einstein** se existe uma função $f : M \rightarrow \mathbb{R}$ tal que

$$\text{Ric} = fg. \quad (13.33)$$

Neste caso, dizemos também que g é uma **métrica de Einstein**. □

Pela Proposição 13.32, variedades de curvatura seccional constante (esta noção de curvatura será definida na próxima seção) são variedades de Einstein. Observe que como

$$S = \text{tr}_g \text{Ric} = f \text{tr}_g g = nf,$$

pois

$$\text{tr}_g g = g^{jk} g_{jk} = g^{jk} g_{kj} = \delta_j^j = n,$$

segue que

$$f = \frac{S}{n}.$$

Logo,

13.26 Proposição. M é uma variedade de Einstein se e somente se

$$\text{Ric} = \frac{S}{n}g. \quad (13.34)$$

O nome *variedades de Einstein* vem da equação de Einstein (em unidades geometrizadas $G = c = 1$) que como deduziremos mais tarde é

$$\text{Ric} - \frac{1}{2}Sg + \Lambda g = 8\pi T,$$

onde T é um 2-tensor simétrico (o *tensor de momento-energia*) que descreve a densidade e fluxo do momento e energia presentes em cada ponto do espaço-tempo e Λ é a *constante cosmológica*. No vácuo, $T = 0$ e a equação de Einstein se torna

$$\text{Ric} - \frac{1}{2}Sg + \Lambda g = 0$$

(e assim, a constante cosmológica mede a densidade de energia do vácuo). Tomando o traço desta equação com respeito à métrica, como $\text{tr}_g \text{Ric} = S$ e $\text{tr}_g g = n = 4$ para o espaço-tempo 4-dimensional, segue que

$$S - 2S + 4\Lambda = 0$$

ou seja,

$$S = 4\Lambda$$

(curvatura escalar do vácuo). Segue que a equação de Einstein para o vácuo é

$$\text{Ric} = \Lambda g,$$

ou seja, a métrica do vácuo é uma métrica de Einstein no sentido da Definição 13.33. Vale a pena observar que Einstein considerou como o seu maior erro a introdução da constante cosmológica na sua equação da relatividade geral. Ele a introduziu para poder produzir um universo estático; quando Hubble observou a expansão do universo através do afastamento mútuo das galáxias, Einstein removeu a constante cosmológica da sua equação (o que equivale efetivamente a considerar $\Lambda = 0$). Quando no final da década de 90 foi observada a aceleração da expansão do universo, ela foi reintroduzida, com valor positivo a ser determinado através de experimentos (correntemente, o valor é bem próximo de zero).

Matematicamente, o interesse em métricas de Einstein provém de um resultado provado por Hilbert: métricas de Einstein são pontos críticos para o funcional curvatura escalar

$$\mathcal{S}(g) = \int_M S dV$$

no espaço de todas as métricas de M com volume constante.

13.4 Curvatura Seccional

13.27 Definição. Dado um espaço vetorial real V com produto interno $\langle \cdot, \cdot \rangle$ e vetores $v, w \in V$ linearmente independentes, a **área do paralelogramo** determinado por v, w é

$$|v \wedge w| = \sqrt{\|v\|^2 \|w\|^2 - \langle v, w \rangle^2}. \quad (13.35)$$

□

13.28 Proposição. $v, w \in V$ são linearmente independentes se e somente se

$$|v \wedge w| \neq 0.$$

Prova: Mostraremos que $|v \wedge w| = 0$ se e somente se v e w são linearmente dependentes. Se $w = \alpha v$, então

$$|v \wedge w| = \sqrt{\|v\|^2 \|\alpha v\|^2 - \langle v, \alpha v \rangle^2} = \alpha \sqrt{\|v\|^4 - \langle v, v \rangle^2} = 0.$$

Reciprocamente, se $|v \wedge w| = 0$ implica que vale a igualdade na desigualdade de Cauchy-Shwarz:

$$\langle v, w \rangle = \|v\| \|w\|.$$

Assuma $w \neq 0$, caso contrário v e w são automaticamente linearmente dependentes e não há nada que precise ser provado. Considere o vetor

$$z = v - \frac{\langle v, w \rangle}{\|w\|^2} w.$$

Temos

$$\begin{aligned}
\|z\|^2 &= \langle z, z \rangle = \left\langle v - \frac{\langle v, w \rangle}{\|w\|^2} w, v - \frac{\langle v, w \rangle}{\|w\|^2} w \right\rangle \\
&= \langle v, v \rangle - \left\langle v, \frac{\langle v, w \rangle}{\|w\|^2} w \right\rangle - \left\langle \frac{\langle v, w \rangle}{\|w\|^2} w, v \right\rangle + \left\langle \frac{\langle v, w \rangle}{\|w\|^2} w, \frac{\langle v, w \rangle}{\|w\|^2} w \right\rangle \\
&= \|v\|^2 - \frac{\langle v, w \rangle^2}{\|w\|^2} - \frac{\langle v, w \rangle^2}{\|w\|^2} + \frac{\langle v, w \rangle^2 \|w\|^2}{\|w\|^4} \\
&= \|v\|^2 - \|v\|^2 - \|v\|^2 + \|v\|^2 \\
&= 0.
\end{aligned}$$

logo

$$v = \frac{\langle v, w \rangle}{\|w\|^2} w.$$

■

13.29 Definição. Seja M uma variedade métrica e σ um plano de $T_p M$. A **curvatura seccional** de M associada a σ é definida por

$$K(X, Y) = \frac{R(X, Y, Y, X)}{|X \wedge Y|^2} = \frac{R(X, Y, Y, X)}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2}. \quad (13.36)$$

onde $X, Y \in T_p M$ são quaisquer vetores que formam uma base para σ . □

A curvatura seccional está bem definida, isto é, independe da base escolhida para o plano σ . De fato, sejam $\{X, Y\}$ e $\{Z, W\}$ duas bases para σ , de modo que existe uma matriz invertível $A = (a_{ij})_{2 \times 2}$ tal que

$$\begin{aligned}
X &= a_{11}Z + a_{21}W, \\
Y &= a_{12}Z + a_{22}W.
\end{aligned}$$

Temos, usando as simetrias do tensor curvatura (em primeiro lugar, o fato que $R(Z, Z, \cdot, \cdot) = R(W, W, \cdot, \cdot) = R(\cdot, \cdot, Z, Z) = R(\cdot, \cdot, W, W) = 0$),

$$\begin{aligned}
R(X, Y, Y, X) &= R(a_{11}Z + a_{21}W, a_{12}Z + a_{22}W, a_{12}Z + a_{22}W, a_{11}Z + a_{21}W) \\
&= R(a_{11}Z, a_{22}W, a_{22}W, a_{11}Z) \\
&\quad + R(a_{11}Z, a_{22}W, a_{12}Z, a_{21}W) \\
&\quad + R(a_{21}W, a_{12}Z, a_{12}Z, a_{21}W) \\
&\quad + R(a_{21}W, a_{12}Z, a_{22}W, a_{11}Z) \\
&= a_{11}^2 a_{12}^2 R(Z, W, W, Z) + a_{11} a_{21} a_{12} a_{22} R(Z, W, Z, W) \\
&\quad + a_{21}^2 a_{22}^2 R(W, Z, Z, W) + a_{11} a_{21} a_{12} a_{22} R(W, Z, W, Z) \\
&= (a_{11}^2 a_{22}^2 - 2a_{11} a_{12} a_{21} a_{22} + a_{12}^2 a_{21}^2) R(Z, W, W, Z) \\
&= (\det A)^2 R(Z, W, W, Z),
\end{aligned}$$

e

$$\begin{aligned}
& |X \wedge Y| \\
&= \|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2 \\
&= \|a_{11}Z + a_{21}W\|^2 \|a_{12}Z + a_{22}W\|^2 - \langle a_{11}Z + a_{21}W, a_{12}Z + a_{22}W \rangle^2 \\
&= \left(a_{11}^2 \|Z\|^2 + 2a_{11}a_{21} \langle Z, W \rangle + a_{21}^2 \|W\|^2 \right) \left(a_{12}^2 \|Z\|^2 + 2a_{12}a_{22} \langle Z, W \rangle + a_{22}^2 \|W\|^2 \right) \\
&\quad - \left[a_{11}a_{12} \|Z\|^2 + (a_{11}a_{22} + a_{21}a_{12}) \langle Z, W \rangle + a_{21}a_{22} \|W\|^2 \right]^2 \\
&= a_{11}^2 a_{12}^2 \|Z\|^4 + 4(a_{11}a_{21}a_{12}a_{22}) \langle Z, W \rangle^2 + a_{21}^2 a_{22}^2 \|W\|^4 + (a_{11}^2 a_{22}^2 + a_{21}^2 a_{12}^2) \|Z\|^2 \|W\|^2 \\
&\quad + 2(a_{11}^2 a_{12}a_{22} + a_{12}^2 a_{11}a_{21}) \|Z\|^2 \langle Z, W \rangle + 2(a_{21}^2 a_{12}a_{22} + a_{22}^2 a_{11}a_{21}) \|W\|^2 \langle Z, W \rangle \\
&\quad - a_{11}^2 a_{12}^2 \|Z\|^4 - (a_{11}^2 a_{22}^2 + 2a_{11}a_{12}a_{21}a_{22} + a_{12}^2 a_{21}^2) \langle Z, W \rangle^2 - a_{21}^2 a_{22}^2 \|W\|^4 - 2(a_{11}a_{12}a_{21}a_{22}) \|Z\|^2 \|W\|^2 \\
&\quad - 2(a_{11}^2 a_{12}a_{22} + a_{12}^2 a_{11}a_{21}) \|Z\|^2 \langle Z, W \rangle - 2(a_{21}^2 a_{12}a_{22} + a_{22}^2 a_{11}a_{21}) \|W\|^2 \langle Z, W \rangle \\
&= (a_{11}^2 a_{22}^2 - 2a_{11}a_{12}a_{21}a_{22} + a_{12}^2 a_{21}^2) [\|Z\|^2 \|W\|^2 - \langle Z, W \rangle^2] \\
&= (\det A)^2 |Z \wedge W|.
\end{aligned}$$

Logo,

$$\frac{R(X, Y, Y, X)}{|X \wedge Y|} = \frac{R(Z, W, W, Z)}{|Z \wedge W|}.$$

O conhecimento da curvatura seccional para todos os planos σ determina o tensor curvatura:

13.30 Proposição. *Se R_1 e R_2 são dois tensores covariantes de ordem 4 em um espaço vetorial V que satisfazem as propriedades de simetria listadas na Proposição 13.14 tais que*

$$K_1(X, Y) = K_2(X, Y)$$

para todos vetores linearmente independentes $X, Y \in V$, então

$$R_1 = R_2.$$

Prova: Por hipótese, para quaisquer vetores $X, Y \in V$,

$$R_1(X, Y, Y, X) = R_2(X, Y, Y, X) \tag{13.37}$$

(segue diretamente da definição no caso em que X, Y são linearmente independentes; se X, Y são linearmente dependentes, um é múltiplo escalar do outro e a identidade é trivial, já que cada lado é igual a zero pelas propriedades de simetria). Daí,

$$R_1(X + Z, Y, Y, X + Z) = R_2(X + Z, Y, Y, X + Z). \tag{13.38}$$

Como

$$\begin{aligned}
R_1(X + W, Y, Y, X + W) &= R_1(X, Y, Y, X) + R_1(X, Y, Y, W) \\
&\quad + R_1(W, Y, Y, X) + R_1(W, Y, Y, W) \\
&= R_1(X, Y, Y, X) + 2R_1(X, Y, Y, W) + R_1(W, Y, Y, W)
\end{aligned}$$

e analogamente para $R_2(X + W, Y, Y, X + W)$, concluímos que

$$R_1(X, Y, Y, W) = R_2(X, Y, Y, W). \tag{13.39}$$

Desta identidade segue que

$$R_1(X, Y + Z, Y + Z, W) = R_2(X, Y + Z, Y + Z, W). \quad (13.40)$$

Como

$$\begin{aligned} R_1(X, Y + Z, Y + Z, W) &= R_1(X, Y, Y, W) + R_1(X, Y, Z, W) \\ &\quad + R_1(X, Z, Y, W) + R_1(X, Z, Z, W) \end{aligned}$$

e analogamente para $R_2(X, Y + Z, Y + Z, W)$, segue que

$$R_1(X, Y, Z, W) + R_1(X, Z, Y, W) = R_2(X, Y, Z, W) + R_2(X, Z, Y, W). \quad (13.41)$$

Defina agora

$$R = R_1 - R_2. \quad (13.42)$$

Então R é um 4-tensor covariante que satisfaz as mesmas propriedades que R_1, R_2 satisfazem. A identidade (13.41) implica que

$$R(X, Y, Z, W) + R(X, Z, Y, W) = 0.$$

Portanto, R é antisimétrico com relação a qualquer par adjacente de índices: a relação acima prova a antisimetricidade de R para o par adjacente interno e das regras de simetria da Proposição 13.14 obtemos a antisimetricidade em relação aos pares adjacentes esquerdo e direito. Daí

$$\begin{aligned} R(Y, Z, X, W) &= -R(Y, X, Z, W) = R(X, Y, Z, W), \\ R(Z, X, Y, W) &= -R(X, Z, Y, W) = R(X, Y, Z, W). \end{aligned}$$

Da identidade de Bianchi

$$R(X, Y, Z, W) + R(Y, Z, X, W) + R(Z, X, Y, W) = 0$$

concluímos então que

$$3R(X, Y, Z, W) = 0$$

para todos os vetores X, Z, Y, W , o que implica que R é o tensor nulo e portanto $R_1 = R_2$. ■

Usando a curvatura seccional, pode-se dar interpretações geométricas para as curvaturas de Ricci e escalar. Seja X um vetor unitário em $T_p M$ e complete ele até uma base ortonormal $\{X, E_2, \dots, E_n\}$ de $T_p M$. Então

$$\begin{aligned} \text{Ric}(X, X) &= R_{11} = g^{km} R_{k11m} = \delta^{km} R(X, E_k, E_m, X) = \sum_{k=1}^n R(X, E_k, E_k, X) = \sum_{k=2}^n R(X, E_k, E_k, X) \\ &= \sum_{k=2}^n \frac{R(X, E_k, E_k, X)}{|X \wedge E_k|}, \end{aligned}$$

ou seja,

$$\text{Ric}(X, X) = \sum_{k=2}^n K(X, E_k). \quad (13.43)$$

Em outras palavras, *para cada vetor unitário $X \in T_p M$, $\text{Ric}(X, X)$ é a soma das curvaturas seccionais de planos gerados por X e os outros vetores de uma base ortonormal*. Por outro lado, como Ric é bilinear e

simétrico, ele é completamente determinado pelos seus valores $\text{Ric}(X, X)$ em vetores unitários X : se X, Y são vetores tais que $X + Y, X - Y \neq 0$,

$$\begin{aligned}\text{Ric}(X, Y) &= \frac{1}{4} (\text{Ric}(X + Y, X + Y) - \text{Ric}(X - Y, X - Y)) \\ &= \frac{1}{4} \left[\|X + Y\|^2 \text{Ric} \left(\frac{X + Y}{\|X + Y\|}, \frac{X + Y}{\|X + Y\|} \right) - \|X - Y\|^2 \text{Ric} \left(\frac{X - Y}{\|X - Y\|}, \frac{X - Y}{\|X - Y\|} \right) \right] \\ &= \frac{1}{4} \left[\|X + Y\|^2 \sum_{k=2}^n K(X + Y, E_k) - \|X - Y\|^2 \sum_{k=2}^n K(X - Y, F_k) \right],\end{aligned}$$

onde $\{X, E_2, \dots, E_n\}, \{X, F_2, \dots, F_n\}$ são bases ortonormais de $T_p M$; se $X + Y = 0$ ou $X - Y = 0$, o resultado continua valendo, já que o termo correspondente é nulo.

$$\text{Ric}(X, Y) = \|X\| \|Y\| \text{Ric} \left(\frac{X}{\|X\|}, \frac{Y}{\|Y\|} \right).$$

Similarmente, se $\{E_1, \dots, E_n\}$ é uma base ortonormal para $T_p M$, então a curvatura escalar é dada por

$$S = g^{ij} R_{ij} = \delta^{ij} R_{ij} = R_{ii} = \sum_{i=1}^n \text{Ric}(E_i, E_i),$$

ou seja,

$$S = \sum_{i \neq j} K(E_i, E_j). \quad (13.44)$$

Em outras palavras, a curvatura escalar é a soma de todas as curvaturas seccionais de planos gerados por pares de vetores de uma base ortonormal.

13.31 Exemplo (Curvaturas de Variedades de Dimensão 2). No caso de variedades de dimensão 2 existe apenas um plano em $T_p M$: o próprio. Para estas variedades, a curvatura seccional pode ser considerada uma função $K : M \rightarrow \mathbb{R}$. Além disso, se $\{E_1, E_2\}$ é uma base ortonormal para $T_p M$, segue da discussão anterior que

$$S(p) = 2K(E_1, E_2) = 2K(p),$$

ou seja, para variedades riemannianas de dimensão 2 a curvatura seccional é essencialmente a curvatura escalar:

$$S = 2K. \quad (13.45)$$

Além disso, temos

$$R(X, Y, Z, W) = K(p) [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle] \quad (13.46)$$

De fato, o único componente independente de R é

$$R_{1221} = R(E_1, E_2, E_2, E_1) = K(E_1, E_2) = K(p),$$

enquanto que os componentes não nulos são dados por

$$\begin{aligned}R_{1212} &= -R_{1221} = -K(p), \\ R_{2112} &= -R_{1212} = R_{1221} = K(p), \\ R_{2121} &= -R_{1221} = -K(p).\end{aligned}$$

Como a equação acima define um tensor 4-covariante de ordem 4 que tem as mesmas componentes que R , ele é R . Segue que

$$\text{Ric}(X, Y) = K(p) g(X, Y), \quad (13.47)$$

pois

$$\begin{aligned}\text{Ric}(X, Y) &= \frac{1}{4} \left[\|X + Y\|^2 \sum_{k=2}^n K(X + Y, E_k) - \|X - Y\|^2 \sum_{k=2}^n K(X - Y, F_k) \right] \\ &= \frac{1}{4} \left[\|X + Y\|^2 K(X + Y, E_2) - \|X - Y\|^2 K(X - Y, F_2) \right] \\ &= K(p) \frac{1}{4} \left[\|X + Y\|^2 - \|X - Y\|^2 \right] \\ &= K(p) g(X, Y).\end{aligned}$$

Em particular, *toda variedade de dimensão 2 é uma variedade de Einstein*:

$$\text{Ric} = \frac{S}{2} g. \quad (13.48)$$

□

13.5 Variedades Riemannianas de Curvatura Seccional Constante

Variedades riemannianas que possuem curvatura seccional constante desempenham papel fundamental em Geometria Riemanniana. Algumas das mais importantes são os *espaços modelo*.

13.32 Proposição. *Seja M uma variedade riemanniana com curvatura seccional constante $K_0 \in \mathbb{R}$. Então as curvaturas de M são dadas pelas fórmulas*

$$R(X, Y)Z = K_0 [\langle Y, Z \rangle X - \langle X, Z \rangle Y]. \quad (13.49)$$

$$R(X, Y, Z, W) = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle]. \quad (13.50)$$

$$\text{Ric} = (n - 1) K_0 g. \quad (13.51)$$

$$S = n(n - 1) K_0. \quad (13.52)$$

Reciprocamente, se (13.49) ou (13.50) valem, então a curvatura seccional é constante.

Em particular, variedades com curvatura seccional constante são variedades de Einstein.

Prova: Para provar (13.50) defina um 4-tensor covariante \tilde{R} em M por

$$\tilde{R}(X, Y, Z, W) = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle]. \quad (13.53)$$

Então \tilde{R} satisfaz todas as propriedades de simetria da Proposição 13.14. De fato,

$$\begin{aligned}\tilde{R}(Y, X, Z, W) &= K_0 [\langle Y, W \rangle \langle X, Z \rangle - \langle Y, Z \rangle \langle X, W \rangle] \\ &= -K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle] \\ &= \tilde{R}(X, Y, Z, W),\end{aligned}$$

$$\begin{aligned}\tilde{R}(X, Y, W, Z) &= K_0 [\langle X, Z \rangle \langle Y, W \rangle - \langle X, W \rangle \langle Y, Z \rangle] \\ &= -K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle] \\ &= \tilde{R}(X, Y, Z, W),\end{aligned}$$

$$\begin{aligned}\tilde{R}(Z, W, X, Y) &= K_0 [\langle Z, Y \rangle \langle W, X \rangle - \langle Z, X \rangle \langle W, Y \rangle] \\ &= K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle] \\ &= \tilde{R}(X, Y, Z, W)\end{aligned}$$

e

$$\begin{aligned} \tilde{R}(X, Y, Z, W) + \tilde{R}(Y, Z, X, W) + \tilde{R}(Z, X, Y, W) \\ = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle + \langle Y, W \rangle \langle Z, X \rangle - \langle Y, X \rangle \langle Z, W \rangle + \langle Z, W \rangle \langle X, Y \rangle - \langle Z, Y \rangle \langle X, W \rangle] \\ = 0. \end{aligned}$$

Como

$$\tilde{K}(X, Y) = \frac{\tilde{R}(X, Y, Y, X)}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2} = \frac{K_0 [\langle X, X \rangle \langle Y, Y \rangle - \langle X, Y \rangle \langle X, Y \rangle]}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2} = K_0, \quad (13.54)$$

segue da Proposição 13.30 que $R = \tilde{R}$ e a recíproca segue então de (13.54).

(13.49) é equivalente a (13.50): claramente (13.50) segue de (13.49) pela definição da curvatura a partir do endomorfismo curvatura; reciprocamente, se (13.50) vale, então

$$\langle R(X, Y) Z, W \rangle = R(X, Y, Z, W) = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle],$$

para todo vetor W e como também

$$\langle K_0 [\langle Y, Z \rangle X - \langle X, Z \rangle Y], W \rangle = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle]$$

para todo vetor W , segue que

$$R(X, Y) Z = K_0 [\langle Y, Z \rangle X - \langle X, Z \rangle Y].$$

Se X, Y são vetores não nulos, como vimos na discussão no final da seção anterior, temos

$$\begin{aligned} \text{Ric}(X, Y) &= \frac{1}{4} \left[\|X + Y\|^2 \sum_{k=2}^n K(X + Y, E_k) - \|X - Y\|^2 \sum_{k=2}^n K(X - Y, F_k) \right] \\ &= \frac{1}{4} \left[\|X + Y\|^2 (n - 1) K_0 - \|X - Y\|^2 (n - 1) K_0 \right] \\ &= (n - 1) K_0 \frac{1}{4} [\|X + Y\|^2 - \|X - Y\|^2] \\ &= (n - 1) K_0 \langle X, Y \rangle; \end{aligned}$$

se um dos vetores é nulo, o resultado vale trivialmente. Além disso,

$$S = \sum_{i \neq j} K(E_i, E_j) = (n^2 - n) K_0 = n(n - 1) K_0.$$

■

13.33 Corolário. *Seja M uma variedade riemanniana. Para cada ponto $p \in M$ escolha uma base ortonormal $\{E_1, E_2, \dots, E_n\}$ de $T_p M$ e defina*

$$R_{ijkl} = R(E_i, E_j, E_k, E_l).$$

Então M possui curvatura seccional constante K_0 se e somente se

$$\begin{aligned} R_{ijji} &= -R_{ijij} = K_0 \quad \text{para todo } i \neq j \text{ e} \\ R_{ijkl} &= 0 \text{ nos outros casos,} \end{aligned}$$

para todo $p \in M$.

Prova: Pela proposição, M possui curvatura seccional constante K_0 se e somente se

$$R(X, Y, Z, W) = K_0 [\langle X, W \rangle \langle Y, Z \rangle - \langle X, Z \rangle \langle Y, W \rangle].$$

Como um tensor é determinado pelos seus valores em uma base, esta equação é válida se e somente se

$$\begin{aligned} R(E_i, E_j, E_k, E_l) &= K_0 [\langle E_i, E_l \rangle \langle E_j, E_k \rangle - \langle E_i, E_k \rangle \langle E_j, E_l \rangle] \\ &= K_0 (\delta_{il}\delta_{jk} - \delta_{ik}\delta_{jl}) \\ &= \begin{cases} 1 & \text{se } i = l \text{ e } j = k \text{ e } i \neq j \\ -1 & \text{se } i = k \text{ e } j = l \text{ e } i \neq j \\ 0 & \text{caso contrário.} \end{cases}. \end{aligned}$$

■

Vemos que em uma variedade riemanniana de curvatura seccional constante o tensor curvatura pode ser expresso diretamente em função da métrica.

13.6 Apêndice: Motivação para a definição do tensor curvatura

13.6.1 Efeitos de Maré em Mecânica Newtoniana

Primeiro, analizaremos os efeitos de maré do ponto de vista Newtoniano. Considere uma família de trajetórias de partículas-teste

$$F(s, t) = (x^1(s, t), \dots, x^n(s, t)).$$

Aqui, $n = 3$. Pela segunda lei de Newton, a equação de movimento para cada trajetória (para cada s fixado) é dada por

$$\frac{\partial^2 F}{\partial t^2}(s, t) + \nabla U(s, t) = 0$$

onde U é o potencial gravitacional Newtoniano. Para calcular a aceleração relativa entre as trajetórias produzidas por forças de maré, calculamos

$$\frac{\partial}{\partial s} \left[\frac{\partial^2 F}{\partial t^2}(s, t) + \nabla U(s, t) \right] = 0.$$

Escrevendo em coordenadas

$$\frac{\partial^2 F}{\partial t^2}(s, t) = (x^1(s, t), \dots, x^n(s, t))$$

e

$$\nabla \phi(s, t) = \left(\frac{\partial U}{\partial x^1}(s, t), \dots, \frac{\partial U}{\partial x^n}(s, t) \right).$$

Segue que, comutando derivadas,

$$\frac{\partial}{\partial s} \frac{\partial^2 x^i}{\partial t^2} = \frac{\partial^2}{\partial t^2} \frac{\partial x^i}{\partial s}$$

enquanto que, pela regra da cadeia,

$$\frac{\partial}{\partial s} \frac{\partial U}{\partial x^i} = \sum_{j=1}^n \frac{\partial^2 U}{\partial x^i \partial x^j} \frac{\partial x^j}{\partial s}.$$

Assim, em coordenadas, obtemos

$$\frac{\partial^2}{\partial t^2} \frac{\partial x^i}{\partial s} + \sum_{j=1}^n \frac{\partial^2 U}{\partial x^i \partial x^j} \frac{\partial x^j}{\partial s} = 0$$

O termo

$$\frac{\partial^2}{\partial t^2} \left(\frac{\partial F}{\partial s} \right) = \frac{\partial^2}{\partial t^2} \left(\frac{\partial x^1}{\partial s}, \dots, \frac{\partial x^n}{\partial s} \right)$$

é a **aceleração relativa das trajetórias**, enquanto que o termo

$$\sum_{i,j=1}^n \frac{\partial^2 U}{\partial x^i \partial x^j} \frac{\partial x^j}{\partial s}$$

corresponde às **forças de gravidade que produzem as marés**.

13.6.2 Efeitos de Maré em Relatividade Geral: Curvatura como uma medida da aceleração relativa de trajetórias geodésicas

Agora analizaremos os efeitos de maré do ponto de vista de Relatividade Geral. Considere uma variação geodésica

$$F(s, t) : (-\varepsilon, \varepsilon) \times [0, 1] \longrightarrow M,$$

ou seja, para cada s fixado, a curva $F_s(t) = F(s, t)$ é uma geodésica de M ; um exemplo é a variação considerada na demonstração do Lema de Gauss (Lema 12.26). O vetor

$$\frac{\partial F}{\partial t}$$

é o vetor tangente em cada geodésica $F_s(t)$, que pode ser visto como a velocidade temporal de uma partícula hipotética ao longo desta trajetória geodésica. Como a trajetória é uma geodésica, temos

$$\frac{D}{dt} \frac{\partial F}{\partial t} = 0,$$

isto é, a aceleração em cada trajetória $F_s(t)$ da variação é nula. Já o vetor

$$\frac{\partial F}{\partial s}$$

mede a velocidade de afastamento, ou taxa de variação do espaço entre as trajetórias geodésicas (a magnitude deste vetor será efetivamente medida na Seção ??). Ou seja, enquanto enxergamos a variável t como uma variável temporal, isto é, $F_s(t)$ é a posição de uma partícula hipotética na trajetória geodésica F_s no instante de tempo t , a variável s é vista como uma variável espacial, medindo a separação entre duas trajetórias geodésicas. Queremos obter a *aceleração relativa das trajetórias no tempo*

$$\frac{D^2}{dt^2} \left(\frac{\partial F}{\partial s} \right) = \frac{D}{dt} \frac{D}{dt} \frac{\partial F}{\partial s},$$

ou seja, a aceleração temporal da taxa de variação espacial (afastamento) das trajetórias geodésicas. Em Relatividade Geral medir isso é importante, pois corresponde a medir os *efeitos gravitacionais de maré*. Em \mathbb{R}^n , como as derivadas parciais comutam, temos para uma variação geral, não necessariamente geodésica, que

$$\frac{\partial^2}{\partial t^2} \left(\frac{\partial F}{\partial s} \right) = \frac{\partial}{\partial s} \left(\frac{\partial^2 F}{\partial t^2} \right),$$

de modo que medir a taxa de variação espacial das acelerações temporais das partículas nas diferentes trajetórias é equivalente a medir a aceleração temporal da taxa de variação espacial das trajetórias; se, em particular, a variação é geodésica, ou seja, uma família de retas, temos

$$\frac{\partial^2}{\partial t^2} \left(\frac{\partial F}{\partial s} \right) = \frac{\partial}{\partial s} \left(\frac{\partial^2 F}{\partial t^2} \right) = 0,$$

ou seja, partículas seguindo ao longo de retas paralelas não sofrem aceleração uma em direção a outra, porque no final das contas as retas paralelas se mantêm à mesma distância uma da outra e $\partial F/\partial s = 0$. Mas em variedades riemannianas (não localmente isométricas a \mathbb{R}^n , como veremos mais tarde) isso não é verdade porque as derivadas covariantes $\frac{D}{dt}$ e $\frac{D}{ds}$ não comutam, de modo que, enquanto que

$$\frac{D}{ds} \left(\frac{D}{dt} \frac{\partial F}{\partial t} \right) = 0,$$

em geral temos

$$\frac{D^2}{dt^2} \left(\frac{\partial F}{\partial s} \right) \neq 0$$

e existe uma aceleração temporal mensurável entre as trajetórias geodésicas na variedade. De fato, a não comutatividade das derivadas covariantes pode ser medida pelo tensor curvatura:

13.34 Lema (Curvatura como medida da não comutatividade da Derivada Covariante). *Seja $F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ uma variação e V um campo vetorial ao longo de F . Então,*

$$\frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} = R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) V.$$

Prova: Escreva

$$F(s, t) = (x^1(s, t), \dots, x^n(s, t))$$

e

$$V = V^i(s, t) \partial_i$$

em uma vizinhança coordenada. Então

$$\frac{DV}{ds} = \frac{D}{ds} (V^i \partial_i) = \frac{\partial V^i}{\partial s} \partial_i + V^i \frac{D}{ds} \partial_i,$$

onde

$$\frac{D}{dt} \frac{DV}{ds} = \frac{\partial^2 V^i}{\partial s \partial t} \partial_i + \frac{\partial V^i}{\partial s} \frac{D}{dt} \partial_i + \frac{\partial V^i}{\partial t} \frac{D}{ds} \partial_i + V^i \frac{D}{dt} \frac{D}{ds} \partial_i.$$

Trocando a letra t pela letra s , também obtemos

$$\frac{D}{ds} \frac{DV}{dt} = \frac{\partial^2 V^i}{\partial t \partial s} \partial_i + \frac{\partial V^i}{\partial t} \frac{D}{ds} \partial_i + \frac{\partial V^i}{\partial s} \frac{D}{dt} \partial_i + V^i \frac{D}{ds} \frac{D}{dt} \partial_i.$$

Logo,

$$\frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} = V^i \left(\frac{D}{dt} \frac{D}{ds} - \frac{D}{ds} \frac{D}{dt} \right) \partial_i. \quad (13.55)$$

Como

$$\begin{aligned} \frac{\partial F}{\partial t} &= \frac{\partial x^j}{\partial t} \partial_j, \\ \frac{\partial F}{\partial s} &= \frac{\partial x^k}{\partial s} \partial_k, \end{aligned}$$

temos

$$\frac{D}{ds} \partial_i = \nabla_{\frac{\partial x^k}{\partial s} \partial_k} \partial_i = \frac{\partial x^k}{\partial s} \nabla_{\partial_k} \partial_i.$$

Daí,

$$\begin{aligned}
 \frac{D}{dt} \frac{D}{ds} \partial_i &= \frac{D}{dt} \left(\frac{\partial x^k}{\partial s} \nabla_{\partial_k} \partial_i \right) \\
 &= \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \frac{\partial x^k}{\partial s} \frac{D}{dt} (\nabla_{\partial_k} \partial_i) \\
 &= \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \frac{\partial x^k}{\partial s} \nabla_{\frac{\partial x^j}{\partial t} \partial_j} (\nabla_{\partial_k} \partial_i) \\
 &= \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \frac{\partial x^k}{\partial s} \frac{\partial x^j}{\partial t} \nabla_{\partial_j} \nabla_{\partial_k} \partial_i.
 \end{aligned}$$

Da mesma forma, trocando t por s e depois j por k ,

$$\begin{aligned}
 \frac{D}{ds} \frac{D}{dt} \partial_i &= \frac{\partial^2 x^k}{\partial t \partial s} \nabla_{\partial_k} \partial_i + \frac{\partial x^k}{\partial t} \frac{\partial x^j}{\partial s} \nabla_{\partial_j} \nabla_{\partial_k} \partial_i \\
 &= \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} \nabla_{\partial_k} \nabla_{\partial_j} \partial_i.
 \end{aligned}$$

Portanto,

$$\begin{aligned}
 \frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} &= V^i \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} (\nabla_{\partial_j} \nabla_{\partial_k} \partial_i - \nabla_{\partial_k} \nabla_{\partial_j} \partial_i) \\
 &= V^i \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} R(\partial_j, \partial_k) \partial_i
 \end{aligned}$$

pois

$$\nabla_{[\partial_j, \partial_k]} \partial_i = 0$$

já que $[\partial_j, \partial_k] = 0$. O fato de que R é um tensor permite agora escrever

$$\begin{aligned}
 \frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} &= R \left(\frac{\partial x^j}{\partial t} \partial_j, \frac{\partial x^k}{\partial s} \partial_k \right) V^i \partial_i \\
 &= R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) V.
 \end{aligned}$$



Observe que definir a curvatura simplesmente por

$$R(X, Y) Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z,$$

de modo que a curvatura mediria apenas e diretamente o grau de não comutatividade da derivada covariante, não é satisfatório, porque esta expressão não define um tensor. De fato, ela não é linear na terceira variável:

$$\begin{aligned}
 \nabla_X \nabla_Y (fZ) - \nabla_Y \nabla_X (fZ) &= \nabla_X (f \nabla_Y Z) - \nabla_X ((Yf) Z) - \nabla_Y (f \nabla_X Z) + \nabla_Y ((Xf) Z) \\
 &= f \nabla_X \nabla_Y Z + (Xf) \nabla_Y Z + (Yf) \nabla_X (Z) + X(Yf) Z \\
 &\quad - f \nabla_Y \nabla_X Z - (Yf) \nabla_X Z - (Xf) \nabla_Y Z - Y(Xf) Z \\
 &= f \nabla_X \nabla_Y Z - f \nabla_Y \nabla_X Z + [X(Yf) - Y(Xf)] Z \\
 &= f (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z) + [X, Y](f) Z.
 \end{aligned}$$

Para obter um tensor, precisamos definir a curvatura como

$$R(X, Y) Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z.$$

pois

$$\nabla_{[X,Y]}(fZ) = f\nabla_{[X,Y]}Z + [X,Y](f)Z$$

e o termo $[X,Y](f)Z$ que restou acima é eliminado, deixando $-f\nabla_{[X,Y]}Z$. A demonstração do Lema 13.34 não funciona se R não é um tensor.

Podemos agora calcular a aceleração temporal entre as geodésicas da variedade.

13.35 Proposição (Equação do Desvio Geodésico). *Seja $F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ uma variação geodésica. Então,*

$$\frac{D^2}{dt^2} \frac{\partial F}{\partial s} + R\left(\frac{\partial F}{\partial s}, \frac{\partial F}{\partial t}\right) \frac{\partial F}{\partial t} = 0$$

Prova: Como cada trajetória de F é uma geodésica, temos

$$\frac{D}{dt} \frac{\partial F}{\partial t} = 0,$$

onde

$$\frac{D}{ds} \frac{D}{dt} \frac{\partial F}{\partial t} = 0.$$

Invocando o Lema de Simetria 12.25 (cuja validade depende de que o tensor torção seja nulo), obtemos

$$\begin{aligned} 0 &= \frac{D}{ds} \left(\frac{D}{dt} \frac{\partial F}{\partial t} \right) = \frac{D}{dt} \left(\frac{D}{ds} \frac{\partial F}{\partial t} \right) - R\left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s}\right) \frac{\partial F}{\partial t} \\ &= \frac{D}{dt} \left(\frac{D}{dt} \frac{\partial F}{\partial s} \right) + R\left(\frac{\partial F}{\partial s}, \frac{\partial F}{\partial t}\right) \frac{\partial F}{\partial t} \\ &= \frac{D^2}{dt^2} \frac{\partial F}{\partial s} + R\left(\frac{\partial F}{\partial s}, \frac{\partial F}{\partial t}\right) \frac{\partial F}{\partial t}. \end{aligned}$$

■

A equação do desvio geodésico da Relatividade Geral é chamada **equação de Jacobi** em Geometria Riemanniana. Em Relatividade Geral, o termo da curvatura

$$R\left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s}\right) \frac{\partial F}{\partial t}$$

é interpretado como sendo o responsável pelas forças que produzem efeitos de maré, isto é, a curvatura do espaço-tempo é que causa os efeitos de maré, fazendo com que partículas-teste que seguem trajetórias geodésicas no espaço-tempo acelerem uma na direção da outra.

13.6.3 Curvatura como uma medida do transporte paralelo em trajetórias fechadas

Outro modo importante de entender o tensor curvatura é como uma medida (infinitesimal) de quanto o transporte paralelo em trajetórias fechadas deixa de retornar o mesmo vetor ou, equivalente, o quanto o transporte paralelo de um vetor de um ponto para outro próximo depende da trajetória que liga estes dois pontos. O enunciado e demonstração dados a seguir é de [Rossmann], Theorem 2.3.5, pp. 106–107 (veja também [MTW], Section 11.4, [Shutz], Section 6.5, pp. 157–161, [Benn-Tucker], Section 6.4, pp. 208–212 e [Dodson-Poston], pp. 309–310).

13.36 Definição. Uma **superfície parametrizada** é uma aplicação diferenciável

$$F : [a, b] \times [c, d] \rightarrow M.$$

Um **campo vetorial ao longo de F** é um campo diferenciável $V : [a, b] \times [c, d] \rightarrow TM$ tal que $V(s, t) \in T_{F(s,t)}M$ para cada (s, t) . □

Uma superfície parametrizada difere em poucos detalhes de uma variação. Em particular, o Lema 13.34 vale para uma superfície parametrizada.

13.37 Teorema (Efeito da Curvatura no Transporte Paralelo em uma Trajetória Fechada). *Seja $F : [s_0, S] \times [t_0, T] \rightarrow M$ uma superfície parametrizada e denote*

$$\begin{aligned} p &= F(s_0, t_0), \\ U &= \left. \frac{\partial F}{\partial s} \right|_{(s_0, t_0)}, \\ V &= \left. \frac{\partial F}{\partial t} \right|_{(s_0, t_0)}. \end{aligned}$$

Seja

$$\tau : T_p M \rightarrow T_p M$$

o transporte paralelo ao longo das curvas da fronteira desta superfície parametrizada, isto é,

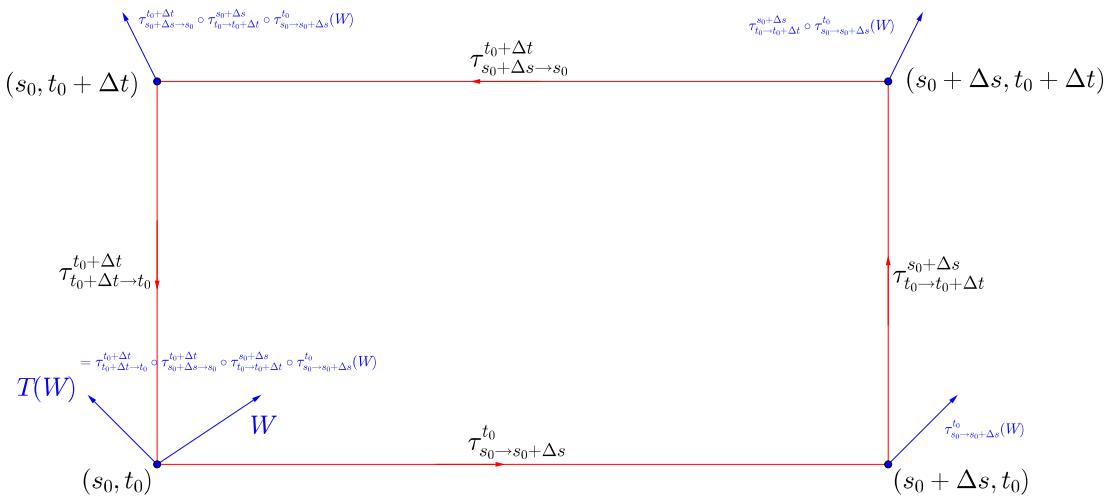
$$\tau = \tau_{t_0 + \Delta t \rightarrow t_0}^{s_0} \circ \tau_{s_0 + \Delta s \rightarrow s_0}^{t_0 + \Delta t} \circ \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} \circ \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0}$$

onde

- $\tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0}$ é o transporte paralelo ao longo da curva $F(s, t_0)$ de s_0 até $s_0 + \Delta s$,
- $\tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s}$ é o transporte paralelo ao longo da curva $F(s_0 + \Delta s, t)$ de t_0 até $t_0 + \Delta t$,
- $\tau_{s_0 + \Delta s \rightarrow s_0}^{t_0 + \Delta t}$ é o transporte paralelo ao longo da curva $F(s, t_0 + \Delta t)$ de $s_0 + \Delta s$ até s_0 ,
- $\tau_{t_0 + \Delta t \rightarrow t_0}^{s_0}$ é o transporte paralelo ao longo da curva $F(s_0, t)$ de $t_0 + \Delta t$ até t_0 .

Se $W \in T_p M$, então

$$R(U, V) W = \lim_{\Delta s, \Delta t \rightarrow 0} \frac{1}{\Delta s \Delta t} (W - \tau(W)).$$



Prova: Como

$$\begin{aligned}\tau_{s_0+\Delta s \rightarrow s_0}^{t_0+\Delta t} \circ \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} &= \text{id}, \\ \tau_{t_0+\Delta t \rightarrow t_0}^{s_0} \circ \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0} &= \text{id},\end{aligned}$$

podemos escrever

$$W - \tau(W) = \tau_{t_0+\Delta t \rightarrow t_0}^{s_0} \circ \tau_{s_0+\Delta s \rightarrow s_0}^{t_0+\Delta t} \circ \left[\tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \circ \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0}(W) - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \circ \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0}(W) \right].$$

Já que

$$\lim_{\Delta s, \Delta t \rightarrow 0} \left(\tau_{t_0+\Delta t \rightarrow t_0}^{s_0} \circ \tau_{s_0+\Delta s \rightarrow s_0}^{t_0+\Delta t} \right) = \text{id},$$

basta então provar que

$$R(U, V) W = \lim_{\Delta s, \Delta t \rightarrow 0} \frac{1}{\Delta s \Delta t} \left[\tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \circ \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0}(W) - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \circ \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0}(W) \right]. \quad (13.56)$$

Isso é equivalente a calcular a diferença entre os transportes paralelos do vetor W de $T_p M$ até $T_q M$, onde $q = F(t_0 + \Delta t, s_0 + \Delta s)$ é o vértice oposto do retângulo ao longo dos dois caminhos possíveis da fronteira do retângulo.

Pelo Lema 11.27 (substituindo a derivada covariante pela derivada covariante ao longo de uma curva e invertendo a aplicação do transporte paralelo), temos para s, t quaisquer

$$\begin{aligned}\frac{DY}{dt} \Big|_{(s, t_0)} &= \lim_{\Delta t \rightarrow 0} \frac{Y_{(s, t_0+\Delta t)} - \tau_{t_0 \rightarrow t_0+\Delta t} Y_{(s, t_0)}}{\Delta t}, \\ \frac{DY}{ds} \Big|_{(s_0, t)} &= \lim_{\Delta t \rightarrow 0} \frac{Y_{(s_0+\Delta s, t)} - \tau_{s_0 \rightarrow s_0+\Delta s} Y_{(s_0, t)}}{\Delta s}\end{aligned}$$

onde Y é um campo ao longo da superfície tal que $Y_{(s_0, t_0)} = W$. Com esta expressão em mente, reescrevemos o lado direito de (13.56) adicionando e subtraindo termos, de modo a decompor os dois caminhos do retângulo nos quatro lados do mesmo:

$$\begin{aligned}&\tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \circ \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0}(W) - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \circ \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0}(W) \\ &= \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \circ \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0}(Y_{(s_0, t_0)}) - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \circ \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0}(Y_{(s_0, t_0)}) \\ &= \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \left[\tau_{t_0 \rightarrow t_0+\Delta t}^{s_0}(Y_{(s_0, t_0)}) - Y_{(s_0, t_0+\Delta t)} \right] \\ &\quad + \tau_{s_0 \rightarrow s_0+\Delta s}^{t_0+\Delta t} \left[Y_{(s_0, t_0+\Delta t)} \right] - Y_{(s_0+\Delta s, t_0+\Delta t)} \\ &\quad - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \left[\tau_{s_0 \rightarrow s_0+\Delta s}^{t_0}(Y_{(s_0, t_0)}) - Y_{(s_0+\Delta s, t_0)} \right] \\ &\quad - \tau_{t_0 \rightarrow t_0+\Delta t}^{s_0+\Delta s} \left[Y_{(s_0+\Delta s, t_0)} \right] + Y_{(s_0+\Delta s, t_0+\Delta t)}.\end{aligned}$$

Portanto, segue do Lema 13.34 que

$$\begin{aligned}
& \lim_{\Delta s, \Delta t \rightarrow 0} \frac{1}{\Delta s \Delta t} \left[\tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0 + \Delta t} \circ \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0} (W) - \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} \circ \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0} (W) \right] \\
&= - \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0 + \Delta t} \left[\lim_{\Delta t \rightarrow 0} \frac{Y_{(s_0, t_0 + \Delta t)} - \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0} (Y_{(s_0, t_0)})}{\Delta t} \right] \\
&\quad - \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \lim_{\Delta s \rightarrow 0} \frac{Y_{(s_0 + \Delta s, t_0 + \Delta t)} - \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0 + \Delta t} (Y_{(s_0, t_0 + \Delta t)})}{\Delta s} \\
&\quad + \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} \left[\lim_{\Delta s \rightarrow 0} \frac{Y_{(s_0 + \Delta s, t_0)} - \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0} (Y_{(s_0, t_0)})}{\Delta s} \right] \\
&\quad + \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \lim_{\Delta t \rightarrow 0} \frac{Y_{(s_0 + \Delta s, t_0 + \Delta t)} - \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} (Y_{(s_0 + \Delta s, t_0)})}{\Delta t} \\
&= - \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0 + \Delta t} \left(\frac{DY}{dt} \Big|_{(s_0, t_0)} \right) \\
&\quad - \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{DY}{ds} \Big|_{(s_0, t_0 + \Delta t)} \\
&\quad + \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} \left(\frac{DY}{ds} \Big|_{(s_0, t_0)} \right) \\
&\quad + \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s} \frac{DY}{dt} \Big|_{(s_0 + \Delta s, t_0)} \\
&= \lim_{\Delta s \rightarrow 0} \frac{\frac{DY}{dt} \Big|_{(s_0 + \Delta s, t_0)} - \tau_{s_0 \rightarrow s_0 + \Delta s}^{t_0 + \Delta t} \frac{DY}{dt} \Big|_{(s_0, t_0)}}{\Delta s} \\
&\quad - \lim_{\Delta t \rightarrow 0} \frac{\frac{DY}{ds} \Big|_{(s_0, t_0 + \Delta t)} - \tau_{t_0 \rightarrow t_0 + \Delta t}^{s_0 + \Delta s} \left(\frac{DY}{ds} \Big|_{(s_0, t_0)} \right)}{\Delta t} \\
&= \frac{D}{ds} \frac{DY}{dt} \Big|_{(s_0, t_0)} - \frac{D}{dt} \frac{DY}{ds} \Big|_{(s_0, t_0)} \\
&= R(U, V) W.
\end{aligned}$$

■

13.7 Exercícios

13.38 Exercício. Mostre que toda aplicação

$$T : \mathcal{T}^1(M) \times \dots \times \mathcal{T}^1(M) \times \mathcal{T}_1(M) \times \dots \times \mathcal{T}_1(M) \longrightarrow C^\infty(M)$$

multilinear sobre $C^\infty(M)$ é induzida por um campo tensorial em $\mathcal{T}_l^k(M)$.

13.39 Exercício. Mostre que a curvatura é um invariante isométrico local. Conclua que

$$R(X, Y)Z = 0$$

para qualquer variedade métrica localmente isométrica a \mathbb{R}^n .

13.40 Exercício. Usando (a)-(d) da Proposição 13.14, prove o Corolário 13.15.

Capítulo 14

Derivada Covariante de Campos Tensoriais

Gostaríamos de definir uma noção completa de *cálculo diferencial* em variedades diferenciáveis, isto é, derivadas segundas e todas as derivadas de ordem superior. A derivada de uma função real f (sua diferencial) é um funcional linear df , isto é, um tensor 1-covariante. Para podermos derivar mais uma vez, precisamos definir a noção da derivada de um tensor 1-covariante. Para definirmos derivadas de ordem superior, precisamos definir as derivadas de tensores de todos os tipos (k, l) . Assim como a derivada de uma função aumenta a ordem covariante de 0 para 1, a derivada de (k, l) -tensores aumentará a ordem covariante de k para $k + 1$: ela será um tensor do tipo $(k + 1, l)$. Por este motivo, ela será chamada *derivada covariante*. Na verdade, ela será construída a partir da derivada covariante de campos vetoriais (a qual recebe este nome em função do seu uso na definição da derivada covariante de tensores) e isso será feito em duas etapas: primeiro generalizaremos a definição de derivada covariante de campos vetoriais (conexões), que é uma noção de derivada direcional, para definir derivadas covariantes de campos tensoriais, ou seja, derivadas direcionais de campos tensoriais (que será chamada uma *conexão* em $\oplus \mathcal{T}_l^k(M)$); esta será usada em seguida para definir a noção propriamente dita de derivada covariante de campos tensoriais, a chamada *derivada covariante total*.

14.1 Conexão nos Fibrados Tensoriais

Por definição, uma conexão em uma variedade diferenciável M é uma maneira de calcular derivadas covariantes de campos vetoriais. Esta conexão permite também definir derivadas covariantes para todos os campos tensoriais de M .

14.1 Lema. *Seja M uma variedade diferenciável dotada de uma conexão ∇ . Então existe uma única conexão*

$$\nabla : \mathcal{T}(M) \times \mathcal{T}_l^k(M) \longrightarrow \mathcal{T}_l^k(M)$$

em cada $\mathcal{T}_l^k(M)$ tal que

- (i) *Em $\mathcal{T}_1(M) = \mathcal{T}(M)$, ∇ coincide com a conexão dada.*
- (ii) *Em $\mathcal{T}_0(M) = C^\infty(M)$,*

$$\nabla_X f = Xf.$$

- (iii) *∇ satisfaaz a regra do produto com relação a produtos tensoriais:*

$$\nabla_X (F \otimes G) = (\nabla_X F) \otimes G + F \otimes (\nabla_X G).$$

- (iv) *∇ comuta com todos os traços: se tr denota o traço com relação a qualquer par de índices, então*

$$\nabla_X (\text{tr } F) = \text{tr} (\nabla_X F).$$

Além disso, esta conexão satisfaz também as propriedades adicionais:

(a) Para todos $Y \in \mathcal{T}(M)$ e $\omega \in \mathcal{T}^1(M)$ vale

$$\nabla_X [\omega(Y)] = (\nabla_X \omega)(Y) + \omega(\nabla_X Y).$$

(b) Para todos $T \in \mathcal{T}_l^k(M)$, $X_i \in \mathcal{T}(M)$ e $\omega^j \in \mathcal{T}^1(M)$ vale

$$\begin{aligned} (\nabla_X T)(X_1, \dots, X_k, \omega^1, \dots, \omega^l) &= X(T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\ &\quad - \sum_{i=1}^k T(X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\ &\quad - \sum_{j=1}^l T(X_1, \dots, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^l) \end{aligned}$$

Prova: Dividiremos a demonstração deste resultado em passos.

Passo 1. Se existe uma conexão que satisfaz as propriedades (i)-(iv), então ela necessariamente satisfaz (a)-(b).

De fato, (a) segue de (iii) e (iv):

$$\begin{aligned} \nabla_X [\omega(Y)] &= \nabla_X \text{tr}(\omega \otimes Y) \\ &= \text{tr} \nabla_X (\omega \otimes Y) \\ &= \text{tr}(\nabla_X \omega \otimes Y) + \text{tr}(\omega \otimes \nabla_X Y) \\ &= \nabla_X \omega(Y) + \omega(\nabla_X Y). \end{aligned}$$

Para provar (b), procedemos por indução separadamente sobre k e l . O caso $(k, l) = (0, 1)$ segue de (a) e (ii):

$$\begin{aligned} (\nabla_X \omega)(Y) &= \nabla_X [\omega(Y)] - \omega(\nabla_X Y) \\ &= X(\omega(Y)) - \omega(\nabla_X Y). \end{aligned}$$

Da mesma forma, o caso $(k, l) = (1, 0)$ segue de (a) e de (ii) (usando a definição da aplicação de um vetor a um covetor via a dualidade entre V e o bidual V^{**}):

$$\begin{aligned} (\nabla_X Y)(\omega) &= \omega(\nabla_X Y) \\ &= \nabla_X [\omega(Y)] - (\nabla_X \omega)(Y) \\ &= X[\omega(Y)] - Y(\nabla_X \omega). \end{aligned}$$

Agora assuma que (b) vale para todos os inteiros $p < k$, $q < l$. Mostraremos que isso implica que (b) vale para k, l . Como todo $T \in \mathcal{T}_l^k(M)$ se escreve na forma

$$T = \sum_{i,j=1}^n F_i \otimes G_j$$

para alguns $F_i \in \mathcal{T}_q^p(M)$ e $G_j \in \mathcal{T}_1^1(M)$, onde $p = k - 1$ e $q = l - 1$, pela linearidade da conexão é suficiente provar o resultado para

$$T = F \otimes G$$

com $F \in \mathcal{T}_q^p(M)$ e $G \in \mathcal{T}_1^1(M)$. Por (iii),

$$\nabla_X T = (\nabla_X F) \otimes G + F \otimes (\nabla_X G),$$

donde

$$\begin{aligned}
 & (\nabla_X T)(X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\
 &= [(\nabla_X F) \otimes G](X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\
 &\quad + [F \otimes (\nabla_X G)](X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\
 &= (\nabla_X F)(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \omega^l) \\
 &\quad + F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1})(\nabla_X G)(X_k, \omega^l).
 \end{aligned}$$

Mas

$$\begin{aligned}
 & (\nabla_X F)(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \omega^l) \\
 &= X(F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1})) G(X_k, \omega^l) \\
 &\quad - \sum_{i=1}^{k-1} F(X_1, \dots, \nabla_X X_i, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \omega^l) \\
 &\quad - \sum_{j=1}^{l-1} F(X_1, \dots, X_{k-1}, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^{l-1}) G(X_k, \omega^l)
 \end{aligned}$$

e

$$\begin{aligned}
 & F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1})(\nabla_X G)(X_k, \omega^l) \\
 &= F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) X(G(X_k, \omega^l)) \\
 &\quad - F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(\nabla_X X_k, \omega^l) \\
 &\quad - F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \nabla_X \omega^l).
 \end{aligned}$$

Portanto,

$$\begin{aligned}
 & (\nabla_X T)(X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\
 &= X(F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1})) G(X_k, \omega^l) + F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) X(G(X_k, \omega^l)) \\
 &\quad - \sum_{i=1}^{k-1} F(X_1, \dots, \nabla_X X_i, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \omega^l) \\
 &\quad - F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(\nabla_X X_k, \omega^l) \\
 &\quad - \sum_{j=1}^{l-1} F(X_1, \dots, X_{k-1}, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^{l-1}) G(X_k, \omega^l) \\
 &\quad - F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \nabla_X \omega^l) \\
 &= X[F(X_1, \dots, X_{k-1}, \omega^1, \dots, \omega^{l-1}) G(X_k, \omega^l)] \\
 &\quad - \sum_{i=1}^{k-1} (F \otimes G)(X_1, \dots, \nabla_X X_i, \dots, X_{k-1}, X_k, \omega^1, \dots, \omega^{l-1}, \omega^l) \\
 &\quad - (F \otimes G)(X_1, \dots, X_{k-1}, \nabla_X X_k, \omega^1, \dots, \omega^{l-1}, \omega^l) \\
 &\quad - \sum_{j=1}^{l-1} (F \otimes G)(X_1, \dots, X_{k-1}, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^{l-1}, \omega^l) \\
 &\quad - (F \otimes G)(X_1, \dots, X_{k-1}, X_k, \omega^1, \dots, \omega^{l-1}, \nabla_X \omega^l)
 \end{aligned}$$

$$\begin{aligned}
&= X \left[(F \otimes G) (X_1, \dots, X_k, \omega^1, \dots, \omega^l) \right] \\
&\quad - \sum_{i=1}^k (F \otimes G) (X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\
&\quad - \sum_{j=1}^{l-1} (F \otimes G) (X_1, \dots, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^l) \\
&= X (T (X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\
&\quad - \sum_{i=1}^k T (X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\
&\quad - \sum_{j=1}^{l-1} T (X_1, \dots, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^l).
\end{aligned}$$

Passo 2. Existência.

Defina $\nabla : \mathcal{T}(M) \times \mathcal{T}_l^k(M) \longrightarrow \mathcal{T}_l^k(M)$ por (b) (o que inclui (a), como visto acima). Mostraremos que ∇ é uma conexão e satisfaz todas as propriedades (i)-(iv).

Inicialmente, as propriedades de uma conexão:

$$(1) \quad \nabla_{fX+gY} T = f\nabla_X T + g\nabla_Y T.$$

Primeiro, para campos covetoriais: para todo Z vale

$$\begin{aligned}
(\nabla_{fX+gY} \omega)(Z) &= (fX + gY)(\omega(Z)) - \omega(\nabla_{fX+gY} Z) \\
&= fX(\omega(Z)) + gY(\omega(Z)) - \omega(f\nabla_X Z + g\nabla_Y Z) \\
&= fX(\omega(Z)) + gY(\omega(Z)) - f\omega(\nabla_X Z) - g\omega(\nabla_Y Z) \\
&= f[X(\omega(Z)) - \omega(\nabla_X Z)] + g[Y(\omega(Z)) - \omega(\nabla_Y Z)] \\
&= f\nabla_X \omega(Z) + g\nabla_Y \omega(Z),
\end{aligned}$$

logo

$$\nabla_{fX+gY} \omega = f\nabla_X \omega + g\nabla_Y \omega.$$

Para $T \in \mathcal{T}_l^k(M)$ temos

$$\begin{aligned}
&(\nabla_{fX+gY} T) (X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\
&= (fX + gY) (T (X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\
&\quad - \sum_{i=1}^k T (X_1, \dots, \nabla_{fX+gY} X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\
&\quad - \sum_{j=1}^{l-1} T (X_1, \dots, X_k, \omega^1, \dots, \nabla_{fX+gY} \omega^j, \dots, \omega^l).
\end{aligned}$$

Como

$$\begin{aligned}
&(fX + gY) (T (X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\
&= fX (T (X_1, \dots, X_k, \omega^1, \dots, \omega^l)) + gY (T (X_1, \dots, X_k, \omega^1, \dots, \omega^l)),
\end{aligned}$$

$$\begin{aligned}
&T (X_1, \dots, \nabla_{fX+gY} X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\
&= T (X_1, \dots, f\nabla_X X_i + g\nabla_Y X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\
&= fT (X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) + gT (X_1, \dots, \nabla_Y X_i, \dots, X_k, \omega^1, \dots, \omega^l)
\end{aligned}$$

e

$$\begin{aligned} & T(X_1, \dots, X_k, \omega^1, \dots, \nabla_{fX+gY}\omega^j, \dots, \omega^l) \\ &= T(X_1, \dots, X_k, \omega^1, \dots, f\nabla_X\omega^j + g\nabla_Y\omega^j, \dots, \omega^l) \\ &= fT(X_1, \dots, X_k, \omega^1, \dots, \nabla_X\omega^j, \dots, \omega^l) + gT(X_1, \dots, X_k, \omega^1, \dots, \nabla_Y\omega^j, \dots, \omega^l), \end{aligned}$$

segue o resultado.

$$(2) \nabla_X(T + S) = \nabla_X T + \nabla_X S.$$

É óbvio da definição.

$$(3) \nabla_X(fT) = f\nabla_X T + (Xf)T.$$

Note que esta regra do produto é um caso especial de (iii). Primeiro, para campos covetoriais: para todo Z vale

$$\begin{aligned} [\nabla_X(f\omega)](Z) &= X[(f\omega)(Z)] - (f\omega)(\nabla_X Z) \\ &= X[f\omega(Z)] - f\omega(\nabla_X Z) \\ &= (Xf)\omega(Z) + fX(\omega(Z)) - f\omega(\nabla_X Z) \\ &= (Xf)\omega(Z) + f[\nabla_X\omega](Z), \end{aligned}$$

logo

$$\nabla_X(f\omega) = f\nabla_X\omega + (Xf)\omega.$$

Para $T \in \mathcal{T}_l^k(M)$ temos

$$\begin{aligned} & (\nabla_X fT)(X_1, \dots, X_k, \omega^1, \dots, \omega^l) \\ &= X(fT(X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\ &\quad - \sum_{i=1}^k (fT)(X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\ &\quad - \sum_{j=1}^l (fT)(X_1, \dots, X_k, \omega^1, \dots, \nabla_X\omega^j, \dots, \omega^l) \\ &= (Xf)T(X_1, \dots, X_k, \omega^1, \dots, \omega^l) + fX(T(X_1, \dots, X_k, \omega^1, \dots, \omega^l)) \\ &\quad - \sum_{i=1}^k (fT)(X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \\ &\quad - \sum_{j=1}^l (fT)(X_1, \dots, X_k, \omega^1, \dots, \nabla_X\omega^j, \dots, \omega^l) \end{aligned}$$

onde segue o resultado.

Agora provemos as propriedades (i)-(iv).

(i) Por definição, assumindo (ii) (provada logo a seguir) e por (a)

$$\begin{aligned} (\nabla_X Y)(\omega) &= X[\omega(Y)] - Y(\nabla_X\omega) \\ &= \nabla_X[\omega(Y)] - (\nabla_X\omega)(Y) \\ &= \omega(\nabla_X Y) \end{aligned}$$

de modo que $\nabla_X Y$ coincide com a conexão dada.

(ii) Por definição,

$$\nabla_X f = Xf.$$

(iii) Por definição, se $F \in \mathcal{T}_l^k(M)$ e $G \in \mathcal{T}_q^p(M)$,

$$\begin{aligned} & [\nabla_X(F \otimes G)](X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q}) \\ &= X[(F \otimes G)(X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q})] \\ &\quad - \sum_{i=1}^{k+p} (F \otimes G)(X_1, \dots, \nabla_X X_i, \dots, X_{k+p}, \omega^1, \dots, \omega^{l+q}) \\ &\quad - \sum_{j=1}^{l+q} (F \otimes G)(X_1, \dots, X_{k+p}, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^{l+q}). \end{aligned}$$

Pela regra do produto para campos vetoriais,

$$\begin{aligned} & X[(F \otimes G)(X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q})] \\ &= X[F(X_1, \dots, X_k, \omega^1, \dots, \omega^l)G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q})] \\ &= X[F(X_1, \dots, X_k, \omega^1, \dots, \omega^l)]G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\ &\quad + F(X_1, \dots, X_k, \omega^1, \dots, \omega^l)X[G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q})]. \end{aligned}$$

Temos também

$$\begin{aligned} & \sum_{i=1}^{k+p} (F \otimes G)(X_1, \dots, \nabla_X X_i, \dots, X_{k+p}, \omega^1, \dots, \omega^{l+q}) \\ &= \sum_{i=1}^k F(X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l)G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\ &\quad + \sum_{i=k+1}^{k+p} F(X_1, \dots, X_k, \omega^1, \dots, \omega^l)G(X_{k+1}, \dots, \nabla_X X_i, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \end{aligned}$$

e

$$\begin{aligned} & \sum_{j=1}^{l+q} (F \otimes G)(X_1, \dots, X_{k+p}, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^{l+q}) \\ &= \sum_{j=1}^l F(X_1, \dots, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^l)G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\ &\quad + \sum_{j=l+1}^{l+q} F(X_1, \dots, X_k, \omega^1, \dots, \omega^l)G(X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \nabla_X \omega^j, \dots, \omega^{l+q}). \end{aligned}$$

Portanto,

$$\begin{aligned}
& [\nabla_X (F \otimes G)] (X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q}) \\
&= \left\{ X [F (X_1, \dots, X_k, \omega^1, \dots, \omega^l)] - \sum_{i=1}^k F (X_1, \dots, \nabla_X X_i, \dots, X_k, \omega^1, \dots, \omega^l) \right. \\
&\quad - \sum_{j=1}^l F (X_1, \dots, X_k, \omega^1, \dots, \nabla_X \omega^j, \dots, \omega^l) \Big\} \times G (X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\
&\quad + F (X_1, \dots, X_k, \omega^1, \dots, \omega^l) \times \{X [G (X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q})] \\
&\quad - \sum_{i=k+1}^{k+p} G (X_{k+1}, \dots, \nabla_X X_i, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\
&\quad - \sum_{j=l+1}^{l+q} G (X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \nabla_X \omega^j, \dots, \omega^{l+q}) \Big\} \\
&= (\nabla_X F) (X_1, \dots, X_k, \omega^1, \dots, \omega^l) G (X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\
&\quad + F (X_1, \dots, X_k, \omega^1, \dots, \omega^l) \nabla_X G (X_{k+1}, \dots, X_{k+p}, \omega^{l+1}, \dots, \omega^{l+q}) \\
&= [(\nabla_X F) \otimes G] (X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q}) \\
&\quad + [F \otimes (\nabla_X G)] (X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q}) \\
&= [(\nabla_X F) \otimes G + F \otimes (\nabla_X G)] (X_1, \dots, X_k, X_{k+1}, \dots, X_{k+p}, \omega^1, \dots, \omega^l, \omega^{l+1}, \dots, \omega^{l+q}).
\end{aligned}$$

(iv) Para provar esta propriedade, estabeleceremos primeiro uma fórmula para $\nabla_X T$ em coordenadas. Se

$$T = T_{i_1 \dots i_k}^{j_1 \dots j_l} dx^{i_1} \otimes \dots \otimes dx^{i_k} \otimes \partial_{j_1} \otimes \dots \otimes \partial_{j_l}$$

e

$$X = X^m \partial_m,$$

então

$$\begin{aligned}
& \nabla_X T (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&= \nabla_{X^m \partial_m} T (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&= X^m (\nabla_{\partial_m} T) (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&= X^m (\nabla_{\partial_m} T)_{i_1 \dots i_k}^{j_1 \dots j_l}
\end{aligned}$$

e a expressão em coordenadas para a derivada covariante de um campo tensorial, como veremos na Proposição 14.9, é

$$(\nabla_{\partial_m} T)_{i_1 \dots i_k}^{j_1 \dots j_l} = \partial_m T_{i_1 \dots i_k}^{j_1 \dots j_l} - \sum_{r=1}^k T_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k}^{j_1 \dots j_l} \Gamma_{mi_r}^p + \sum_{s=1}^l T_{i_1 \dots i_k}^{j_1 \dots j_{s-1} p j_{s+1} \dots j_l} \Gamma_{mp}^{j_s}. \quad (14.1)$$

Seja $F \in \mathcal{T}_l^k(M)$, de modo que $\text{tr } F \in \mathcal{T}_{l-1}^{k-1}(M)$. Por linearidade é suficiente provar que

$$[\nabla_{\partial_m} (\text{tr } F)] (\partial_{i_1}, \dots, \partial_{i_{k-1}}, dx^{j_1}, \dots, dx^{j_{l-1}}) = [\text{tr} (\nabla_{\partial_m} F)] (\partial_{i_1}, \dots, \partial_{i_{k-1}}, dx^{j_1}, \dots, dx^{j_{l-1}}),$$

isto é, que

$$[\nabla_{\partial_m} (\text{tr } F)]_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} = [\text{tr} (\nabla_{\partial_m} F)]_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}}.$$

Escrevendo

$$F = \sum_{\substack{i_1, \dots, i_k=1 \\ j_1, \dots, j_l=1}}^n F_{i_1 \dots i_k}^{j_1 \dots j_l} dx^{i_1} \otimes \dots \otimes dx^{i_k} \otimes \partial_{j_1} \otimes \dots \otimes \partial_{j_l},$$

temos, assumindo que o traço é tomado em relação aos índices p, q ,

$$(\text{tr } F)_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} = F_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}}.$$

Logo,

$$\begin{aligned} & [\nabla_{\partial_m} (\text{tr } F)]_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} \\ &= \frac{\partial (\text{tr } F)_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}}}{\partial x^m} - \sum_{r=1}^k (\text{tr } F)_{i_1 \dots i_{r-1} p i_r + 1 \dots i_k}^{j_1 \dots j_l} \Gamma_{mi_r}^p + \sum_{s=1}^l (\text{tr } F)_{i_1 \dots i_k}^{j_1 \dots j_{s-1} p j_s + 1 \dots j_l} \Gamma_{mp}^{j_s} \\ &= \frac{\partial F_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}}}{\partial x^m} - \sum_{r=1}^k F_{i_1 \dots i_{p-1} u i_p \dots i_{r-1} p i_r + 1 \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}} \Gamma_{mi_r}^p + \sum_{s=1}^l F_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{s-1} p j_s + 1 \dots j_{k-1}} \Gamma_{mp}^{j_s} \end{aligned}$$

e

$$\begin{aligned} & [\text{tr} (\nabla_{\partial_m} F)]_{i_1 \dots i_{k-1}}^{j_1 \dots j_{l-1}} \\ &= (\nabla_{\partial_m} F)_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}} \\ &= \frac{\partial F_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}}}{\partial x^m} - \sum_{r=1}^k F_{i_1 \dots i_{p-1} u i_p \dots i_{r-1} p i_r + 1 \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{k-1}} \Gamma_{mi_r}^p + \sum_{s=1}^l F_{i_1 \dots i_{p-1} u i_p \dots i_{l-1}}^{j_1 \dots j_{q-1} u j_q \dots j_{s-1} p j_s + 1 \dots j_{k-1}} \Gamma_{mp}^{j_s}. \end{aligned}$$

Comparando as expressões, vemos que elas são idênticas.

Passo 3. Unicidade.

Se existir uma conexão $\nabla : \mathcal{T}(M) \times \mathcal{T}_l^k(M) \rightarrow \mathcal{T}_l^k(M)$ que satisfaz (i)-(iv), então ela também satisfaz (a)-(b), logo será igual à conexão definida no Passo 2. ■

De agora em diante, quando nos referirmos a uma conexão em uma variedade diferenciável M , estaremos nos referindo à conexão do lema, definida em todas os fibrados tensoriais da variedade.

14.2 Proposição. *Seja M uma variedade diferenciável e ω um campo covetorial em M . Então*

$$\nabla_X \omega = \left(X^i \frac{\partial \omega_k}{\partial x^i} - X^i \omega_j \Gamma_{ik}^j \right) dx^k. \quad (14.2)$$

Em particular,

$$\nabla_{\partial_i} dx^j = -\Gamma_{ik}^j dx^k. \quad (14.3)$$

Prova: De fato,

$$\begin{aligned} (\nabla_X \omega)(Y) &= \nabla_X [\omega(Y)] - \omega(\nabla_X Y) \\ &= \nabla_X [\omega(Y^k \partial_k)] - \omega [(X(Y^k) + X^i Y^j \Gamma_{ij}^k) \partial_k] \\ &= X[Y^k \omega(\partial_k)] - (X(Y^k) + X^i Y^j \Gamma_{ij}^k) \omega(\partial_k) \\ &= X(Y^k \omega_k) - (X(Y^k) + X^i Y^j \Gamma_{ij}^k) \omega_k \\ &= X(Y^k) \omega_k + Y^k X(\omega_k) - X(Y^k) \omega_k + X^i Y^j \omega_k \Gamma_{ij}^k \\ &= Y^k X^i \frac{\partial \omega_k}{\partial x^i} - X^i Y^j \omega_k \Gamma_{ij}^k \\ &= \left(X^i \frac{\partial \omega_k}{\partial x^i} - X^i \omega_j \Gamma_{ik}^j \right) Y^k \\ &= \left(X^i \frac{\partial \omega_k}{\partial x^i} - X^i \omega_j \Gamma_{ik}^j \right) dx^k(Y). \end{aligned}$$

■

14.2 Derivada Covariante Total

14.3 Definição. Seja M uma variedade diferenciável dotada de uma conexão ∇ . Dado um campo (k, l) -tensorial $T \in \mathcal{T}_l^k(M)$, a **derivada covariante total** de T é o campo $(k+1, l)$ -tensorial

$$\nabla T : \mathcal{T}_1(M) \times \dots \times \mathcal{T}_1(M) \times \mathcal{T}^1(M) \times \dots \times \mathcal{T}^1(M) \longrightarrow C^\infty(M)$$

definido por

$$\nabla T(Y_1, \dots, Y_k, X, \omega^1, \dots, \omega^l) = \nabla_X T(Y_1, \dots, Y_k, \omega^1, \dots, \omega^l). \quad (14.4)$$

□

O nome *derivada covariante* pode ser agora compreendido: a derivada covariante total de um tensor aumenta em um a sua ordem covariante.

A derivada covariante total do tensor métrica g é o tensor identicamente nulo:

14.4 Proposição (Derivada Covariante Total do Tensor Métrica). Se (M, g) é uma variedade métrica, então

$$\nabla g = 0.$$

Prova: Pela definição e pela compatibilidade da conexão com a métrica, temos

$$\begin{aligned} \nabla g(X, Y, Z) &= \nabla_Z g(X, Y) \\ &= Z[g(X, Y)] - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \\ &= Z\langle X, Y \rangle - \langle \nabla_Z X, Y \rangle - \langle X, \nabla_Z Y \rangle \\ &= 0. \end{aligned}$$

■
Da demonstração vemos que a conexão riemanniana pode ser equivalentemente caracterizada como a única conexão simétrica tal que $\nabla g = 0$. O fato que $\nabla g = 0$ foi um sério empecilho para Einstein no seu caminho em direção à descoberta da Teoria da Relatividade Geral, como veremos no Capítulo 16.

14.5 Proposição (Significado da Derivada Covariante Total Nula). Se (M, g) é uma variedade métrica e $X \in \mathcal{T}(M)$, então

$$\nabla X = 0$$

se e somente se X é um campo paralelo.

Prova: Veja a observação que se segue à Definição 14.10 e observe que se $X \in \mathcal{T}(M)$, então por definição $\nabla X(Y) = \nabla_Y X$. ■

14.6 Definição. Dizemos que um tensor $T \in \mathcal{T}_l^k(M)$ é **paralelo** se $\nabla T = 0$. □

Assim, a compatibilidade da métrica é equivalente ao paralelismo do tensor métrica.

14.7 Exemplo (Hessiana e Laplaciano). Se $f \in C^\infty(M)$, então a derivada covariante total ∇f de f é simplesmente o covetor (1-forma) df , isto é, a diferencial de f . De fato, ambos os tensores tem o mesmo efeito sobre campos vetoriais:

$$df(X) = Xf = \nabla_X f = (\nabla f)(X).$$

O 2-tensor

$$\nabla^2 f = \nabla(\nabla f)$$

é chamado a **hessiana** (covariante) de f . Por definição,

$$\nabla^2 f(X, Y) = \nabla_Y(\nabla f)(X) = Y((\nabla f)(X)) - \nabla f(\nabla_Y X),$$

ou seja,

$$\nabla^2 f(X, Y) = Y(Xf) - (\nabla_Y X)f. \quad (14.5)$$

O **laplaciano** de f é definido por

$$\Delta f = \text{tr}_g \nabla^2 f. \quad (14.6)$$

□

14.8 Notação.

Seja

$$T = T_{i_1 \dots i_k}^{j_1 \dots j_l} dx^{i_1} \otimes \dots \otimes dx^{i_k} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}}$$

um campo (k, l) -tensorial. Então temos duas notações bastante difundidas para escrever a expressão em coordenadas da derivada covariante total de T :

$$\nabla T = \nabla_m T_{i_1 \dots i_k}^{j_1 \dots j_l} dx^{i_1} \otimes \dots \otimes dx^{i_k} \otimes dx^m \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \quad (14.7)$$

e

$$\nabla T = T_{i_1 \dots i_k; m}^{j_1 \dots j_l} dx^{i_1} \otimes \dots \otimes dx^{i_k} \otimes dx^m \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_l}} \quad (14.8)$$

Isto é, cada componente do campo $(k + 1, l)$ -tensorial derivada covariante total é denotado por

$$\nabla_m T_{i_1 \dots i_k}^{j_1 \dots j_l} = T_{i_1 \dots i_k; m}^{j_1 \dots j_l}. \quad (14.9)$$

Por exemplo, se

$$X = X^i \frac{\partial}{\partial x^i},$$

então

$$\nabla X = \nabla_j X^i dx^j \otimes \frac{\partial}{\partial x^i} = X^i_{;j} dx^j \otimes \frac{\partial}{\partial x^i}. \quad (14.10)$$

□

14.9 Proposição. *Seja M uma variedade diferenciável dotada de uma conexão ∇ . Então as componentes da derivada covariante total em um sistema de coordenadas são dadas por*

$$\nabla_m T_{i_1 \dots i_k}^{j_1 \dots j_l} = T_{i_1 \dots i_k; m}^{j_1 \dots j_l} = \partial_m T_{i_1 \dots i_k}^{j_1 \dots j_l} + \sum_{s=1}^l T_{i_1 \dots i_k}^{j_1 \dots j_{s-1} p j_{s+1} \dots j_l} \Gamma_{mp}^{j_s} - \sum_{r=1}^k T_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k}^{j_1 \dots j_l} \Gamma_{mi_r}^p. \quad (14.11)$$

Em particular, em coordenadas normais,

$$\nabla_m T_{i_1 \dots i_k}^{j_1 \dots j_l}(p) = \partial_m T_{i_1 \dots i_k}^{j_1 \dots j_l}(p).$$

Prova: Por definição,

$$\begin{aligned}
T_{i_1 \dots i_k; m}^{j_1 \dots j_l} &= \nabla T(\partial_{i_1}, \dots, \partial_{i_k}, \partial_m, dx^{j_1}, \dots, dx^{j_l}) \\
&= (\nabla_{\partial_m} T)(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&= \partial_m T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) - \sum_{r=1}^k T(\partial_{i_1}, \dots, \nabla_{\partial_m} \partial_{i_r}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&\quad - \sum_{s=1}^l T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, \nabla_{\partial_m} dx^{j_s}, \dots, dx^{j_l}) \\
&= \partial_m T_{i_1 \dots i_k}^{j_1 \dots j_l} - \sum_{r=1}^k T(\partial_{i_1}, \dots, \Gamma_{m i_r}^p \partial_p, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \\
&\quad - \sum_{s=1}^l T(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, -\Gamma_{m p}^{j_s} dx^p, \dots, dx^{j_l}) \\
&= \partial_m T_{i_1 \dots i_k}^{j_1 \dots j_l} - \sum_{r=1}^k T_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k}^{j_1 \dots j_l} \Gamma_{m i_r}^p + \sum_{s=1}^l T_{i_1 \dots i_k}^{j_1 \dots j_{s-1} p j_{s+1} \dots j_l} \Gamma_{m p}^{j_s}.
\end{aligned}$$

Em coordenadas normais, os coeficientes de Christoffel em p se anulam. ■

Por exemplo, para uma função real:

$$\nabla_i f = f_{;i} = \frac{\partial f}{\partial x^i};$$

para um campo vetorial (campo 1-tensorial contravariante)

$$\nabla_j X^i = X_{;j}^i = \partial_j X^i + X^p \Gamma_{jp}^i;$$

para um campo covetorial (campo 1-tensorial covariante)

$$\nabla_j \omega_i = \omega_{i;j} = \partial_j \omega_i - \omega_p \Gamma_{ij}^p;$$

para um campo 2-tensorial covariante:

$$\nabla_k T_{ij} = T_{ij;k} = \partial_k T_{ij} - T_{pj} \Gamma_{ik}^p - T_{ip} \Gamma_{jk}^p;$$

para um campo $(3, 1)$ -tensorial:

$$\nabla_m R_{ijk}^l = R_{ijk;m}^l = \partial_m R_{ijk}^l + R_{ijk}^p \Gamma_{pm}^l - R_{pj}^l \Gamma_{im}^p - R_{ip}^l \Gamma_{jm}^p - R_{ijp}^l \Gamma_{km}^p;$$

e para um campo 4-tensorial covariante:

$$\nabla_m R_{ijkl} = R_{ijkl;m} = \partial_m R_{ijkl} - R_{pjkl} \Gamma_{im}^p - R_{ipkl} \Gamma_{jm}^p - R_{ijpl} \Gamma_{km}^p - R_{ijkp} \Gamma_{lm}^p.$$

14.10 Proposição (Regra do Produto). *Seja M uma variedade diferenciável dotada de uma conexão ∇ . Então*

$$\nabla_m \left(F_{i_1 \dots i_k}^{j_1 \dots j_l} G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} \right) = \left(\nabla_m F_{i_1 \dots i_k}^{j_1 \dots j_l} \right) G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} + F_{i_1 \dots i_k}^{j_1 \dots j_l} \left(\nabla_m G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} \right).$$

Prova: Seja $T = F \otimes G$, de modo que

$$T_{i_1 \dots i_k i_{k+1} \dots i_{k+p}}^{j_1 \dots j_l j_{l+1} \dots j_{l+q}} = F_{i_1 \dots i_k}^{j_1 \dots j_l} G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}}.$$

Daí,

$$\begin{aligned}
& \nabla_m T_{i_1 \dots i_k i_{k+1} \dots i_{k+p}}^{j_1 \dots j_l j_{l+1} \dots j_{l+q}} \\
&= \nabla T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, \partial_m, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= \nabla_{\partial_m} T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= \partial_m T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&\quad - \sum_{r=1}^{k+p} T (\partial_{i_1}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&\quad - \sum_{s=1}^{l+q} T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}).
\end{aligned}$$

Como

$$\begin{aligned}
& \partial_m T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= \partial_m (F_{i_1 \dots i_k}^{j_1 \dots j_l} G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}}) \\
&= (\partial_m F_{i_1 \dots i_k}^{j_1 \dots j_l}) G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} + F_{i_1 \dots i_k}^{j_1 \dots j_l} (\partial_m G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}}), \\
& \sum_{r=1}^{k+p} T (\partial_{i_1}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= - \sum_{r=1}^k F (\partial_{i_1}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) G (\partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&\quad - \sum_{r=k+1}^{k+p} F (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) G (\partial_{i_{k+1}}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= - \sum_{r=1}^k F (\partial_{i_1}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} \\
&\quad - \sum_{r=k+1}^{k+p} F_{i_1 \dots i_k}^{j_1 \dots j_l} G (\partial_{i_{k+1}}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}})
\end{aligned}$$

e

$$\begin{aligned}
& \sum_{s=1}^{l+q} T (\partial_{i_1}, \dots, \partial_{i_k}, \partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_1}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_l}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&= - \sum_{s=1}^l F (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_l}) G (\partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \\
&\quad - \sum_{s=l+1}^{l+q} F (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) G (\partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_{l+q}}) \\
&= - \sum_{s=1}^l F (\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_l}) G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} \\
&\quad - \sum_{s=l+1}^{l+q} F_{i_1 \dots i_k}^{j_1 \dots j_l} G (\partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_{l+q}}),
\end{aligned}$$

segue que

$$\begin{aligned}
& \nabla_m T_{i_1 \dots i_k i_{k+1} \dots i_{k+p}}^{j_1 \dots j_l j_{l+1} \dots j_{l+q}} \\
&= \left[\partial_m F_{i_1 \dots i_k}^{j_1 \dots j_l} - \sum_{r=1}^k F(\partial_{i_1}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_k}, dx^{j_1}, \dots, dx^{j_l}) \right. \\
&\quad \left. - \sum_{s=1}^l F(\partial_{i_1}, \dots, \partial_{i_k}, dx^{j_1}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_l}) \right] G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} \\
&\quad + F_{i_1 \dots i_k}^{j_1 \dots j_l} \left[\partial_m G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} - \sum_{r=k+1}^{k+p} G(\partial_{i_{k+1}}, \dots, \nabla_{\partial_{i_m}} \partial_{i_r}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, dx^{j_{l+q}}) \right. \\
&\quad \left. - \sum_{s=l+1}^{l+q} G(\partial_{i_{k+1}}, \dots, \partial_{i_{k+p}}, dx^{j_{l+1}}, \dots, \nabla_{\partial_{i_m}} dx^{j_s}, \dots, dx^{j_{l+q}}) \right] \\
&= (\nabla_m F_{i_1 \dots i_k}^{j_1 \dots j_l}) G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}} + F_{i_1 \dots i_k}^{j_1 \dots j_l} (\nabla_m G_{i_{k+1} \dots i_{k+p}}^{j_{l+1} \dots j_{l+q}}).
\end{aligned}$$

■

Este resultado justifica a notação ∇_m .

14.11 Lema. *Seja M uma variedade métrica. Então*

$$\nabla_k g_{ij} = g_{ij;k} = 0$$

e

$$\nabla_k g^{ij} = g_{;k}^{ij} = 0$$

para todos os índices i, j, k .

Prova: Embora já tenhamos demonstrado a primeira afirmação na Proposição 14.4, forneceremos outra demonstração que a derivada covariante total do tensor métrica g é o tensor identicamente nulo usando a fórmula para as componentes da derivada covariante. De fato, pelo Corolário 11.25,

$$\partial_k g_{ij} = g_{ip} \Gamma_{jk}^p + g_{pj} \Gamma_{ik}^p.$$

Logo,

$$\nabla_k g_{ij} = g_{ij;k} = \partial_k g_{ij} - g_{ip} \Gamma_{jk}^p - g_{pj} \Gamma_{ik}^p = 0.$$

Para calcular $\nabla_k g^{ij}$, primeiro observamos que

$$\begin{aligned}
\nabla_k \delta_i^j &= \delta_{i;k}^j = \partial_k \delta_i^j + \delta_i^p \Gamma_{pk}^j - \delta_p^j \Gamma_{ik}^p \\
&= \delta_i^i \Gamma_{ik}^j - \delta_j^j \Gamma_{ik}^i \\
&= 0.
\end{aligned}$$

Portanto,

$$\nabla_k (g_{ip} g^{pj}) = \nabla_k \delta_i^j = 0.$$

Mas, pela regra do produto,

$$\nabla_k (g_{ip} g^{pj}) = g^{pj} \nabla_k g_{ip} + g_{ip} \nabla_k g^{pj} = g_{ip} \nabla_k g^{pj}$$

pela primeira afirmação. Concluímos que

$$g_{ip} \nabla_k g^{pj} = 0.$$

Contraindo esta expressão através da inversa da métrica,

$$g^{li} g_{ip} \nabla_k g^{pj} = 0,$$

segue que

$$0 = \delta_p^l \nabla_k g^{pj} = \nabla_k g^{lj}.$$

■

Em vista dos resultados acima, a derivada covariante se comporta bem com a operação de subir ou descer um índice:

14.12 Proposição. *Seja M uma variedade métrica. Então*

$$\nabla_m T_{i_1 \dots i_k i_{k+1}}^{j_1 \dots j_{l-1}} = \nabla_m \left(g_{i_{k+1} p} T_{i_1 \dots i_k}^{j_1 \dots j_{l-1} p} \right) = g_{i_{k+1} p} \nabla_m T_{i_1 \dots i_k}^{j_1 \dots j_{l-1} p}, \quad (14.12)$$

e

$$\nabla_m T_{i_1 \dots i_{k-1}}^{j_1 \dots j_l j_{l+1}} = \nabla_m \left(g^{j_{l+1} p} T_{i_1 \dots i_{k-1} p}^{j_1 \dots j_l} \right) = g^{j_{l+1} p} \nabla_m T_{i_1 \dots i_{k-1} p}^{j_1 \dots j_l}. \quad (14.13)$$

Prova: Segue das Proposições 14.10 e 14.11. ■

No caso de um $(1, 1)$ -tensor, as fórmulas são

$$\begin{aligned} \nabla_m T_i^j &= \nabla_m (g_{ik} T^{kj}) = g_{ik} (\nabla_m T^{kj}), \\ \nabla_m T_i^j &= \nabla_m (g^{kj} T_{ik}) = g^{kj} (\nabla_m T_{ik}). \end{aligned}$$

Na notação de ponto e vírgula:

$$\begin{aligned} T_{i;m}^j &= g_{ik} T_{;m}^{kj}, \\ \nabla_m T_{i;m}^j &= g^{kj} T_{ik;m}. \end{aligned}$$

14.13 Exemplo (Hessiana e Laplaciano em coordenadas). As componentes de ∇f são

$$\nabla f = \partial_i f dx^i.$$

É comum denotar

$$f_i = \partial_i f, \quad (14.14)$$

de modo que

$$\nabla^2 f = \nabla_j f_i dx^i \otimes dx^j = f_{i;j} dx^i \otimes dx^j.$$

Temos

$$\nabla_j f_i = f_{i;j} = \partial_j f_i - f_p \Gamma_{ij}^p = \partial_{ij}^2 f - \Gamma_{ij}^k \partial_k f. \quad (14.15)$$

Ou seja,

$$\nabla^2 f = (\partial_{ij}^2 f - \Gamma_{ij}^k \partial_k f) dx^i \otimes dx^j. \quad (14.16)$$

Também é comum escrever

$$f_{ij} = \partial_{ij}^2 f, \quad (14.17)$$

de modo que

$$\nabla^2 f = (f_{ij} - f_k \Gamma_{ij}^k) dx^i \otimes dx^j. \quad (14.18)$$

Daí obtemos a fórmula do laplaciano em coordenadas:

$$\begin{aligned} \Delta f &= \text{tr}_g \nabla^2 f = g^{ij} (\nabla^2 f)_{ij} \\ &= g^{ij} \partial_{ij}^2 f - g^{ij} \Gamma_{ij}^k \partial_k f. \end{aligned}$$

□

14.14 Proposição (Identidade de Bianchi Diferencial). *Seja M uma variedade métrica. A derivada covariante total do tensor curvatura satisfaz a seguinte propriedade:*

$$\nabla R(X, Y, Z, V, W) + \nabla R(X, Y, V, W, Z) + \nabla R(X, Y, W, Z, V) = 0. \quad (14.19)$$

Em termos de componentes,

$$R_{ijkl;m} + R_{ijlm;k} + R_{ijmk;l} = \nabla_m R_{ijkl} + \nabla_k R_{ijlm} + \nabla_l R_{ijmk} = 0. \quad (14.20)$$

Prova: Em primeiro lugar, devido à propriedade de simetria 14.14 (c) do tensor curvatura, a identidade do enunciado é equivalente à identidade

$$\nabla R(Z, V, X, Y, W) + \nabla R(V, W, X, Y, Z) + \nabla R(W, Z, X, Y, V) = 0. \quad (14.21)$$

A vantagem desta identidade é que o campo na quarta posição é o mesmo (isto é, Y) em todos os termos. Por multilinearidade, basta provar que

$$\nabla R(\partial_i, \partial_j, \partial_k, \partial_l, \partial_m) + \nabla R(\partial_j, \partial_m, \partial_k, \partial_l, \partial_i) + \nabla R(\partial_m, \partial_i, \partial_k, \partial_l, \partial_j) = 0$$

para todos os campos vetoriais $\partial_i, \partial_j, \partial_k, \partial_l, \partial_m$. Por definição e compatibilidade da métrica,

$$\begin{aligned} \nabla R(\partial_i, \partial_j, \partial_k, \partial_l, \partial_m) &= \nabla_{\partial_m} R(\partial_i, \partial_j, \partial_k, \partial_l) \\ &= \nabla_{\partial_m} \langle R(\partial_i, \partial_j) \partial_k, \partial_l \rangle \\ &= \langle \nabla_{\partial_m} R(\partial_i, \partial_j) \partial_k, \partial_l \rangle + \langle R(\partial_i, \partial_j) \partial_k, \nabla_{\partial_m} \partial_l \rangle. \end{aligned}$$

Para facilitar os cálculos consideravelmente, consideramos coordenadas normais em p . Em coordenadas normais em p ,

$$\nabla_{\partial_m} \partial_l(p) = \Gamma_{lm}^q(p) \partial_q = 0.$$

Logo, em p ,

$$\nabla R(\partial_i, \partial_j, \partial_k, \partial_l, \partial_m) = \langle \nabla_{\partial_m} R(\partial_i, \partial_j) \partial_k, \partial_l \rangle.$$

Escrevendo

$$\begin{aligned} R(\partial_i, \partial_j) \partial_k &= \nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_j} \nabla_{\partial_i} \partial_k - \nabla_{[\partial_i, \partial_j]} \partial_k \\ &= \nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_j} \nabla_{\partial_i} \partial_k, \end{aligned}$$

porque $[\partial_i, \partial_j] = 0$, segue que

$$\nabla R(\partial_i, \partial_j, \partial_k, \partial_l, \partial_m) = \langle \nabla_{\partial_m} \nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_m} \nabla_{\partial_j} \nabla_{\partial_i} \partial_k, \partial_l \rangle.$$

Daí, em p ,

$$\begin{aligned} \nabla R(\partial_i, \partial_j, \partial_k, \partial_l, \partial_m) + \nabla R(\partial_j, \partial_m, \partial_k, \partial_l, \partial_i) + \nabla R(\partial_m, \partial_i, \partial_k, \partial_l, \partial_j) \\ &= \langle \nabla_{\partial_m} \nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_m} \nabla_{\partial_j} \nabla_{\partial_i} \partial_k + \nabla_{\partial_i} \nabla_{\partial_j} \nabla_{\partial_m} \partial_k - \nabla_{\partial_i} \nabla_{\partial_m} \nabla_{\partial_j} \partial_k \\ &\quad + \nabla_{\partial_j} \nabla_{\partial_m} \nabla_{\partial_i} \partial_k - \nabla_{\partial_j} \nabla_{\partial_i} \nabla_{\partial_m} \partial_k, \partial_l \rangle \\ &= \langle \nabla_{\partial_j} \nabla_{\partial_m} \nabla_{\partial_i} \partial_k - \nabla_{\partial_m} \nabla_{\partial_j} \nabla_{\partial_i} \partial_k + \nabla_{\partial_m} \nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_i} \nabla_{\partial_m} \nabla_{\partial_j} \partial_k \\ &\quad + \nabla_{\partial_i} \nabla_{\partial_j} \nabla_{\partial_m} \partial_k - \nabla_{\partial_j} \nabla_{\partial_i} \nabla_{\partial_m} \partial_k, \partial_l \rangle \\ &= \langle R(\partial_j, \partial_m) \nabla_{\partial_i} \partial_k + R(\partial_m, \partial_i) \nabla_{\partial_j} \partial_k + R(\partial_i, \partial_j) \nabla_{\partial_m} \partial_k, \partial_l \rangle \\ &= \langle R(\partial_j, \partial_m) 0 + R(\partial_m, \partial_i) 0 + R(\partial_i, \partial_j) 0, \partial_l \rangle \\ &= 0, \end{aligned}$$

já que $\nabla_{\partial_i} \partial_k = \nabla_{\partial_j} \partial_k = \nabla_{\partial_m} \partial_k = 0$ em p . ■

14.2.1 Divergência de Campos Tensoriais

Dado um campo vetorial X em \mathbb{R}^n , sua divergência é a função

$$\operatorname{div} X = \partial_i X^i = \nabla_i X^i.$$

Isto é,

$$\operatorname{div} X = \operatorname{tr} \nabla X.$$

Dado um campo vetorial $X \in \mathcal{T}(M)$, de maneira análoga definimos sua divergência por

$$\operatorname{div} X = \operatorname{tr} \nabla X.$$

Em coordenadas,

$$\operatorname{div} X = \nabla_i X^i = X_{;i}^i.$$

Para definir a divergência de um campo covetorial $\omega \in \mathcal{T}_1(M)$, observe que ao descermos índices no último termo obtemos

$$X_{;i}^i = g^{ij} X_{j;i} = g^{ij} \left(X^b \right)_{j;i} = \operatorname{tr}_g \nabla \left(X^b \right),$$

de modo que faz sentido definir

$$\operatorname{div} \omega = \operatorname{tr}_g \nabla \omega.$$

Em coordenadas,

$$\operatorname{div} \omega = g^{ij} \nabla_i \omega_j = g^{ij} \omega_{j;i}.$$

Usando estas idéias, definimos a divergência de tensores contravariantes e covariantes da seguinte forma:

14.15 Definição. Dado um campo l -tensorial contravariante, a **divergência** de T é o campo $(l - 1)$ -tensorial contravariante

$$\operatorname{div} T = \operatorname{tr} \nabla T.$$

Dado um campo k -tensorial covariante T , a **divergência** de T é o campo $(k - 1)$ -tensorial covariante

$$\operatorname{div} T = \operatorname{tr}_g \nabla T.$$

□

Observe que

$$\Delta f = \operatorname{tr}_g \nabla^2 f = \operatorname{div} \nabla f.$$

Em coordenadas, se T é um 2-tensor contravariante,

$$(\operatorname{div} T)_i = \nabla_j T^{ij} = T_{;j}^{ij}.$$

se T é um 2-tensor covariante,

$$(\operatorname{div} T)_i = g^{jk} \nabla_k T_{ij}.$$

14.16 Proposição. Seja M uma variedade métrica. Então

$$\nabla S = 2 \operatorname{div} \operatorname{Ric}. \quad (14.22)$$

Prova: Por definição,

$$(\operatorname{div} \operatorname{Ric})_i = g^{jk} \nabla_k R_{ij}.$$

Pela identidade de Bianchi diferencial, temos

$$\nabla_m R_{ijkl} + \nabla_k R_{ijlm} + \nabla_l R_{ijmk} = 0.$$

Daí,

$$g^{jl} (\nabla_m R_{ijkl} + \nabla_k R_{ijlm} + \nabla_l R_{ijmk}) = 0,$$

onde, pela regra do produto e pelo fato que $\nabla_p g^{rs} = 0$,

$$\nabla_m (g^{jl} R_{ijkl}) + \nabla_k (g^{jl} R_{ijlm}) + g^{jl} \nabla_l R_{ijmk} = 0. \quad (14.23)$$

Como

$$\begin{aligned} R_{pq} &= g^{rs} R_{rpqs}, \\ R_{ijkl} &= -R_{jikl}, \\ R_{ijlm} &= R_{jiml}, \end{aligned}$$

o primeiro termo desta soma é

$$-\nabla_m R_{ik},$$

enquanto que o segundo termo é

$$\nabla_k R_{im}.$$

Logo a identidade (14.23) é equivalente a

$$\nabla_m R_{ik} = \nabla_k R_{im} + g^{jl} \nabla_l R_{ijmk}.$$

Daí,

$$\begin{aligned} g^{ik} \nabla_m R_{ik} &= g^{ik} \nabla_k R_{im} + g^{ik} g^{jl} \nabla_l R_{ijmk} \\ &= (\text{div Ric})_m + g^{jl} \nabla_l (g^{ik} R_{ijmk}) \\ &= (\text{div Ric})_m + g^{jl} \nabla_l R_{jm} \\ &= (\text{div Ric})_m + (\text{div Ric})_m \\ &= 2 (\text{div Ric})_m. \end{aligned}$$

Portanto,

$$\nabla_m S = \nabla_m (g^{ik} R_{ik}) = g^{ik} \nabla_m R_{ik} = 2 (\text{div Ric})_m.$$

■

14.17 Proposição. Se M^n é uma variedade de Einstein conexa e $n \geq 3$, então M tem curvatura escalar constante.

Prova: Temos

$$\text{Ric} = \frac{S}{n} g,$$

onde

$$R_{ij} = \frac{S}{n} g_{ij}$$

e, portanto,

$$\nabla_k R_{ij} = \frac{1}{n} g_{ij} \nabla_k S.$$

Logo,

$$(\text{div Ric})_i = g^{jk} \nabla_k R_{ij} = \frac{1}{n} g^{kj} g_{ij} \nabla_k S = \frac{1}{n} \delta_i^k \nabla_k S = \frac{1}{n} \nabla_i S.$$

Segue da proposição anterior que

$$\frac{\nabla S}{2} = \frac{\nabla S}{n}.$$

Se $n \neq 2$, esta identidade implica que

$$\nabla S = 0$$

e portanto $S \equiv$ constante, se M é conexa. ■

Como toda variedade riemanniana de dimensão 2 é uma variedade de Einstein e existem variedades de dimensão 2 que não possuem curvatura escalar constante, este resultado não vale se $n = 2$.

Capítulo 15

O Tensor Momento-Energia

Na Teoria da Relatividade Geral, o *tensor momento-energia* (também chamado *tensor tensão-energia*) descreve a densidade e fluxo de energia e momento no espaço-tempo e é responsável por criar a curvatura no espaço-tempo. Nas palavras de Wheeler, *o espaço-tempo age na matéria dizendo a ela como se mover* [isto é, na ausência de forças atuando, apenas sob o efeito de sua própria inércia, ao longo das geodésicas do espaço-tempo]; *por sua vez, a matéria age sobre o espaço-tempo, dizendo a ele como se curvar* ([MTW], Secção 1.2, p. 5). Exceto que matéria na Teoria da Relatividade possui um significado mais geral, não restrito apenas à massa.

15.1 4-Vetores

Tanto na Teoria da Relatividade Especial quanto na Geral, é conveniente descrever o movimento de um objeto em um referencial inercial (em Relatividade Especial) ou em um sistema de coordenadas geral (em Relatividade Geral) em coordenadas do espaço-tempo dependendo do seu tempo próprio τ , isto é, de acordo com o tempo medido em seu próprio referencial, e portanto independente de referencial. Assim, em um referencial inercial a trajetória de uma partícula livre é uma curva parametrizada pelo tempo próprio

$$\alpha(\tau) = (t(\tau), x(\tau), y(\tau), z(\tau)),$$

e a geodésica γ que uma partícula livre segue no espaço-tempo curvo seria parametrizada pelo tempo próprio $\gamma(\tau)$.

Vetores no espaço-tempo tem 4 coordenadas e para distingui-los dos vetores espaciais (projeções espaciais), são frequentemente chamados de *quadrivetores* ou 4-vetores. Assim, temos o **quadrivetor deslocamento**, ou seja deslocamento no espaço-tempo

$$\Delta \mathbf{s} = (\Delta t, \Delta x, \Delta y, \Delta z)$$

e o **quadrivetor velocidade**

$$\mathbf{u} = \frac{d\mathbf{s}}{d\tau} = \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau} \right), \quad (15.1)$$

ou $\alpha'(\tau)$ para uma trajetória curva $\alpha(\tau)$ no espaço-tempo. O quadrivetor velocidade é por definição tangente à trajetória da partícula no espaço-tempo.

Observe que se a trajetória do objeto no espaço-tempo é uma curva, o seu referencial próprio obviamente não é um referencial inercial. Entretanto, em cada instante de tempo existe um referencial inercial que tem a mesma velocidade que o objeto; este referencial inercial que *comove* com o objeto é temporário, pois no próximo instante de tempo ele não está mais se comovendo com o objeto. Ele é chamado o seu *referencial comóvel temporário* (em inglês: *momentarily comoving reference frame* (MCRF)). O quadrivetor velocidade

em um determinado instante de tempo (tangente à trajetória da partícula naquele instante de tempo) é justamente o vetor base e_0 do seu referencial comóvel temporário naquele momento.

Como o tempo próprio τ do objeto e o tempo t de um referencial inercial arbitrário estão relacionados através da equação

$$\tau = \sqrt{1 - v^2} t, \quad (15.2)$$

onde \mathbf{v} é a velocidade espacial do objeto em relação a este referencial inercial ($v = \|\mathbf{v}\|$), isto é,

$$\mathbf{v} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right),$$

segue que

$$d\tau = \sqrt{1 - v^2} dt,$$

e portanto as coordenadas do quadrvetor velocidade no referencial inercial também podem ser escritas na forma

$$\mathbf{u} = \frac{1}{\sqrt{1 - v^2}} \left(1, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = \frac{1}{\sqrt{1 - v^2}} (1, \mathbf{v})$$

ou

$$\mathbf{u} = \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z) = \frac{1}{\sqrt{1 - v^2}} (1, \mathbf{v}). \quad (15.3)$$

Observe que quando $v \ll 1$ temos $\tau \approx t$ e

$$\mathbf{u} \approx (1, v^x, v^y, v^z) = (1, \mathbf{v}),$$

e o quadrvetor velocidade de um objeto em repouso em relação ao referencial inercial (trajetória vertical no espaço-tempo) tem neste referencial coordenadas

$$\mathbf{u} = (1, 0, 0, 0).$$

Em particular, segue também que se $v \neq 1$ o quadrvetor velocidade é um *vetor unitário*, isto é,

$$\|\mathbf{u}\| = 1 \quad (15.4)$$

pois

$$\begin{aligned} \langle \mathbf{u}, \mathbf{u} \rangle &= \left\langle \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z), \frac{1}{\sqrt{1 - v^2}} (1, v^x, v^y, v^z) \right\rangle \\ &= \frac{1}{1 - v^2} \left[-1 + (v^x)^2 + (v^y)^2 + (v^z)^2 \right] \\ &= \frac{1}{1 - v^2} [-1 + v^2] \\ &= -1. \end{aligned}$$

15.2 Momento-Energia

15.1 Definição. Se m é a massa de um objeto, definimos o **quadrvetor momento-energia**

$$\mathbf{P} = m\mathbf{u}. \quad (15.5)$$

□

Em um referencial inercial arbitrário, segue que o quadrivetor momento-energia tem coordenadas

$$\mathbf{P} = \frac{m}{\sqrt{1-v^2}} (1, v^x, v^y, v^z) = \frac{m}{\sqrt{1-v^2}} (1, \mathbf{v}).$$

Observe que como velocidade é adimensional em relatividade (tempo e espaço são medidos com a mesma unidade), o quadrivetor momento-energia tem unidades de massa. De fato, se $v \neq 1$

$$\langle \mathbf{P}, \mathbf{P} \rangle = \langle m\mathbf{u}, m\mathbf{u} \rangle = m^2 \langle \mathbf{u}, \mathbf{u} \rangle = -m^2, \quad (15.6)$$

de modo que

$$\|\mathbf{P}\| = m. \quad (15.7)$$

Note que a massa m de um objeto, sendo a norma do quadrivetor momento-energia (um vetor do espaço-tempo de Minkowski), tem o mesmo valor em qualquer referencial inercial. Em outras palavras, ela não depende da velocidade do referencial: massa não aumenta com a velocidade e na teoria da relatividade não se define “massa relativística” (veja [Okun1] para uma exposição clara, incluindo a origem histórica deste termo e até seu uso inicial por Einstein para inferir a atração da luz por uma massa gravitacional (ele raciocinou que fótons teriam massa $m = E/c^2$), antes de formular a Teoria da Relatividade Geral, quando abandonou essa idéia inconsistente, e uma coletânea de artigos criticando os conceitos de “massa de repouso” e “massa relativística” em [Okun2]). Por outro lado, massa é mais facilmente medida em um referencial em que o objeto está em repouso.

Observe que $m = 0$ se e somente se \mathbf{P} é um vetor do tipo luz, e consequentemente \mathbf{u} também é, donde $v = 1$. Portanto, fótons não possuem massa.

15.2 Definição. A componente temporal do quadrivetor momento-energia é chamada a **energia relativística** E do objeto e sua projeção espacial o seu **momento relativístico** \mathbf{p} . \square

Assim,

$$E = \frac{m}{\sqrt{1-v^2}},$$

$$\mathbf{p} = \frac{m}{\sqrt{1-v^2}} \mathbf{v} = \left(\frac{mv^x}{\sqrt{1-v^2}}, \frac{mv^y}{\sqrt{1-v^2}}, \frac{mv^z}{\sqrt{1-v^2}} \right),$$

de modo que

$$m^2 = E^2 - p^2. \quad (15.8)$$

Note que tanto a energia relativística quanto o momento relativístico dependem do referencial, já que são componentes (coordenadas) de um quadrivetor. Observe também que massa não é a mesma coisa que energia: energia é a componente temporal do quadrivetor momento-energia, enquanto que massa é a magnitude deste vetor. Apenas que, em relatividade, elas são medidas com as mesmas unidades de medida (quilogramas), da mesma forma que tempo e espaço em relatividade são medidos com as mesmas unidades de medida (metros), apesar de não serem a mesma coisa.

Em unidades do espaço-tempo (ct, x, y, z) em que a velocidade da luz é c e não normalizada para 1, a relação entre o tempo próprio e o tempo em um referencial inercial arbitrário é

$$\tau = \sqrt{1 - \left(\frac{v}{c}\right)^2} t$$

e o quadrivetor velocidade tem coordenadas

$$\mathbf{u} = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, v^x, v^y, v^z) = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, \mathbf{v}),$$

de modo que o momento-energia é

$$\mathbf{P} = \frac{m}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, v^x, v^y, v^z) = \frac{m}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (c, \mathbf{v}),$$

mas a energia relativística é

$$E = \frac{mc^2}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$$

(para que as unidades de medida de energia fiquem corretas) e o momento relativístico é

$$\mathbf{p} = \frac{mv}{\sqrt{1 - \left(\frac{v}{c}\right)^2}},$$

ou seja,

$$\mathbf{P} = \left(\frac{E}{c}, \mathbf{v} \right).$$

Quando $v = 0$, temos

$$E_0 = m$$

ou, quando a velocidade da luz não é normalizada, a célebre equação

$$E_0 = mc^2$$

para a energia de repouso; ou seja, a energia de um objeto não desaparece quando ele está em repouso, como na mecânica de Newton, mas ao contrário é enorme (por causa do fator c^2).

15.3 Fluxos

15.3.1 Definição

Fluxo (*flux*) é a taxa de transferência de um material (uma grandeza concreta tal como, por exemplo, massa, carga elétrica, número de partículas) ou propriedade (uma grandeza abstrata tal como, por exemplo, energia ou momento) ao longo de uma direção, por unidade de área, por unidade de tempo. Ele é portanto uma grandeza vetorial, cuja unidade é quantidade de material ou propriedade por unidade de área por unidade de tempo. Por exemplo, o material pode estar sendo transferido da esquerda para a direita no eixo x , ou na direção normal a uma certa superfície de referência. [Às vezes, fluxo por unidade de tempo é chamado simplesmente de *fluxo*, enquanto que fluxo por unidade de área por unidade de tempo é chamado de *densidade de fluxo*.]

15.3 Exemplo. No **fluxo de massa** o material transferido comumente se refere a um fluido (taxa de transferência de fluido, ou seja, taxa de escoamento) ou a uma certa substância dissolvida em um fluido. A taxa de transferência de massa pode ser medida, por exemplo, em $\text{kg}/\text{m}^2\text{s}$ ou, o que é equivalente no caso de um fluido de densidade constante, em $\text{litros}/\text{m}^2\text{s}$.

Um exemplo simples de **fluxo de energia** é a taxa de transferência de calor, que é energia térmica, medida em $\text{joules}/\text{m}^2\text{s}$.

No **fluxo de momento**, a taxa de transferência dada por unidades de momento por unidade de área por unidade de tempo, tem as mesmas unidades de força por área ($(\text{kg}\cdot\text{m}/\text{s})/\text{m}^2\text{s} = (\text{kg}\cdot\text{m}/\text{s}^2)/\text{m}^2$), ou seja, pressão. Assim, a força aplicada a um objeto pode ser vista como a taxa de transferência de momento para o objeto. O fluxo de momento também chamado estresse. \square

No que se segue, convencionaremos que o fluxo será positivo quando a transferência de material ou propriedade é no sentido orientado positivo do eixo x^i . Se o fluxo tem sinal negativo, isso significa que a transferência de material está-se realizando no sentido orientado negativo do eixo x^i .

15.4 Tensor Momento-Energia

Em cada evento do espaço-tempo existe um **tensor de momento-energia** T , que contém um conhecimento da densidade de energia, densidade de momento e tensão, associadas a todas as formas de matéria e todos os campos não-gravitacionais (como, por exemplo, o campo eletromagnético) medida por qualquer e todos os observadores naquele evento (já que é um tensor, logo sua definição e significado transcende todos os sistemas de coordenadas e referências iniciais). Mas qualquer observador, limitado pelo seu referencial inercial de Lorentz, presta mais atenção às componentes de T do que ao próprio T , atribuindo a cada componente um significado físico especial.

O tensor momento-energia é um 2-tensor contravariante (embora também possa ser visto como um 2-tensor covariante ou mesmo como um $(1,1)$ -tensor covariante e contravariante, através de contração de índices). O componente T^{ij} é o fluxo da i -ésima componente do momento-energia passando através da superfície $x^j \equiv \text{constante}$. Logo, T^{00} é a **densidade de energia** ($t = \text{constante}$), enquanto que T^{0j} para $i = 1, 2, 3$ é o fluxo da energia através de uma superfície no espaço. T^{i0} para $i = 1, 2, 3$ é a **densidade do momento** ($t = \text{constante}$). T^{ij} para $i, j = 1, 2, 3$ é o fluxo de momento através de uma superfície espacial, também chamado **tensão (stress)**: T^{ii} é a **tensão normal** (pressão normal à superfície) e T^{ij} , $i \neq j$, é a **tensão de cisalhamento** (pressão tangente à superfície). Assim, o tensor momento-energia representado em forma matricial é

$$T = \left[\begin{array}{c|ccc} T^{00} & T^{01} & T^{02} & T^{03} \\ \hline T^{10} & T^{11} & T^{12} & T^{13} \\ T^{20} & T^{21} & T^{22} & T^{23} \\ T^{30} & T^{31} & T^{32} & T^{33} \end{array} \right],$$

onde

$$T = \left[\begin{array}{cc|ccc} T^{00} & \text{densidade de energia} & T^{01} & T^{02} & T^{03} & \text{fluxo de energia} \\ \hline T^{10} & & T^{11} & T^{12} & T^{13} & \\ T^{20} & \text{densidade de momento} & T^{21} & T^{22} & T^{23} & \text{fluxo de momento} \\ T^{30} & & T^{31} & T^{32} & T^{33} & \end{array} \right].$$

15.4.1 Tensor Momento-Energia da Poeira

O tensor momento-energia mais simples é o **tensor momento-energia da poeira (dust)**. Poeira é simplesmente uma coleção de partículas em repouso em algum referencial (portanto elas se movem com a mesma velocidade em qualquer outro referencial), o seu referencial próprio. É um exemplo de um fluido perfeito (a ser definido no próximo parágrafo) com pressão nula.

A densidade do número de partículas (fluxo através das superfícies $t = \text{constante}$) é definida por

$$n = \text{número de partículas por unidade de volume no referencial próprio.}$$

Em um referencial inercial em que as partículas não estão em repouso, digamos em que as partículas tem velocidade v , o volume é alterado devido à contração de Lorentz: por exemplo, um paralelepípedo sólido de volume $\Delta x \Delta y \Delta z$ terá um volume $\Delta x \Delta y \Delta z \sqrt{1 - v^2}$ medido em um referencial em que as partículas tem velocidade v em uma das direções x, y, z . A densidade do número de partículas será então

$$\frac{n}{\sqrt{1 - v^2}}.$$

O **fluxo de partículas** é o número de partículas que atravessam uma superfície por unidade de área, por unidade de tempo. Além da orientação da superfície, em Relatividade área e tempo dependem do referencial. Obviamente, no referencial em que as partículas estão em repouso o fluxo é zero. Em um referencial em que as partículas movem-se com todas com a velocidade \mathbf{v} em uma mesma direção, digamos na direção x' , se considerarmos a superfície normal a x' a área será a mesma (já que não há contração nas direções transversais

ao movimento do referencial) $\Delta A = \Delta y \Delta z = \Delta y' \Delta z'$. Mas o volume que contém as partículas durante um intervalo $\Delta t'$ será

$$V = \Delta x' \Delta A = v \Delta t' \Delta A,$$

e como a densidade de partículas por unidade de volume neste referencial é $n/\sqrt{1-v^2}$ segue que o fluxo através de superfícies $x = \text{constante}$ por unidade de área por unidade de tempo será

$$\phi = \frac{nv}{\sqrt{1-v^2}}.$$

Mesmo que as partículas tenham componentes de velocidade nas direções y, z , como a componente do fluxo na direção x (através das superfícies $x = \text{constante}$) só depende da componente da velocidade ortogonal à superfície, isto é, v^x , segue que a componente do fluxo na direção x é dada por

$$\phi^x = \frac{nv^x}{\sqrt{1-v^2}}.$$

15.4 Definição. O quadrivetor fluxo de partículas é definido por

$$\mathbf{N} = n\mathbf{u}$$

□

Logo,

$$\mathbf{N} = \frac{n}{\sqrt{1-v^2}} (1, v^x, v^y, v^z) = \frac{n}{\sqrt{1-v^2}} (1, \mathbf{v}).$$

Note que a densidade do número de partículas no referencial próprio é a norma do quadrivetor fluxo e portanto independente do referencial inercial (o que também é óbvio da definição dela):

$$\langle \mathbf{N}, \mathbf{N} \rangle = \langle n\mathbf{u}, n\mathbf{u} \rangle = n^2 \langle \mathbf{u}, \mathbf{u} \rangle = -n^2,$$

de modo que

$$\|\mathbf{N}\| = n. \quad (15.9)$$

Para obter o tensor momento-energia da poeira, vamos assumir por simplicidade que as partículas tem a mesma massa m . No referencial comóvel temporário das partículas, em que elas se encontram em repouso, a energia de cada partícula é a energia de repouso m . Portanto, a energia por unidade de volume, ou seja, a *densidade de energia*, é

$$\rho = nm.$$

Em fluidos mais gerais, existe um movimento aleatório das partículas, logo energia cinética de movimento (mesmo em um referencial média dos referenciais) e a densidade de energia do fluido terá uma expressão bem diferente. Em um referencial em que as partículas se movem com a mesma velocidade v , a densidade de energia ainda é o número de partículas por unidade de volume vezes a energia de cada partícula, que é dada por

$$\rho = \frac{n}{\sqrt{1-v^2}} \frac{m}{\sqrt{1-v^2}} = \frac{nm}{1-v^2}.$$

Para obter o tensor momento-energia da poeira, basta determinar os seus componentes em algum referencial. No caso da poeira, a determinação dos seus componentes no seu referencial comóvel temporário é particularmente simples. Não há movimento das partículas, logo os momentos são todos nulos. Assim, no referencial comóvel temporário, o tensor momento-energia tem componentes

$$T = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Estas são exatamente as componentes do tensor

$$\mathbf{P} \otimes \mathbf{N} = (m\mathbf{u}) \otimes (n\mathbf{u}) = nm(\mathbf{u} \otimes \mathbf{u}) = \rho(\mathbf{u} \otimes \mathbf{u}),$$

pois $\mathbf{u} = (1, 0, 0, 0)$ no referencial comóvel temporário e $(\mathbf{F} \otimes \mathbf{G})^{ij} = F^i G^j$ para quaisquer quadrvetores (1-tensores contravariantes) \mathbf{F}, \mathbf{G} . Portanto, este é o tensor momento-energia da poeira:

$$T = \rho(\mathbf{u} \otimes \mathbf{u}).$$

Podemos agora calcular seus componentes em um referencial em que as partículas se movem com velocidade \mathbf{v} . Como

$$\begin{aligned} u^0 &= \frac{1}{\sqrt{1-v^2}}, \\ u^i &= \frac{v^i}{\sqrt{1-v^2}}, \end{aligned}$$

e

$$T^{ij} = \rho u^i u^j,$$

segue que

$$\begin{aligned} T^{00} &= \rho u^0 u^0 = \frac{\rho}{1-v^2}, \\ T^{0j} &= \rho u^0 u^j = \frac{\rho v^j}{1-v^2}, \\ T^{i0} &= \rho u^i u^0 = \frac{\rho v^i}{1-v^2}, \\ T^{ij} &= \rho u^i u^j = \frac{\rho v^i v^j}{1-v^2}, \quad \text{para } i, j \neq 0, \end{aligned}$$

e

$$T = \frac{\rho}{1-v^2} \begin{bmatrix} 1 & v^x & v^y & v^z \\ v^x & (v^x)^2 & v^x v^y & v^x v^z \\ v^y & v^x v^y & (v^y)^2 & v^y v^z \\ v^z & v^x v^z & v^y v^z & (v^z)^2 \end{bmatrix}.$$

Note que o tensor momento-energia da poeira é simétrico. Veremos que isso sempre acontece.

15.4.2 Simetria do Tensor Momento-Energia

Para provar que o tensor momento-energia é simétrico, isto é, que

$$T^{ij} = T^{ji}$$

para todos os índices i, j , basta provar isso em um referencial, pois isso implica

$$T(\omega, \sigma) = T(\sigma, \omega)$$

para todos os covetores ω, σ , o que será válido em todo referencial. O referencial mais simples de trabalhar é o referencial comóvel temporário.

Para ver que

$$T^{i0} = T^{0i},$$

ou seja, que a densidade de energia é igual ao fluxo de energia, observe que o fluxo de energia é igual à densidade de energia vezes a velocidade com que ela escoa. Mas como massa e energia são a mesma coisa, isto é a densidade de massa vezes a velocidade com que ela se move, ou seja a densidade de momento.

A simetria

$$T^{ij} = T^{ji},$$

para $i, j = 1, 2, 3$ é deduzida como em Mecânica Newtoniana (veja, por exemplo, [Shutz], pp. 97-98).

15.4.3 Conservação do Momento-Energia

Pela lei de conservação do momento-energia segue que

$$\operatorname{div} T = 0,$$

já que o divergente mede infinitesimalmente o fluxo total por unidade de volume (veja o apêndice a este capítulo).

15.4.4 Tensor Momento-Energia de um Fluido Perfeito

Um **fluido perfeito** é um fluido em que não há condução de calor ou viscosidade. A primeira condição implica que o fluxo da energia através de uma superfície no espaço é zero, logo T^{0j} no referencial comóvel temporário e segunda condição implica que não há tensão de cisalhamento, ou seja, $T^{ij} = 0$ para $i \neq j$, $i, j = 1, 2, 3$ neste referencial. Em outras palavras, um fluido perfeito é caracterizado apenas por densidade e pressão; o exemplo mais familiar de fluido perfeito é um gás perfeito. Usando simetria, concluímos que no referencial comóvel temporário o tensor momento-energia de um fluido perfeito tem a forma

$$T = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{bmatrix},$$

onde P é a pressão. É fácil ver que

$$T = (\rho + P)(\mathbf{u} \otimes \mathbf{u}) + P\eta,$$

onde η é o tensor métrica de Lorentz do espaço de Minkowski. Portanto, em um referencial em que o fluido se move com velocidade \mathbf{v} temos

$$T = \frac{\rho + P}{1 - v^2} \begin{bmatrix} 1 & v^x & v^y & v^z \\ v^x & (v^x)^2 & v^x v^y & v^x v^z \\ v^y & v^x v^y & (v^y)^2 & v^y v^z \\ v^z & v^x v^z & v^y v^z & (v^z)^2 \end{bmatrix} + P \begin{bmatrix} \tilde{\eta}_{00} & \tilde{\eta}_{01} & \tilde{\eta}_{02} & \tilde{\eta}_{03} \\ \tilde{\eta}_{10} & \tilde{\eta}_{11} & \tilde{\eta}_{12} & \tilde{\eta}_{13} \\ \tilde{\eta}_{20} & \tilde{\eta}_{21} & \tilde{\eta}_{22} & \tilde{\eta}_{23} \\ \tilde{\eta}_{30} & \tilde{\eta}_{31} & \tilde{\eta}_{32} & \tilde{\eta}_{33} \end{bmatrix},$$

onde $\tilde{\eta}_{ij}$ são os componentes da métrica de Lorentz no referencial.

Para qualquer métrica de Lorentz \mathbf{g} em um espaço-tempo curvo teremos

$$T = (\rho + P)(\mathbf{u} \otimes \mathbf{u}) + P\mathbf{g}.$$

15.5 Apêndice: Divergente e Fluxo

É bastante esclarecedor estudar o escoamento de um fluido para entender a matemática do fluxo. Para um fluido movendo-se no espaço, denote por $\mathbf{V} = (V^x, V^y, V^z)$ o seu **campo de velocidades**. Para cada ponto (x, y, z) no espaço e para cada instante de tempo t , $\mathbf{V}(t, x, y, z)$ denota a velocidade do fluido ao passar por aquele ponto naquele instante de tempo. De fato, fixado um ponto no espaço, a velocidade do fluido ao passar por aquele ponto pode ser variável no tempo. Quando isso não ocorre, isto é, se a velocidade não depende do tempo, $\mathbf{V} = \mathbf{V}(x, y, z)$, dizemos que o campo é **estacionário** (ou que o fluido está em *regime permanente*). Mais geralmente

$$V^x = V^x(t, x, y, z),$$

$$V^y = V^y(t, x, y, z),$$

$$V^z = V^z(t, x, y, z),$$

e $\mathbf{V} : \Omega \times I \rightarrow \mathbb{R}^3$ é uma função vetorial definida no produto cartesiano $\Omega \times I$, onde $\Omega \subset \mathbb{R}^3$ é uma região aberta do espaço e I é um intervalo contendo 0 (que pode ser simplesmente o extremo esquerdo deste intervalo).

Denote por $\mathbf{F}(\mathbf{x}, t)$ o **campo de escoamento** (*flow field*) de um campo de velocidades \mathbf{V} de um fluido. Por definição, isso significa que

$$\begin{cases} \frac{d\mathbf{F}}{dt}(t, \mathbf{x}) = \mathbf{V}(t, \mathbf{F}(t, \mathbf{x})), \\ \mathbf{F}(0, \mathbf{x}) = \mathbf{x}. \end{cases}$$

Em outras palavras, fixado $\mathbf{x} \in \Omega$, $\mathbf{F}(t, \mathbf{x})$ é a **trajetória** do campo de escoamento que passa por \mathbf{x} no instante $t = 0$: se uma partícula é posicionada dentro do fluido na posição \mathbf{x} , a sua posição após haver decorrido um intervalo de tempo t é dada por $\mathbf{F}(t, \mathbf{x})$.

Para um campo de velocidades estacionário, as trajetórias do campo de escoamento são fixas no espaço, não se modificando com o tempo. Assim, se uma partícula for posicionada na posição \mathbf{x} no instante de tempo t_1 e, logo em seguida, no instante de tempo t_2 uma segunda partícula é posicionada na mesma posição \mathbf{x} , as duas partículas seguirão a mesma trajetória, ou seja, decorrido um intervalo de tempo Δt após o instante de tempo de posicionamento inicial de cada partícula, elas ocuparão a mesma posição $\mathbf{F}(t, \mathbf{x})$ no espaço. Para um campo de velocidades não-estacionário, as trajetórias do campo não têm uma forma fixa, variando com o tempo: em cada instante de tempo t , o campo de escoamento tem uma forma diferente. Desta vez, se uma partícula for posicionada na posição \mathbf{x} no instante de tempo t_1 , a velocidade do fluido naquele instante é $\mathbf{V}(t_1, \mathbf{x})$ e a partícula seguirá uma certa trajetória que tem esta velocidade inicial, enquanto que, se logo em seguida no instante de tempo t_2 uma segunda partícula é posicionada na mesma posição \mathbf{x} , a velocidade do campo naquele momento será $\mathbf{V}(t_2, \mathbf{x})$, possivelmente diferente da velocidade inicial da primeira partícula, e muito provavelmente as duas partículas seguirão trajetórias completamente diferentes.

É importante não confundir o conceito de trajetórias do campo de escoamento com o conceito de *linhas de corrente* do escoamento. As **linhas de corrente** (*streamlines*) de um escoamento são as curvas tangentes ao campo de velocidades do escoamento. É claro que no caso de um campo estacionário as linhas de corrente são exatamente as trajetórias do campo. No caso de um campo não-estacionário, entretanto, como o vetor velocidade varia com o tempo, a trajetória que uma partícula seguirá após ser posicionada em um ponto dentro do fluido em um determinado instante de tempo em geral não coincidirá com a linha de corrente do escoamento que passa por aquele ponto naquele momento. Para um campo de velocidades não-estacionário, as linhas de corrente variam em forma com o passar do tempo.

Outra maneira de enxergar o campo de escoamento é fixar t e considerar a função $\mathbf{F}_t : \Omega \rightarrow \Omega_t$, onde $\Omega_t = \mathbf{F}_t(\Omega)$. $\mathbf{F}_t(\mathbf{x})$ dá a posição da partícula que ocupava a posição \mathbf{x} no instante inicial após decorrido um intervalo de tempo t . A função \mathbf{F}_t pode ser vista como a função que nos permite ver como regiões do fluido são deformadas através do escoamento: uma região $\Omega_0 \subset \Omega$ do fluido é deformada pelo escoamento, de modo que após decorrido um intervalo de tempo t ela é transformada na região $\mathbf{F}_t(\Omega_0)$. Sob condições muito gerais, esta transformação é um difeomorfismo (a inversa é diferenciável porque basta seguir as trajetórias no sentido reverso para de $\mathbf{F}_t(\Omega_0)$ voltar em Ω_0).

As trajetórias de um campo de escoamento são encontradas resolvendo-se a equação diferencial ordinária com condição inicial acima (é claro que, sendo \mathbf{F} e \mathbf{V} grandezas vetoriais no espaço, esta equação diferencial ordinária é na verdade um sistema de 3 equações diferenciais com 3 condições iniciais). Desde que \mathbf{V} satisfaça certas condições razoáveis de regularidade (por exemplo, se \mathbf{V} for continuamente diferenciável), a equação diferencial ordinária acima sempre tem uma única solução, pelo menos localmente. Portanto, dado \mathbf{V} , podemos encontrar \mathbf{F} . Em geral, não é possível encontrar \mathbf{F} explicitamente (a função \mathbf{V} pode ser muito complicada), mas podemos obtê-lo aproximadamente através de métodos numéricos.

Reciprocamente, conhecendo-se o campo de escoamento \mathbf{F} , é possível encontrar o campo de velocidades \mathbf{V} . Essencialmente, basta derivar \mathbf{F} em relação ao tempo:

$$\mathbf{V}(t, \mathbf{x}) = \frac{d\mathbf{F}}{dt}(t, \mathbf{F}_t^{-1}(\mathbf{x})).$$

Mais especificamente, $\mathbf{V}(t, \mathbf{x})$ é a derivada temporal da trajetória do campo que passa pelo ponto $\mathbf{F}_t^{-1}(\mathbf{x})$

no instante inicial $t = 0$, esta derivada sendo calculada no instante de tempo t . Se o campo de escoamento é dado através de uma expressão fechada (e não por uma expressão aproximada), então podemos conhecer o campo de velocidades com exatidão. Obviamente, é mais difícil encontrar \mathbf{F} , dado \mathbf{V} , do que encontrar \mathbf{V} , dado \mathbf{F} .

15.5.1 Significado Geométrico do Divergente Espacial

Nossos campos vetoriais em geral serão funções dependendo de três coordenadas espaciais e uma temporal: $\mathbf{V} = \mathbf{V}(t, x, y, z)$. Considere o **divergente espacial** de um campo vetorial, isto é, o divergente em relação às três coordenadas espaciais, sem envolver a coordenada temporal,

$$\operatorname{div}_{\mathbf{x}} \mathbf{V} = \frac{\partial V^x}{\partial x} + \frac{\partial V^y}{\partial y} + \frac{\partial V^z}{\partial z}.$$

Defina

$$\mathcal{V}(t, \mathbf{x}) = \det d\mathbf{F}_t(\mathbf{x}),$$

isto é,

$$d\mathbf{F}_t(\mathbf{x}) = \begin{bmatrix} \frac{\partial F_t^1(\mathbf{x})}{\partial x} & \frac{\partial F_t^1(\mathbf{x})}{\partial y} & \frac{\partial F_t^1(\mathbf{x})}{\partial z} \\ \frac{\partial F_t^2(\mathbf{x})}{\partial x} & \frac{\partial F_t^2(\mathbf{x})}{\partial y} & \frac{\partial F_t^2(\mathbf{x})}{\partial z} \\ \frac{\partial F_t^3(\mathbf{x})}{\partial x} & \frac{\partial F_t^3(\mathbf{x})}{\partial y} & \frac{\partial F_t^3(\mathbf{x})}{\partial z} \end{bmatrix},$$

onde escrevemos $\mathbf{F}_t(\mathbf{x}) = (F_t^1(\mathbf{x}), F_t^2(\mathbf{x}), F_t^3(\mathbf{x}))$. Intuitivamente, $\mathcal{V}(t, \mathbf{x})$ mede o volume de um elemento infinitesimal do fluido em diferentes pontos do escoamento (produto misto de três vetores no espaço). Este volume varia à medida que o elemento move-se ao longo de uma trajetória do escoamento.

Para um campo de velocidades estacionário $\mathbf{V} = \mathbf{V}(\mathbf{x})$, esta variação infinitesimal por unidade de volume é dada precisamente pelo divergente espacial $\operatorname{div}_{\mathbf{x}} \mathbf{V}$ que neste caso de um campo estacionário é simplesmente o divergente $\operatorname{div} \mathbf{V}$: o divergente do campo de velocidades estacionário de um fluido é a taxa de variação do volume de um elemento do fluido que se move ao longo de uma trajetória do campo, por unidade de volume.

15.5 Teorema. *Se $\mathbf{V} = \mathbf{V}(\mathbf{x})$ é um campo vetorial estacionário, então*

$$\operatorname{div} \mathbf{V}(\mathbf{x}) = \frac{1}{\mathcal{V}(0)} \left. \frac{d\mathcal{V}(t)}{dt} \right|_{t=0}.$$

Mais geralmente,

$$\operatorname{div} \mathbf{V}(\mathbf{F}_t(\mathbf{x})) = \frac{1}{\det d(\mathbf{F}_t)} \frac{d}{dt} \det d(\mathbf{F}_t)_{\mathbf{x}}.$$

Prova 1: Consideraremos primeiro uma demonstração elementar e bastante intuitiva da primeira equação.

Fixe um ponto \mathbf{x} no espaço. Considere $\varepsilon > 0$ pequeno e sejam $\mathbf{v}_1 = \varepsilon \mathbf{i}, \mathbf{v}_2 = \varepsilon \mathbf{j}, \mathbf{v}_3 = \varepsilon \mathbf{k}$, onde $\mathbf{i}, \mathbf{j}, \mathbf{k}$ são os vetores da base canônica de \mathbb{R}^3 . O ponto \mathbf{x} e os três vetores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ definem um paralelepípedo $P(0)$ no espaço com um de seus vértices localizado no ponto \mathbf{x} e os outros vértices localizados nos pontos $\mathbf{x} + \mathbf{v}_1, \mathbf{x} + \mathbf{v}_2$ e $\mathbf{x} + \mathbf{v}_3$. Para cada t fixado, o escoamento \mathbf{F}_t é uma função de \mathbb{R}^3 em \mathbb{R}^3 , ou seja, uma função que leva pontos do espaço em outros pontos do espaço ao longo das trajetórias do escoamento; o efeito visual de \mathbf{F}_t é que cada ponto se move ao longo de sua trajetória enquanto decorre o intervalo de tempo t , após o que ele pára. Em particular, cada ponto do paralelepípedo move-se ao longo da trajetória à qual ele pertence, distorcendo o paralelepípedo no processo, de modo que $\mathbf{F}_t(P(0))$ não é mais, em geral, um paralelepípedo. Por outro lado, se ε é pequeno (ou seja, vendo o paralelepípedo $P(0)$ como um elemento infinitesimal de volume aproximado), a imagem de $P(0)$ sob \mathbf{F}_t pode ser aproximada pela sua imagem sob a derivada $d\mathbf{F}_t$ de \mathbf{F}_t : de fato, se $\mathbf{v} = (\mathbf{x} + \Delta\mathbf{x}) - \mathbf{x} = \Delta\mathbf{x}$ é um pequeno vetor baseado em \mathbf{x} , temos

$$\mathbf{F}_t(\mathbf{x} + \Delta\mathbf{x}) - \mathbf{F}_t(\mathbf{x}) \approx d\mathbf{F}_t(\mathbf{x}) \Delta\mathbf{x} = d\mathbf{F}_t(\mathbf{x}) \mathbf{v},$$

pois, por definição de derivada direcional,

$$d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v} = \frac{\partial \mathbf{F}_t}{\partial \mathbf{v}} = \lim_{s \rightarrow 0} \frac{\mathbf{F}_t(\mathbf{x} + s\Delta\mathbf{x}) - \mathbf{F}_t(\mathbf{x})}{s}.$$

Portanto, se ε é pequeno, o paralelepípedo $P(0)$ é levado pelo escoamento do campo depois de decorrido um intervalo de tempo t no paralelepípedo $P(t)$ com um vértice em $\mathbf{F}_t(\mathbf{x})$ e os outros vértices localizados nos pontos $\mathbf{F}_t(\mathbf{x}) + \mathbf{v}_1(t)$, $\mathbf{F}_t(\mathbf{x}) + \mathbf{v}_2(t)$ e $\mathbf{F}_t(\mathbf{x}) + \mathbf{v}_3(t)$, onde

$$\begin{aligned}\mathbf{v}_1(t) &= d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v}_1, \\ \mathbf{v}_2(t) &= d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v}_2, \\ \mathbf{v}_3(t) &= d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v}_3.\end{aligned}$$

O volume do paralelepípedo $P(t)$ é dado pelo produto misto:

$$\mathcal{V}(t) = \mathbf{v}_1(t) \cdot [\mathbf{v}_2(t) \times \mathbf{v}_3(t)].$$

Logo, pela regra da derivada do produto e a propriedade de comutatividade do produto escalar com o produto vetorial, segue que

$$\begin{aligned}\frac{d\mathcal{V}}{dt} &= \frac{d\mathbf{v}_1}{dt} \cdot (\mathbf{v}_2 \times \mathbf{v}_3) + \mathbf{v}_1 \cdot \left(\frac{d\mathbf{v}_2}{dt} \times \mathbf{v}_3 \right) + \mathbf{v}_1 \cdot \left(\mathbf{v}_2 \times \frac{d\mathbf{v}_3}{dt} \right) \\ &= \frac{d\mathbf{v}_1}{dt} \cdot (\mathbf{v}_2 \times \mathbf{v}_3) + \frac{d\mathbf{v}_2}{dt} \cdot (\mathbf{v}_3 \times \mathbf{v}_1) + \frac{d\mathbf{v}_3}{dt} \cdot (\mathbf{v}_1 \times \mathbf{v}_2).\end{aligned}$$

Em $t = 0$, temos

$$\begin{aligned}\mathbf{v}_2 \times \mathbf{v}_3 &= \varepsilon^2 \mathbf{i}, \\ \mathbf{v}_3 \times \mathbf{v}_1 &= \varepsilon^2 \mathbf{j}, \\ \mathbf{v}_1 \times \mathbf{v}_2 &= \varepsilon^2 \mathbf{k}.\end{aligned}$$

Para calcular $\frac{d\mathbf{v}_i(t)}{dt}$, lembramos a definição do campo de escoamento $\mathbf{F}(t, \mathbf{x})$ do campo de velocidades $\mathbf{V}(\mathbf{x})$:

$$\frac{d\mathbf{F}_t}{dt}(\mathbf{x}) = \mathbf{V}(\mathbf{F}_t(\mathbf{x})).$$

Derivando com relação a \mathbf{x} , temos

$$d \left[\frac{d\mathbf{F}_t}{dt} \right] (\mathbf{x}) = d[\mathbf{V}(\mathbf{F}_t)]_{\mathbf{x}}.$$

Invertendo a ordem de derivação na derivada parcial mista do lado esquerdo e aplicando a regra da cadeia no lado direito, temos

$$\frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}] = d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}. \quad (15.10)$$

Aplicando ambos os lados desta equação ao vetor $\mathbf{v}_i = \mathbf{v}_i(0)$ obtemos no lado esquerdo (pois \mathbf{v}_i é um vetor constante)

$$\left(\frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}] \right) \mathbf{v}_i = \frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v}_i] = \frac{d\mathbf{v}_i}{dt}(t),$$

enquanto que no lado direito obtemos

$$d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}} \mathbf{v}_i = d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} \mathbf{v}_i(t).$$

Segue que

$$\frac{d\mathbf{v}_i}{dt}(t) = d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})}\mathbf{v}_i(t). \quad (15.11)$$

Em particular, em $t = 0$ temos $\mathbf{F}_0(\mathbf{x}) = \mathbf{x}$, donde

$$\left. \frac{d\mathbf{v}_i}{dt} \right|_{t=0} = d\mathbf{V}_{\mathbf{x}}\mathbf{v}_i,$$

Estamos agora em condições de concluir a demonstração do teorema:

$$\begin{aligned} \left. \frac{d\mathcal{V}(t)}{dt} \right|_{t=0} &= \varepsilon^2 [(d\mathbf{V}(\mathbf{x})(\varepsilon\mathbf{i})) \cdot \mathbf{i} + (d\mathbf{V}(\mathbf{x})(\varepsilon\mathbf{j})) \cdot \mathbf{j} + (d\mathbf{V}(\mathbf{x})(\varepsilon\mathbf{k})) \cdot \mathbf{k}] \\ &= \varepsilon^3 [(d\mathbf{V}(\mathbf{x})\mathbf{i}) \cdot \mathbf{i} + (d\mathbf{V}(\mathbf{x})\mathbf{j}) \cdot \mathbf{j} + (d\mathbf{V}(\mathbf{x})\mathbf{k}) \cdot \mathbf{k}] \\ &= \varepsilon^3 \left[\frac{\partial V^x}{\partial x} + \frac{\partial V^y}{\partial y} + \frac{\partial V^z}{\partial z} \right] \\ &= \mathcal{V}(0) \operatorname{div} \mathbf{V}(\mathbf{x}). \end{aligned}$$

Prova 2: Mostraremos que

$$\frac{d}{dt} \det d(\mathbf{F}_t)_{\mathbf{x}} = \operatorname{div} \mathbf{V}(\mathbf{F}_t(\mathbf{x})) \det d(\mathbf{F}_t)_{\mathbf{x}}.$$

Pela regra da cadeia,

$$\frac{d}{dt} \det d(\mathbf{F}_t)_{\mathbf{x}} = d(\det)_{d(\mathbf{F}_t)_{\mathbf{x}}} \frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}].$$

Lembremos a derivada da função determinante $\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$. Escrevendo $\det X = \det(X_1, X_2, X_3)$, onde X_1, X_2, X_3 são as linhas da matriz X , a derivada do determinante em X é o funcional linear

$$d(\det)_X : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$$

cujo valor na matriz $A = (A_1, A_2, A_3)$ é

$$d(\det)_X A = \det(A_1, X_2, X_3) + \det(X_1, A_2, X_3) + \det(X_1, X_2, A_3).$$

Segue que

$$d(\det)_{d(\mathbf{F}_t)_{\mathbf{x}}} A = \det([d(\mathbf{F}_t)_{\mathbf{x}}]_1, X_2, X_3) + \det(X_1, [d(\mathbf{F}_t)_{\mathbf{x}}]_2, X_3) + \det(X_1, X_2, [d(\mathbf{F}_t)_{\mathbf{x}}]_3)$$

onde $[d(\mathbf{F}_t)_{\mathbf{x}}]_i$ denota a i -ésima linha da matriz $d(\mathbf{F}_t)_{\mathbf{x}}$. Por outro lado, pela comutatividade dos operadores diferenciais d e $\frac{d}{dt}$ e pela regra da cadeia, temos

$$\frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}] = d \left[\frac{d\mathbf{F}_t}{dt}(\mathbf{x}) \right] = d[\mathbf{V}(\mathbf{F}_t(\mathbf{x}))] = d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}.$$

Logo,

$$\begin{aligned} d(\det)_{d(\mathbf{F}_t)_{\mathbf{x}}} \frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}] &= \det ([d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d(\mathbf{F}_t)_{\mathbf{x}}]_3) \\ &+ \det ([d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d(\mathbf{F}_t)_{\mathbf{x}}]_3) \\ &+ \det ([d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_3). \end{aligned}$$

onde $[d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_i$ denota a i -ésima linha da matriz $d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}$. Daí, usando a linearidade da função determinante com relação a uma linha fixada, temos

$$\begin{aligned}
 & \det([d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d(\mathbf{F}_t)_{\mathbf{x}}]_3) \\
 &= \det \left[\begin{array}{ccc} \sum_{j=1}^3 \frac{\partial V^x}{\partial x^j} \frac{\partial F_t^j}{\partial x} & \sum_{j=1}^3 \frac{\partial V^x}{\partial x^j} \frac{\partial F_t^j}{\partial y} & \sum_{j=1}^3 \frac{\partial V^x}{\partial x^j} \frac{\partial F_t^j}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] \\
 &= \det \left[\begin{array}{ccc} \frac{\partial V^x}{\partial x} \frac{\partial F_t^1}{\partial x} & \frac{\partial V^x}{\partial x} \frac{\partial F_t^1}{\partial y} & \frac{\partial V^x}{\partial x} \frac{\partial F_t^1}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] + \det \left[\begin{array}{ccc} \frac{\partial V^x}{\partial y} \frac{\partial F_t^2}{\partial x} & \frac{\partial V^x}{\partial y} \frac{\partial F_t^2}{\partial y} & \frac{\partial V^x}{\partial y} \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] \\
 &+ \det \left[\begin{array}{ccc} \frac{\partial V^x}{\partial z} \frac{\partial F_t^3}{\partial x} & \frac{\partial V^x}{\partial z} \frac{\partial F_t^3}{\partial y} & \frac{\partial V^x}{\partial z} \frac{\partial F_t^3}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] \\
 &= \frac{\partial V^x}{\partial x} \det \left[\begin{array}{ccc} \frac{\partial F_t^1}{\partial x} & \frac{\partial F_t^1}{\partial y} & \frac{\partial F_t^1}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] + \frac{\partial V^x}{\partial y} \det \left[\begin{array}{ccc} \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] + \frac{\partial V^x}{\partial z} \det \left[\begin{array}{ccc} \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \\ \frac{\partial F_t^2}{\partial x} & \frac{\partial F_t^2}{\partial y} & \frac{\partial F_t^2}{\partial z} \\ \frac{\partial F_t^3}{\partial x} & \frac{\partial F_t^3}{\partial y} & \frac{\partial F_t^3}{\partial z} \end{array} \right] \\
 &= \frac{\partial V^x}{\partial x} \det d(\mathbf{F}_t)_{\mathbf{x}}.
 \end{aligned}$$

Similarmente,

$$\begin{aligned}
 \det([d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d(\mathbf{F}_t)_{\mathbf{x}}]_3) &= \frac{\partial V^y}{\partial y} \det d(\mathbf{F}_t)_{\mathbf{x}}, \\
 \det([d(\mathbf{F}_t)_{\mathbf{x}}]_1, [d(\mathbf{F}_t)_{\mathbf{x}}]_2, [d\mathbf{V}_{\mathbf{F}_t(\mathbf{x})} d(\mathbf{F}_t)_{\mathbf{x}}]_3) &= \frac{\partial V^z}{\partial z} \det d(\mathbf{F}_t)_{\mathbf{x}}.
 \end{aligned}$$

Portanto,

$$\begin{aligned}
 d(\det)_{d(\mathbf{F}_t)_{\mathbf{x}}} \frac{d}{dt} [d(\mathbf{F}_t)_{\mathbf{x}}] &= \left(\frac{\partial V^x}{\partial x} \mathbf{F}_t(\mathbf{x}) + \frac{\partial V^y}{\partial y} \mathbf{F}_t(\mathbf{x}) + \frac{\partial V^z}{\partial z} \mathbf{F}_t(\mathbf{x}) \right) \det d(\mathbf{F}_t)_{\mathbf{x}} \\
 &= \operatorname{div} \mathbf{V}(\mathbf{F}_t(\mathbf{x})) \det d(\mathbf{F}_t)_{\mathbf{x}}.
 \end{aligned}$$

15.5.2 Significado Físico do Divergente

Pelo Teorema de Gauss,

$$\int_{\Omega} \operatorname{div} \mathbf{V} = \int_{\partial\Omega} \mathbf{V} \cdot \eta,$$

onde, por definição, o lado direito é o **fluxo total** do campo \mathbf{V} através da superfície $\partial\Omega$. Em particular, se $\operatorname{div} \mathbf{V} = 0$ em um aberto de \mathbb{R}^n , o fluxo total é nulo através de qualquer superfície fechada contida no aberto. Pelo Teorema do Valor Médio para Integrais,

$$\frac{1}{\operatorname{vol}\Omega} \int_{\Omega} \operatorname{div} \mathbf{V} = \operatorname{div} \mathbf{V}(\mathbf{x})$$

para algum $\mathbf{x} \in \Omega$. Portanto,

$$\operatorname{div} \mathbf{V}(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\operatorname{vol} B_{\varepsilon}(\mathbf{x})} \int_{B_{\varepsilon}(\mathbf{x})} \operatorname{div} \mathbf{V}.$$

Consequentemente,

$$\operatorname{div} \mathbf{V}(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\operatorname{vol} B_{\varepsilon}(\mathbf{x})} \int_{S_{\varepsilon}(\mathbf{x})} \mathbf{V} \cdot \eta.$$

Assim, o divergente mede infinitesimalmente o fluxo total por unidade de volume.

Capítulo 16

Equação de Einstein

16.1 Equação de Newton para o Potencial Gravitacional

O análogo Newtoniano da equação de Einstein é a equação de Poisson para o potencial gravitacional ϕ

$$\Delta\phi = 4\pi G\rho, \quad (16.1)$$

onde ρ é a densidade de massa e G denota a constante gravitacional de Newton. Em Mecânica Newtoniana, a gravidade é gerada apenas por massa.

Vamos ver como esta equação é deduzida. No caso de uma única fonte puntual de massa M localizada na origem 0, é comprovado experimentalmente que a força gravitacional atraiendo uma partícula teste de massa m é

$$\mathbf{F} = m\mathbf{g} = -\frac{GMm}{r^2}\hat{\mathbf{r}},$$

onde \mathbf{g} é a aceleração da gravidade, \mathbf{r} é o vetor posição radial partindo da fonte e $\hat{\mathbf{r}}$ denota o seu vetor unitário. O campo gravitacional \mathbf{G} , isto é, a força gravitacional agindo em um certo ponto do espaço localizado em \mathbf{r} por unidade de massa, é portanto dado por

$$\mathbf{G} = -\frac{GM}{r^2}\hat{\mathbf{r}}.$$

[Note que \mathbf{G} e G denotam conceitos completamente diferentes, não relacionados.] Assim, pela segunda lei de Newton, o campo gravitacional \mathbf{G} por definição nada mais é que a aceleração da gravidade \mathbf{g} :

$$\mathbf{G} = \frac{\mathbf{F}}{m} = \mathbf{g}.$$

Considerando um aberto Ω contendo a fonte 0 tal que sua fronteira é uma superfície regular fechada $\partial\Omega$, segue que

$$\int_{\partial\Omega} \mathbf{G} \cdot \mathbf{n} ds = -GM \int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} ds.$$

16.1 Lema. Se $\Omega \subset \mathbb{R}^3$ é um aberto tal que sua fronteira $\partial\Omega$ é uma superfície regular fechada e $0 \notin \partial\Omega$, vale

$$\int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} ds = \begin{cases} 4\pi & \text{se } 0 \in \Omega, \\ 0 & \text{se } 0 \notin \Omega. \end{cases}$$

Prova: Primeiro observamos que

$$\operatorname{div} \frac{\hat{\mathbf{r}}}{r^2} = 0.$$

De fato,

$$\frac{\hat{\mathbf{r}}}{r^2} = \frac{x}{\|x\|^3} = \frac{(x^1, x^2, x^3)}{\left[\sum_{i=1}^3 (x^i)^2\right]^{3/2}},$$

de modo que

$$\begin{aligned}\partial_i \left[\left(\frac{\hat{\mathbf{r}}}{r^2} \right)^i \right] &= \partial_i \left(\frac{x^i}{\left[\sum_{i=1}^3 (x^i)^2 \right]^{3/2}} \right) = \frac{\left[\sum_{i=1}^3 (x^i)^2 \right]^{3/2} - 3(x^i)^2 \left[\sum_{i=1}^3 (x^i)^2 \right]^{1/2}}{\left[\sum_{i=1}^3 (x^i)^2 \right]^{5/2}} \\ &= \frac{1}{\|x\|^2} - \frac{3(x^i)^2}{\|x\|^4},\end{aligned}$$

logo

$$\operatorname{div} \frac{\hat{\mathbf{r}}}{r^2} = \sum_{i=1}^3 \partial_i \left(\frac{\hat{\mathbf{r}}}{r^2} \right)^i = \frac{3}{\|x\|^2} - 3 \frac{\sum_{i=1}^3 (x^i)^2}{\|x\|^4} = 0.$$

Considere primeiro o caso $0 \notin \Omega$, de modo que o campo vetorial $\frac{1}{r^2} \hat{\mathbf{r}}$ é um campo suave em $\bar{\Omega}$. Podemos então aplicar o Teorema da Divergência para obter

$$\int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds = \int_{\Omega} \operatorname{div} \frac{\hat{\mathbf{r}}}{r^2} \, dv = 0.$$

Se $0 \in \Omega$, o campo vetorial $\frac{1}{r^2} \hat{\mathbf{r}}$ tem uma singularidade em 0 e o Teorema da Divergência não se aplica à região $\bar{\Omega}$. Considere então uma esfera de raio R centrada em 0 e contida em Ω . Então o aberto $\Omega \setminus \overline{B_R(0)}$ cuja fronteira é $\partial\Omega \cup S_R(0)$ não contém a singularidade 0, logo segue do resultado anterior que

$$\int_{\partial\Omega \cup S_R(0)} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds = 0.$$

Como

$$\int_{\partial\Omega \cup S_R(0)} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds = \int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds + \int_{S_R(0)} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds,$$

temos que

$$\int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds = - \int_{S_R(0)} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds.$$

Nesta última integral, para manter a convenção de \mathbf{n} ser o vetor normal apontando para fora da superfície em relação ao aberto que ela engloba, o vetor \mathbf{n} aponta para dentro da esfera, isto é,

$$\mathbf{n} = -\hat{\mathbf{r}},$$

de modo que

$$\int_{\partial\Omega} \frac{1}{r^2} \hat{\mathbf{r}} \cdot \mathbf{n} \, ds = \frac{1}{R^2} \int_{\partial\Omega} \, ds = \frac{1}{R^2} 4\pi R^2 = 4\pi.$$

■

Segue do Lema 16.1 que

$$\int_{\partial\Omega} \mathbf{G} \cdot \mathbf{n} \, ds = -4\pi GM.$$

Também segue dela que se Ω é um aberto tal que $\bar{\Omega}$ não contém a fonte, teremos

$$\int_{\partial\Omega} \mathbf{G} \cdot \mathbf{n} ds = 0.$$

Isso significa que cada partícula de massa M dentro da região Ω contribui para a integral de superfície o valor $-4\pi GM$, enquanto que partículas fora da região Ω não contribuem nada. Portanto, o valor da integral é $-4\pi G$ vezes a massa total dentro da região Ω . Para uma distribuição contínua de massa com densidade de massa $\rho(x)$, a massa total dentro da região Ω é dada por

$$\int_{\Omega} \rho dv.$$

Consequentemente,

$$\int_{\partial\Omega} \mathbf{G} \cdot \mathbf{n} ds = -4\pi G \int_{\Omega} \rho dv$$

Mas, pelo Teorema da Divergência, também temos que

$$\int_{\partial\Omega} \mathbf{G} \cdot \mathbf{n} ds = \int_{\Omega} \operatorname{div} \mathbf{G} dv,$$

logo,

$$\int_{\Omega} [\operatorname{div} \mathbf{G} + 4\pi G \rho] dv = 0.$$

Como isso é válido para qualquer domínio Ω com fronteira regular arbitrário, segue que

$$\operatorname{div} \mathbf{G} = -4\pi G \rho.$$

Mas o potencial gravitacional é definido por

$$\mathbf{G} = -\nabla \phi$$

(como o campo gravitacional é conservativo, isto é, satisfaz $\operatorname{rot} \mathbf{G} = 0$, a existência de ϕ está assegurada), donde

$$\operatorname{div}(\nabla \phi) = 4\pi G \rho,$$

e daí

$$\Delta \phi = 4\pi G \rho. \tag{16.2}$$

16.2 Dedução da Equação de Einstein

A primeira tentativa de Einstein em obter uma equação equivalente para a sua nova teoria de gravitação foi simplesmente substituir o potencial gravitacional ϕ por g_{ij} . O problema é que $\nabla g = 0$, logo $\Delta g = 0$ também. Este motivo levou Einstein a abandonar a busca por uma equação covariante, contentando-se em uma equação invariante apenas por transformações de coordenadas especiais (veja mais detalhes em [Pais], p. 221; para a trajetória completa de Einstein até chegar na equação correta, veja na mesma referência citada os Capítulos 12-14). Ele levou alguns anos até chegar na equação (16.4) que veremos a seguir.

Na Teoria da Relatividade Geral de Einstein, gravidade é simplesmente a curvatura do espaço-tempo e esta é gerada pelo tensor de momento-energia T . Como vimos no capítulo anterior, ele é um 2-tensor simétrico e pela lei de conservação da energia satisfaz

$$\operatorname{div} T = 0.$$

Na equação de Einstein, a fonte da curvatura T do espaço-tempo é colocada no lado direito da equação, enquanto que no lado esquerdo teremos um objeto que caracteriza a curvatura do espaço-tempo, logo este

deve ser também um 2-tensor simétrico com divergência nula. Chamamos este tensor de **tensor de Einstein** e o denotamos por G . Desta forma, a equação para a geração da gravidade (isto é, para o efeito da massa-energia sobre a curvatura do espaço-tempo) é

$$G = \kappa T,$$

onde κ é uma constante de proporcionalidade a ser determinada mais tarde (o fato que irá determiná-la é simplesmente a exigência que no limite para campos gravitacionais estáticos, relativamente fracos, ela deverá corresponder à lei da gravitação universal de Newton). Portanto, o 2-tensor simétrico G deve ser construído apenas a partir do tensor da curvatura de Riemann e da métrica, deve satisfazer $\text{div } G = 0$ e deve ser em princípio uma função linear da curvatura (se vamos medir a curvatura, não há motivo para medi-la de uma forma complicada, não linear). Estas exigências naturais determinam o tensor G :

$$G = \text{Ric} - \frac{1}{2} Sg. \quad (16.3)$$

Em termos de componentes, a equação de Einstein se escreve na forma covariante

$$R_{ij} - \frac{1}{2} Sg_{ij} = \kappa T_{ij}, \quad (16.4)$$

e na forma contravariante

$$R^{ij} - \frac{1}{2} Sg^{ij} = \kappa T^{ij}. \quad (16.5)$$

A equação de Einstein é portanto um sistema de 10 equações independentes, já que os tensores envolvidos são simétricos. Cada uma destas equações é uma equação parcial diferencial não linear de segunda ordem, isto é, envolvendo g_{ij} , $\partial_k g_{ij}$ e $\partial_{kl}^2 g_{ij}$, as componentes da métrica e suas derivadas parciais de primeira e segunda ordem. Por outro lado, g_{ij} são as componentes da métrica em algum sistema de coordenadas particular, e como há 4 coordenadas, temos 4 graus de liberdade, o que em princípio deveria tornar impossível obter uma solução única a partir de dados iniciais e de fronteira, como em problemas bem determinados típicos de EDPs. No entanto, a condição $\nabla G = 0$, que corresponde a 4 equações, faz com que o sistema efetivamente possua apenas 6 equações independentes para as 6 componentes g_{ij} que permitem caracterizar a geometria independentemente das coordenadas.

Podemos obter a curvatura escalar exclusivamente em termos do tensor momento-energia, o que ajuda muito em certos cálculos. De fato, fazendo a contração em relação à métrica da equação de Einstein, como

$$\begin{aligned} \text{tr}_g \text{Ric} &= S, \\ \text{tr}_g g &= 4, \end{aligned}$$

obtemos

$$S = -\kappa \text{tr}_g T. \quad (16.6)$$

Em termos de componentes,

$$S = -\kappa g^{ij} T_{ij} = -\kappa T_i^i = -\kappa g_{ij} T^{ij}.$$

Assim, a equação de Einstein pode reescrita na forma

$$\text{Ric} = \kappa \left[T - \frac{1}{2} (\text{tr}_g T) g \right]. \quad (16.7)$$

Em particular, a equação de Einstein no vácuo ($T = 0$) é

$$\text{Ric} = 0.$$

Isso não significa que a curvatura do espaço-tempo no vácuo é nula, pois o tensor de Ricci não contém toda a informação da curvatura: podemos ter $\text{Ric} = 0$ mas $R \neq 0$ (isso acontece com a métrica de Schwarzschild, como veremos mais tarde).

16.2.1 O Tensor de Einstein tem Divergência Nula

16.2 Proposição. Vale

$$\operatorname{div} G = 0.$$

Prova: Temos

$$\operatorname{div} G = \operatorname{div} \operatorname{Ric} - \frac{1}{2} \operatorname{div} (Sg)$$

Pela Proposição 14.16,

$$\operatorname{div} \operatorname{Ric} = \frac{\nabla S}{2}.$$

Pela Definição 14.15,

$$[\operatorname{div} (Sg)]_i = g^{jk} \nabla_k (Sg_{ij}) = g^{jk} (\nabla_k S) g_{ij} + Sg^{jk} \nabla_k g_{ij}.$$

Como (Proposição 14.4 ou 14.11)

$$\nabla g = 0,$$

segue que

$$[\operatorname{div} (Sg)]_i = g_{ij} g^{jk} (\nabla_k S) = \delta_i^k \nabla_k S = \nabla_i S,$$

onde

$$\operatorname{div} (Sg) = \nabla S.$$

Daí,

$$\operatorname{div} G = \frac{\nabla S}{2} - \frac{\nabla S}{2} = 0.$$

■

16.2.2 O Limite Newtoniano

Consideraremos agora um campo gravitacional estático, fraco.

O movimento de uma partícula livre no espaço-tempo segue uma geodésica. Parametrizando pelo tempo próprio τ , a equação da geodésica é dada por

$$\frac{d^2 x^k}{d\tau^2} + \Gamma_{ij}^k \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} = 0,$$

para $k = 0, 1, 2, 3$, $x^0 = ct$.

Primeiro vamos usar o princípio de equivalência para obter uma expressão para a aceleração da gravidade. Considere uma partícula livre no seu referencial comovente temporário em que ela está em repouso, de modo que

$$\begin{aligned} \frac{dx^j}{d\tau} &= 0, \quad i = 1, 2, 3, \\ \frac{dx^0}{d\tau} &= c \frac{dt}{d\tau}. \end{aligned}$$

Para $k = 1, 2, 3$, temos então que

$$\frac{d^2 x^k}{d\tau^2} = -c^2 \Gamma_{00}^k \left(\frac{dt}{d\tau} \right)^2.$$

Escolhendo $t = \tau$, segue que

$$\frac{d^2 x^k}{d\tau^2} = -c^2 \Gamma_{00}^k,$$

isto é, a aceleração da partícula é

$$\mathbf{a} = -c^2 (\Gamma_{00}^1, \Gamma_{00}^2, \Gamma_{00}^3). \quad (16.8)$$

Esta é, pelo princípio de equivalência, a **aceleração da gravidade g**. Suas componentes são os símbolos de Christoffel $\Gamma_{00}^1, \Gamma_{00}^2, \Gamma_{00}^3$. Vamos calcular estes símbolos em um campo gravitacional fraco.

A hipótese de um campo gravitacional fraco permite escrever a métrica do espaço-tempo na forma

$$g_{ij} = \eta_{ij} + h_{ij},$$

onde η_{ij} é a métrica de Minkowski e h_{ij} é uma pequena perturbação, isto é,

$$|h_{ij}| \ll 1.$$

Assumindo que o tensor métrica g e h são diagonais (o que é razoável, já que o tensor momento-energia é, conforme veremos mais abaixo), temos

$$g_{ij} = 0, \quad \text{se } i \neq j,$$

onde

$$\begin{aligned} g^{ii} &= \frac{1}{g_{ii}}, \\ g^{ij} &= 0, \quad \text{se } i \neq j, \end{aligned}$$

e o símbolo de Christoffel Γ_{00}^k , $k = 1, 2, 3$, é então dado por

$$\begin{aligned} \Gamma_{00}^k &= \frac{1}{2} (\partial_0 g_{0k} + \partial_0 g_{0k} - \partial_k g_{00}) g^{mk} \\ &= -\frac{1}{2} \frac{\partial_k g_{00}}{g_{kk}} \\ &= -\frac{1}{2} \frac{\partial_k (\eta_{00} + h_{00})}{\eta_{kk} + h_{kk}} \\ &= -\frac{1}{2} \frac{\partial_k h_{00}}{1 + h_{kk}}. \end{aligned}$$

Como $|h_{kk}| \ll 1$, podemos escrever

$$\Gamma_{00}^k = -\frac{1}{2} \partial_k h_{00}.$$

Portanto,

$$\mathbf{g} = \frac{c^2}{2} (\partial_1 h_{00}, \partial_2 h_{00}, \partial_3 h_{00}).$$

Mas

$$\mathbf{g} = -\nabla \phi,$$

logo

$$-\partial_k \phi = \frac{c^2}{2} \partial_k h_{00},$$

onde

$$\phi = -\frac{c^2}{2} h_{00}. \tag{16.9}$$

Isso mostra que apenas a componente h_{00} do tensor h contribui para a aceleração da gravidade no limite Newtoniano. É razoável assumir

$$|h_{ij}| \ll |h_{0j}| \ll |h_{00}|$$

isto é, todas as componentes $h_{ij} = 0$ exceto h_{00} , porque o tensor momento-energia satisfaz esta condição, como veremos mais abaixo (veja [Shutz], p. 194, para maiores detalhes), obtendo a métrica g na forma

$$ds^2 = -(1 - h_{00}) c^2 dt^2 + dx^2 + dy^2 + dz^2,$$

ou seja

$$ds^2 = - \left(1 + \frac{2}{c^2} \phi \right) c^2 dt^2 + dx^2 + dy^2 + dz^2. \quad (16.10)$$

Queremos apenas determinar κ . Usando as equações de Einstein na forma (16.7), escolhemos trabalhar com a equação

$$R_{00} = \kappa \left(T_{00} - \frac{1}{2} g_{00} \operatorname{tr}_g T \right).$$

Vamos calcular primeiro R_{00} . Como

$$R_{ij} = R_{kij}^k,$$

segue que

$$R_{00} = R_{k00}^k,$$

já que $R_{000}^0 = 0$, pelas simetrias do tensor endomorfismo curvatura. Temos

$$R_{ijk}^l = (\Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l) + \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j}.$$

Mas, como $|h_{00}| \ll 1$, a métrica é aproximadamente de Minkowski e podemos assumir que os símbolos de Christoffel se anulam, pois

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk}$$

e as componentes da métrica g_{ij} são aproximadamente constantes, logo suas derivadas se anulam. Portanto,

$$R_{ijk}^l = \frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j}.$$

Daí,

$$R_{k00}^k = \frac{\partial \Gamma_{00}^k}{\partial x^k} - \frac{\partial \Gamma_{k0}^k}{\partial x^0}.$$

Como estamos supondo um campo estático, que não depende da variável temporal x^0 , qualquer derivada em relação a esta variável é nula, logo

$$R_{k00}^k = \frac{\partial \Gamma_{00}^k}{\partial x^k}$$

e

$$\begin{aligned} \Gamma_{00}^k &= \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk} \\ &= \frac{1}{2} (\partial_0 g_{0m} + \partial_0 g_{0m} - \partial_m g_{00}) g^{mk} \\ &= -\frac{1}{2} (\partial_m g_{00}) g^{mk} \\ &= -\frac{1}{2} (\partial_k g_{00}) g^{kk}. \end{aligned}$$

Assim, para $k = 1, 2, 3$,

$$R_{k00}^k = \frac{1}{2} (\partial_{kk}^2 g_{00}) g^{kk} = \frac{1}{c^2} (\partial_{kk}^2 \phi).$$

Consequentemente,

$$R_{00} = \frac{1}{c^2} \Delta \phi. \quad (16.11)$$

O tensor momento-energia para fluidos perfeitos velocidades baixas ($v/c \approx 0$) é dominado pelo componente $T^{00} = \rho c^2$ e as outras componentes podem ser consideradas nulas. Portanto, usando $g_{00} \approx \eta_{00} = -1$, obtemos

$$\text{tr}_g T = \eta_{00} T^{00} = -\rho c^2$$

e

$$\begin{aligned} T_{00} - \frac{1}{2} g_{00} \text{tr}_g T &= T_{00} - \frac{1}{2} \eta_{00} (\eta_{00} T^{00}) \\ &= \frac{1}{2} T_{00} \\ &= \frac{1}{2} \rho c^2. \end{aligned}$$

Concluímos que no limite Newtoniano a equação

$$R_{00} = \kappa \left(T_{00} + \frac{1}{2} g_{00} \text{tr}_g T \right)$$

é

$$\Delta\phi = \frac{1}{2} \kappa c^2 \rho. \quad (16.12)$$

Comparando com a equação da gravitação universal de Newton (16.2)

$$\Delta\phi = 4\pi G\rho,$$

concluímos que

$$\kappa = \frac{8\pi G}{c^2}, \quad (16.13)$$

às vezes chamada a *constante gravitacional de Einstein*. Usando unidades geometrizadas $G = c = 1$, a equação de Einstein se escreve na forma

$$\text{Ric} - \frac{1}{2} Sg = 8\pi T. \quad (16.14)$$

16.3 Postulado Geodésico como Consequência da Equação de Einstein

O princípio que partículas livres percorrem trajetória geodésicas no espaço-tempo é às vezes conhecido como o *postulado geodésico*. Na verdade ele é uma consequência lógica da equação de Einstein, como veremos.

A equação de Einstein na forma contravariante quando $T = \rho(\mathbf{u} \otimes \mathbf{u})$ é o tensor poeira é

$$G = \kappa \rho u^i u^j.$$

Como $\text{div } G = 0$, segue que

$$\text{div} [\rho(\mathbf{u} \otimes \mathbf{u})] = 0.$$

Em coordenadas,

$$(\rho u^i u^j)_{;i} = 0.$$

Pela regra do produto,

$$(\rho u^i)_{;i} u^j + \rho u^i u^j_{;i} = 0$$

ou

$$(\rho u^i)_{;i} u^j + \rho u^i u^j_{;i} = 0. \quad (16.15)$$

Derivando

$$\|\mathbf{u}\|^2 = g_{ij} u^i u^j = u_i u^i = -1, \quad (16.16)$$

obtemos

$$(u_j u^j)_{;i} = 0,$$

onde, novamente pela regra do produto,

$$u_{j;i} u^j + u_j u^j_{;i} = 0.$$

Descendo e subindo índices no primeiro termo desta equação

$$\begin{aligned} u_{j;i} u^j &= u^j u_{j;i} = g^{kj} u_k u_{j;i} \\ &= u_k g^{kj} u_{j;i} = u_k u^k_{;i} \\ &= u_j u^j_{;i}, \end{aligned}$$

vemos que ela é igual ao segundo termo, donde

$$u_{j;i} u^j = 0. \quad (16.17)$$

Voltando a (16.15), multiplicamos ela por u_j , obtendo

$$u_j u^j (\rho u^i)_{;i} + \rho u^i u_j u^j_{;i} = 0$$

que se reduz por (16.16) e (16.17) a

$$-(\rho u^i)_{;i} = 0.$$

Este é o primeiro termo em (16.15), logo concluímos que

$$\rho u^i u^j_{;i} = 0.$$

Como $\rho \neq 0$, segue que

$$u^i u^j_{;i} = 0. \quad (16.18)$$

Lembrando que

$$u^j_{;i} = \frac{\partial u^j}{\partial x^i} + u^k \Gamma^j_{ki}$$

e

$$u^j = \frac{\partial x^j}{\partial \tau},$$

temos que

$$u^j_{;i} = \frac{\partial^2 x^j}{\partial x^i \partial \tau} + \frac{\partial x^k}{\partial \tau} \Gamma^j_{ik}.$$

Notando que

$$\frac{\partial^2 x^j}{\partial x^i \partial \tau} = \begin{cases} 0 & \text{se } i \neq 0, \\ \frac{\partial^2 x^j}{\partial \tau^2} & \text{se } i = 0, \end{cases}$$

vemos que (16.18) é a equação geodésica:

$$\frac{\partial^2 x^j}{\partial \tau^2} + \frac{\partial x^i}{\partial \tau} \frac{\partial x^k}{\partial \tau} \Gamma^j_{ik} = 0.$$

16.4 A Constante Cosmológica

Quando o espaço-tempo é plano, isto é, $R = 0$, segue que $\text{Ric} = 0$ e $S = 0$. Portanto, em um espaço-tempo plano, o tensor de Einstein se anula:

$$G = 0.$$

Assim, G é uma medida da métrica e também da curvatura. Além disso, no vácuo onde $T = 0$, pelas equações (16.2) e (16.7) temos

$$G = 0,$$

assim como

$$S = -8\pi \text{tr}_g T = 0$$

e

$$\text{Ric} = 8\pi \left[T - \frac{1}{2} (\text{tr}_g T) g \right] = 0.$$

Isso não implica necessariamente que $R = 0$, já que o tensor de Ricci não contém toda a informação do tensor curvatura de Riemann, mas é um resultado esperado.

Por outro lado, a equação de Einstein aplicada ao universo todo prediz um universo dinâmico, não estático (veja [MTW], Cap. 27). Quando Einstein percebeu isso em 1917, ele modificou sua equação, introduzindo uma **constante cosmológica** Λ para que o universo permanecesse estático:

$$G + \Lambda g = 8\pi T.$$

Observe que o tensor $G + \Lambda g$ continua sendo livre de divergência. Esta equação modificada admite uma solução estática para o universo (veja [MTW], Cap. 27, Box 27.5). Por outro lado, no vácuo onde $T = 0$, a equação modificada de Einstein se torna

$$\text{Ric} - \frac{1}{2} S g + \Lambda g = 0,$$

ou seja,

$$G = -\Lambda g$$

e não mais $G = 0$, de modo que podemos dizer que a constante cosmológica mede a densidade de energia do vácuo (o que é consistente com a noção de que o vácuo não é realmente vazio de acordo com a teoria quântica). Tomando o traço desta equação com respeito à métrica, como $\text{tr}_g \text{Ric} = S$ e $\text{tr}_g g = 4$, segue que

$$S - 2S + 4\Lambda = 0$$

ou seja,

$$S = 4\Lambda.$$

Segue que a equação de Einstein para o vácuo é

$$\text{Ric} = \Lambda g.$$

Portanto, a curvatura do vácuo não é a mais a curvatura nula do espaço-tempo de Minkowski.

Imediatamente depois que os resultados de Hubble sobre a expansão do universo (desvio para o vermelho da luz das galáxias, indicando o afastamento mútuo delas) se tornaram conhecidos em 1930, Einstein considerou como o seu maior erro a introdução da constante cosmológica na sua equação da relatividade geral (se não a tivesse removido, ele poderia reinvindicar a expansão do universo como uma das previsões da teoria) e novamente a removeu, o que equivale efetivamente a considerar $\Lambda = 0$. Quando no final da década de 90 foi observada a aceleração da expansão do universo, ela foi reintroduzida, com valor positivo a ser determinado através de experimentos. Mas as observações recentes indicam que o seu valor é bem próximo de zero (da ordem de 10^{-122} em unidades naturais adimensionais), de modo que no vácuo o espaço-tempo é realmente bem próximo ao espaço-tempo de Minkowski.

Capítulo 17

A Solução de Schwarzschild

A primeira solução para a equação de Einstein

$$\text{Ric} - \frac{1}{2} Sg = 8\pi T$$

foi obtida em dezembro de 1915 por Schwarzschild, para a surpresa de Einstein, que achava que dificilmente seriam obtidas soluções exatas para a sua equação.

O espaço-tempo de Schwarzschild é o modelo relativístico mais simples de um universo contendo uma única estrela. A estrela é assumida ser estática, isto é, sem rotação (ou com velocidade de rotação fisicamente desprezível), e esfericamente simétrica. O modelo resultante pode ser aplicado a regiões em torno de qualquer objeto astronômico isolado que satisfaz estas condições aproximadamente, tais como planetas, estrelas e buracos negros. No caso do sistema solar, ela provê um modelo mais preciso que o Newtoniano.

17.1 Métrica Plana em Coordenadas Esféricas

O espaço tridimensional menos a origem $\mathbb{R}^3 \setminus \{0\}$ é difeomorfo ao produto cartesiano $\mathbb{R}_+ \times \mathbb{S}^2$, onde $\mathbb{R}_+ = (0, +\infty)$. Isso permite definir coordenadas esféricas para $\mathbb{R}^3 \setminus \{0\}$. Desta forma, podemos definir um difeomorfismo de $\mathbb{R} \times \mathbb{R}_+ \times \mathbb{S}^2$ para o espaço-tempo de Minkowski \mathbb{M} ou um atlas para \mathbb{M} de sistemas de coordenadas esféricas por

$$\varphi(t, r, \theta, \phi) = (t, r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta).$$

Por exemplo, uma carta deste atlas seria $(\varphi, \mathbb{R} \times (0, +\infty) \times (0, \pi) \times (0, 2\pi))$; variando o domínio de θ e ϕ , cobrimos todo \mathbb{M} com exceção da origem. Note que aqui $\theta = \text{constante}$ denota paralelos e $\phi = \text{constante}$ denota meridianos (o exemplo de carta dado deixa de cobrir um semiplano fechado de \mathbb{R}^3). Neste tipo de carta temos

$$\begin{aligned}\partial_0 \varphi &= \partial_t \varphi = (1, 0, 0, 0), \\ \partial_1 \varphi &= \partial_r \varphi = (0, \sin \theta \cos \phi, \sin \theta \sin \phi, \cos \phi), \\ \partial_2 \varphi &= \partial_\theta \varphi = (0, r \cos \theta \cos \phi, r \cos \theta \sin \phi, -r \sin \theta), \\ \partial_3 \varphi &= \partial_\phi \varphi = (0, -r \sin \theta \sin \phi, r \sin \theta \cos \phi, 0),\end{aligned}$$

e os coeficientes g_{ij} da métrica de Lorentz são dados por

$$\begin{aligned} g_{00} &= -\langle \partial_0 \varphi, \partial_0 \varphi \rangle = -1, \\ g_{01} = g_{10} &= \langle \partial_0 \varphi, \partial_1 \varphi \rangle = 0, \\ g_{02} = g_{20} &= \langle \partial_0 \varphi, \partial_2 \varphi \rangle = 0, \\ g_{03} = g_{30} &= \langle \partial_0 \varphi, \partial_3 \varphi \rangle = 0, \\ g_{11} &= \langle \partial_1 \varphi, \partial_1 \varphi \rangle = 1, \\ g_{12} = g_{21} &= \langle \partial_1 \varphi, \partial_2 \varphi \rangle = 0, \\ g_{13} = g_{31} &= \langle \partial_1 \varphi, \partial_3 \varphi \rangle = 0, \\ g_{22} &= \langle \partial_2 \varphi, \partial_2 \varphi \rangle = r^2, \\ g_{23} = g_{32} &= \langle \partial_2 \varphi, \partial_3 \varphi \rangle = 0, \\ g_{33} &= \langle \partial_3 \varphi, \partial_3 \varphi \rangle = r^2 \operatorname{sen}^2 \theta. \end{aligned}$$

Portanto, em coordenadas esféricas, a métrica de Lorentz para o espaço-tempo de Minkowski se escreve na forma

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \operatorname{sen}^2 \theta d\phi^2.$$

Denotando o elemento de área da esfera por

$$d\sigma^2 = d\theta^2 + \operatorname{sen}^2 \theta d\phi^2,$$

podemos também escrever de forma mais sucinta

$$ds^2 = -dt^2 + dr^2 + r^2 d\sigma^2. \quad (17.1)$$

17.2 Métrica de Schwarzschild

No caso de uma estrela esféricamente simétrica, o espaço-tempo resultante deve ser esféricamente simétrico. Como a estrela também é suposta estática, a métrica não deve depender da variável temporal t . Assim, a métrica de Schwarzschild deve ser do tipo

$$ds^2 = -A(r) dt^2 + B(r) dr^2 + C(r) d\sigma^2.$$

Fazendo uma mudança de variável em \mathbb{R}_+ , podemos considerar $C(r) = r^2$, logo podemos escrever a métrica de Schwarzschild na forma

$$ds^2 = -T(r) dt^2 + D(r) dr^2 + r^2 d\sigma^2.$$

Nesta normalização, em cada espaço de repouso $t = \text{constante}$, a superfície $r = \text{constante}$ é a esfera padrão \mathbb{S}_r^2 de raio r , com curvatura gaussiana $1/r^2$ e área $4\pi r^2$.

A única fonte de gravitação no universo de Schwarzschild é a própria estrela, que não modelamos. Logo o espaço-tempo é um vácuo e a equação de Einstein implica que

$$\operatorname{Ric} = 0,$$

isto é, a métrica de Schwarzschild tem curvatura de Ricci nula.

Finalmente, longe da fonte de gravitação, a influência deve ser arbitrariamente pequena, de modo que a métrica de Schwarzschild deve aproximar a métrica plana do espaço de Minkowski (17.1) e portanto

$$\begin{aligned} \lim_{r \rightarrow +\infty} T(r) &\rightarrow -1, \\ \lim_{r \rightarrow +\infty} D(r) &\rightarrow 1. \end{aligned}$$

Veremos que estas condições determinam as funções $T(r)$ e $D(r)$.

17.1 Proposição. Uma métrica de Lorentz

$$ds^2 = -T(r)dt^2 + D(r)dr^2 + r^2d\sigma^2$$

em $\mathbb{R} \times \mathbb{R}_+ \times \mathbb{S}^2$ satisfaz $\text{Ric} = 0$ e

$$\begin{aligned}\lim_{r \rightarrow +\infty} T(r) &\rightarrow -1, \\ \lim_{r \rightarrow +\infty} D(r) &\rightarrow 1.\end{aligned}$$

se e somente se

$$\begin{aligned}T(r) &= -f(r), \\ D(r) &= \frac{1}{f(r)},\end{aligned}$$

com

$$f(r) = 1 - \frac{2M}{r}.$$

Prova: Veja [ONeill], Lemma 13.1, p. 366. ■

17.2 Definição. Considere a variedade diferenciável $\mathcal{M} = \mathbb{R} \times (2M, +\infty) \times \mathbb{S}^2$. \mathcal{M} torna-se uma variedade de Lorentz quando definimos nela a **métrica de Schwarzschild**

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{1}{1 - \frac{2M}{r}}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2. \quad (17.2)$$

□

Para as coordenadas de Schwarzschild (t, r, θ, ϕ) temos

$$\begin{aligned}g_{00} &= -\left(1 - \frac{2M}{r}\right) = -\frac{r-2M}{r}, \\ g_{11} &= \frac{1}{1 - \frac{2M}{r}} = \frac{r}{r-2M}, \\ g_{22} &= r^2, \\ g_{33} &= r^2\sin^2\theta, \\ g_{ij} &= 0 \quad \text{se } i \neq j.\end{aligned} \quad (17.3)$$

A métrica de Schwarzschild é ortogonal e, portanto, é fácil trabalhar com ela. Em particular,

$$\begin{aligned}g^{00} &= -\frac{r}{r-2M}, \\ g^{11} &= \frac{r-2M}{r}, \\ g^{22} &= \frac{1}{r^2}, \\ g^{33} &= \frac{1}{r^2\sin^2\theta}, \\ g^{ij} &= 0 \quad \text{se } i \neq j.\end{aligned} \quad (17.4)$$

Verificaremos agora que a métrica de Schwarzschild tem curvatura de Ricci nula, mas o seu tensor curvatura não é nulo, logo \mathcal{M} não é localmente isométrica à métrica plana.

17.3 Lema. Os 40 símbolos de Christoffel para a métrica de Schwarzschild são

$$\begin{aligned}\Gamma_{ij}^0 &= \begin{bmatrix} 0 & \frac{M}{r(r-2M)} & 0 & 0 \\ \frac{M}{r(r-2M)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \Gamma_{ij}^1 &= \begin{bmatrix} \frac{M(r-2M)}{r^3} & 0 & 0 & 0 \\ 0 & -\frac{M}{r(r-2M)} & 0 & 0 \\ 0 & 0 & -(r-2M) & 0 \\ 0 & 0 & 0 & -(r-2M)\sin^2\theta \end{bmatrix}, \\ \Gamma_{ij}^2 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin\theta\cos\theta \end{bmatrix}, \quad \Gamma_{ij}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot\theta \\ 0 & \frac{1}{r} & \cot\theta & 0 \end{bmatrix}.\end{aligned}\tag{17.5}$$

Prova: Os símbolos de Christoffel são dados por

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk}.$$

Como $g^{mk} = 0$ se $m \neq k$, segue que

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) g^{kk}.\tag{17.6}$$

Temos

$$\begin{aligned}\partial_1 g_{00} &= \frac{2M}{r^2}, \\ \partial_1 g_{11} &= -\frac{2M}{(r-2M)^2}, \\ \partial_1 g_{22} &= 2r, \\ \partial_1 g_{33} &= 2r \sin^2\theta, \\ \partial_2 g_{33} &= 2r^2 \sin\theta \cos\theta,\end{aligned}\tag{17.7}$$

e, caso contrário,

$$\partial_k g_{ij} = 0.\tag{17.8}$$

Logo, de (17.3), (17.4), (17.6) e (17.7) segue que, para $k = 0$,

$$\Gamma_{ij}^0 = \frac{1}{2} (\partial_i g_{j0} + \partial_j g_{i0} - \partial_0 g_{ij}) g^{00} = \frac{1}{2} (\partial_i g_{j0} + \partial_j g_{i0}) g^{00},$$

$$\begin{aligned}
\Gamma_{00}^0 &= (\partial_0 g_{00}) g^{00} = 0, \\
\Gamma_{01}^0 &= \Gamma_{10}^0 = \frac{1}{2} (\partial_0 g_{10} + \partial_1 g_{00}) g^{00} = \frac{1}{2} (\partial_1 g_{00}) g^{00} \\
&= \frac{M}{r(r-2M)}, \\
\Gamma_{02}^0 &= \Gamma_{20}^0 = \frac{1}{2} (\partial_0 g_{20} + \partial_2 g_{00}) g^{00} = 0, \\
\Gamma_{03}^0 &= \Gamma_{30}^0 = \frac{1}{2} (\partial_0 g_{30} + \partial_3 g_{00}) g^{00} = 0, \\
\Gamma_{11}^0 &= (\partial_1 g_{10}) g^{00} = 0, \\
\Gamma_{12}^0 &= \Gamma_{21}^0 = \frac{1}{2} (\partial_1 g_{20} + \partial_2 g_{10}) g^{00} = 0, \\
\Gamma_{13}^0 &= \Gamma_{31}^0 = \frac{1}{2} (\partial_1 g_{30} + \partial_3 g_{10}) g^{00} = 0, \\
\Gamma_{22}^0 &= (\partial_2 g_{20}) g^{00} = 0, \\
\Gamma_{23}^0 &= \Gamma_{32}^0 = \frac{1}{2} (\partial_2 g_{30} + \partial_3 g_{20}) g^{00} = 0, \\
\Gamma_{33}^0 &= (\partial_3 g_{30}) g^{00} = 0;
\end{aligned}$$

para $k = 1$,

$$\Gamma_{ij}^1 = \frac{1}{2} (\partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij}) g^{11},$$

$$\begin{aligned}
\Gamma_{00}^1 &= \frac{1}{2} (2\partial_0 g_{01} - \partial_1 g_{00}) g^{11} = -\frac{1}{2} (\partial_1 g_{00}) g^{11} \\
&= \frac{M}{r^2} - \frac{2M^2}{r^3} = \frac{M(r-2M)}{r^3}, \\
\Gamma_{01}^1 &= \Gamma_{10}^1 = \frac{1}{2} (\partial_0 g_{11} + \partial_1 g_{01} - \partial_1 g_{01}) g^{11} = 0, \\
\Gamma_{02}^1 &= \Gamma_{20}^1 = \frac{1}{2} (\partial_0 g_{21} + \partial_2 g_{01} - \partial_1 g_{02}) g^{11} = 0, \\
\Gamma_{03}^1 &= \Gamma_{30}^1 = \frac{1}{2} (\partial_0 g_{31} + \partial_3 g_{01} - \partial_1 g_{03}) g^{11} = 0, \\
\Gamma_{11}^1 &= \frac{1}{2} (\partial_1 g_{11}) g^{11} = -\frac{M}{r(r-2M)}, \\
\Gamma_{12}^1 &= \Gamma_{21}^1 = \frac{1}{2} (\partial_1 g_{21} + \partial_2 g_{11} - \partial_2 g_{12}) g^{11} = 0, \\
\Gamma_{13}^1 &= \Gamma_{31}^1 = \frac{1}{2} (\partial_1 g_{31} + \partial_3 g_{11} - \partial_1 g_{13}) g^{11} = 0, \\
\Gamma_{22}^1 &= \frac{1}{2} (2\partial_2 g_{21} - \partial_1 g_{22}) g^{11} = -\frac{1}{2} (\partial_1 g_{22}) g^{11} = 2M - r, \\
\Gamma_{23}^1 &= \Gamma_{32}^1 = \frac{1}{2} (\partial_2 g_{31} + \partial_3 g_{21} - \partial_1 g_{23}) g^{11} = 0, \\
\Gamma_{33}^1 &= \frac{1}{2} (2\partial_3 g_{31} - \partial_1 g_{33}) g^{11} = -\frac{1}{2} (\partial_1 g_{33}) g^{11} = (2M - r) \operatorname{sen}^2 \theta;
\end{aligned}$$

para $k = 2$,

$$\Gamma_{ij}^2 = \frac{1}{2} (\partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij}) g^{22},$$

$$\begin{aligned}
\Gamma_{00}^2 &= \frac{1}{2} (\partial_0 g_{02} + \partial_0 g_{02} - \partial_2 g_{00}) g^{22} = 0, \\
\Gamma_{01}^2 &= \Gamma_{10}^2 = \frac{1}{2} (\partial_0 g_{12} + \partial_1 g_{02} - \partial_2 g_{01}) g^{22} = 0, \\
\Gamma_{02}^2 &= \Gamma_{20}^2 = \frac{1}{2} (\partial_0 g_{22} + \partial_2 g_{02} - \partial_2 g_{02}) g^{22} = 0, \\
\Gamma_{03}^2 &= \Gamma_{30}^2 = \frac{1}{2} (\partial_0 g_{32} + \partial_3 g_{02} - \partial_2 g_{03}) g^{22} = 0, \\
\Gamma_{11}^2 &= \frac{1}{2} (\partial_1 g_{12} + \partial_1 g_{12} - \partial_2 g_{11}) g^{22} = 0, \\
\Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{1}{2} (\partial_1 g_{22} + \partial_2 g_{12} - \partial_2 g_{12}) g^{22} = \frac{1}{2} (\partial_1 g_{22}) g^{22} = \frac{1}{r}, \\
\Gamma_{13}^2 &= \Gamma_{31}^2 = \frac{1}{2} (\partial_1 g_{32} + \partial_3 g_{12} - \partial_2 g_{13}) g^{22} = 0, \\
\Gamma_{22}^2 &= \frac{1}{2} (\partial_2 g_{22}) g^{22} = 0, \\
\Gamma_{23}^2 &= \Gamma_{32}^2 = \frac{1}{2} (\partial_2 g_{32} + \partial_3 g_{22} - \partial_2 g_{23}) g^{22} = 0, \\
\Gamma_{33}^2 &= \frac{1}{2} (\partial_3 g_{32} + \partial_3 g_{32} - \partial_2 g_{33}) g^{22} = -\frac{1}{2} (\partial_2 g_{33}) g^{22} = -\operatorname{sen} \theta \cos \theta;
\end{aligned}$$

e, finalmente, para $k = 3$,

$$\Gamma_{ij}^3 = \frac{1}{2} (\partial_i g_{j3} + \partial_j g_{i3} - \partial_3 g_{ij}) g^{33} = \frac{1}{2} (\partial_i g_{j3} + \partial_j g_{i3}) g^{33},$$

$$\begin{aligned}
\Gamma_{00}^3 &= \frac{1}{2} (\partial_0 g_{03} + \partial_0 g_{03}) g^{33} = 0, \\
\Gamma_{01}^3 &= \Gamma_{10}^3 = \frac{1}{2} (\partial_0 g_{13} + \partial_1 g_{03}) g^{33} = 0, \\
\Gamma_{02}^3 &= \Gamma_{20}^3 = \frac{1}{2} (\partial_0 g_{23} + \partial_2 g_{03}) g^{33} = 0, \\
\Gamma_{03}^3 &= \Gamma_{30}^3 = \frac{1}{2} (\partial_0 g_{33} + \partial_3 g_{03}) g^{33} = 0, \\
\Gamma_{11}^3 &= (\partial_1 g_{13}) g^{33} = 0, \\
\Gamma_{12}^3 &= \Gamma_{21}^3 = \frac{1}{2} (\partial_1 g_{23} + \partial_2 g_{13}) g^{33} = 0, \\
\Gamma_{13}^3 &= \Gamma_{31}^3 = \frac{1}{2} (\partial_1 g_{33} + \partial_3 g_{13}) g^{33} = \frac{1}{2} (\partial_1 g_{33}) g^{33} = \frac{1}{r}, \\
\Gamma_{22}^3 &= (\partial_2 g_{23}) g^{33} = 0, \\
\Gamma_{23}^3 &= \Gamma_{32}^3 = \frac{1}{2} (\partial_2 g_{33} + \partial_3 g_{23}) g^{33} = \frac{1}{2} (\partial_2 g_{33}) g^{33} = \frac{\cos \theta}{\operatorname{sen} \theta}, \\
\Gamma_{33}^3 &= (\partial_3 g_{33}) g^{33} = 0.
\end{aligned}$$

■

17.4 Proposição. A métrica de Schwarzschild satisfaz

$$\operatorname{Ric} = S = 0. \quad (17.9)$$

Ela não é uma métrica plana, pois o seu tensor curvatura não é nulo, mas é uma métrica assintoticamente plana, isto é,

$$R_{ijk}^l \rightarrow 0 \quad \text{quando } r \rightarrow \infty.$$

Prova: As componentes do tensor endomorfismo curvatura são dadas por

$$R_{ijk}^l = \Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l + \partial_i \Gamma_{jk}^l - \partial_j \Gamma_{ik}^l$$

A curvatura de Ricci é dada por

$$R_{ij} = R_{kij}^k,$$

portanto vamos calcular apenas

$$R_{kij}^k = \Gamma_{ij}^m \Gamma_{km}^k - \Gamma_{kj}^m \Gamma_{im}^k + \partial_k \Gamma_{ij}^k - \partial_i \Gamma_{kj}^k.$$

Lembramos que $R_{kkj}^k = 0$. Usando (17.5), para $ij = 00$ temos

$$\begin{aligned} R_{k00}^k &= \Gamma_{00}^m \Gamma_{km}^k - \Gamma_{k0}^m \Gamma_{0m}^k + \partial_k \Gamma_{00}^k - \partial_0 \Gamma_{k0}^k \\ &= \Gamma_{00}^1 \Gamma_{k1}^k - \Gamma_{k0}^0 \Gamma_{00}^k - \Gamma_{k0}^1 \Gamma_{01}^k + \partial_k \Gamma_{00}^k, \end{aligned}$$

onde

$$\begin{aligned} R_{000}^0 &= 0, \\ R_{100}^1 &= \Gamma_{00}^1 \Gamma_{11}^1 - \Gamma_{10}^0 \Gamma_{00}^1 - \Gamma_{10}^1 \Gamma_{01}^1 + \partial_1 \Gamma_{00}^1 = \Gamma_{00}^1 \Gamma_{11}^1 - \Gamma_{10}^0 \Gamma_{00}^1 + \partial_1 \Gamma_{00}^1 \\ &= -\frac{M(r-2M)}{r^3} \frac{M}{r(r-2M)} - \frac{M}{r(r-2M)} \frac{M(r-2M)}{r^3} + \left(-\frac{2M}{r^3} + \frac{6M^2}{r^4} \right) \\ &= -\frac{2M^2}{r^4} - \frac{2M}{r^3} + \frac{6M^2}{r^4} = \frac{4M^2}{r^4} - \frac{2M}{r^3}, \\ R_{200}^2 &= \Gamma_{00}^1 \Gamma_{21}^2 - \Gamma_{20}^0 \Gamma_{00}^2 - \Gamma_{20}^1 \Gamma_{01}^2 + \partial_2 \Gamma_{00}^2 = \Gamma_{00}^1 \Gamma_{21}^2 \\ &= \frac{M(r-2M)}{r^3} \frac{1}{r} = \frac{M}{r^3} - \frac{2M^2}{r^4}, \\ R_{300}^3 &= \Gamma_{00}^1 \Gamma_{31}^3 - \Gamma_{30}^0 \Gamma_{00}^3 - \Gamma_{30}^1 \Gamma_{01}^3 + \partial_3 \Gamma_{00}^3 = \Gamma_{00}^1 \Gamma_{31}^3 \\ &= \frac{M(r-2M)}{r^3} \frac{1}{r} = \frac{M}{r^3} - \frac{2M^2}{r^4}, \end{aligned}$$

logo

$$R_{00} = 0; \quad (17.10)$$

para $ij = 01$ temos

$$\begin{aligned} R_{k01}^k &= \Gamma_{01}^m \Gamma_{km}^k - \Gamma_{k1}^m \Gamma_{0m}^k + \partial_k \Gamma_{01}^k - \partial_0 \Gamma_{k1}^k \\ &= \Gamma_{01}^1 \Gamma_{k1}^k - \Gamma_{k1}^0 \Gamma_{00}^k - \Gamma_{k1}^1 \Gamma_{01}^k + \partial_k \Gamma_{01}^k, \end{aligned}$$

onde

$$\begin{aligned} R_{001}^0 &= 0, \\ R_{101}^1 &= \Gamma_{01}^1 \Gamma_{11}^1 - \Gamma_{10}^0 \Gamma_{00}^1 - \Gamma_{11}^1 \Gamma_{01}^1 + \partial_1 \Gamma_{01}^1 = 0, \\ R_{201}^2 &= \Gamma_{01}^1 \Gamma_{21}^2 - \Gamma_{20}^0 \Gamma_{00}^2 - \Gamma_{21}^1 \Gamma_{01}^2 + \partial_2 \Gamma_{01}^2 = 0, \\ R_{301}^3 &= \Gamma_{01}^1 \Gamma_{31}^3 - \Gamma_{30}^0 \Gamma_{00}^3 - \Gamma_{31}^1 \Gamma_{01}^3 + \partial_3 \Gamma_{01}^3 = 0, \end{aligned}$$

logo

$$R_{01} = R_{10} = 0; \quad (17.11)$$

para $ij = 02$ temos

$$\begin{aligned} R_{k02}^k &= \Gamma_{02}^m \Gamma_{km}^k - \Gamma_{k2}^m \Gamma_{0m}^k + \partial_k \Gamma_{02}^k - \partial_0 \Gamma_{k2}^k \\ &= -\Gamma_{k2}^1 \Gamma_{01}^k, \end{aligned}$$

donde

$$\begin{aligned} R_{002}^0 &= 0, \\ R_{102}^1 &= -\Gamma_{12}^1 \Gamma_{01}^1 = 0, \\ R_{202}^2 &= -\Gamma_{22}^1 \Gamma_{01}^2 = 0, \\ R_{302}^3 &= -\Gamma_{32}^1 \Gamma_{01}^3 = 0, \end{aligned}$$

logo

$$R_{02} = R_{20} = 0; \quad (17.12)$$

para $ij = 03$ temos

$$\begin{aligned} R_{k03}^k &= \Gamma_{03}^m \Gamma_{km}^k - \Gamma_{k3}^m \Gamma_{0m}^k + \partial_k \Gamma_{03}^k - \partial_0 \Gamma_{k3}^k \\ &= -\Gamma_{k3}^1 \Gamma_{01}^k, \end{aligned}$$

donde

$$\begin{aligned} R_{003}^0 &= 0, \\ R_{103}^1 &= -\Gamma_{13}^1 \Gamma_{01}^1 = 0, \\ R_{203}^2 &= -\Gamma_{23}^1 \Gamma_{01}^2 = 0, \\ R_{303}^3 &= -\Gamma_{33}^1 \Gamma_{01}^3 = 0, \end{aligned}$$

logo

$$R_{03} = R_{30} = 0; \quad (17.13)$$

para $ij = 11$ temos

$$\begin{aligned} R_{k11}^k &= \Gamma_{11}^m \Gamma_{km}^k - \Gamma_{k1}^m \Gamma_{1m}^k + \partial_k \Gamma_{11}^k - \partial_1 \Gamma_{k1}^k \\ &= \Gamma_{11}^1 \Gamma_{k1}^k - \Gamma_{k1}^0 \Gamma_{10}^k - \Gamma_{k1}^1 \Gamma_{11}^k - \Gamma_{k1}^2 \Gamma_{12}^k - \Gamma_{k1}^3 \Gamma_{13}^k + \partial_k \Gamma_{11}^k - \partial_1 \Gamma_{k1}^k, \end{aligned}$$

donde

$$\begin{aligned} R_{011}^0 &= \Gamma_{11}^1 \Gamma_{01}^0 - \Gamma_{01}^0 \Gamma_{10}^0 - \Gamma_{01}^1 \Gamma_{11}^0 - \Gamma_{11}^2 \Gamma_{12}^0 - \Gamma_{11}^3 \Gamma_{13}^0 + \partial_0 \Gamma_{11}^0 - \partial_1 \Gamma_{01}^0 \\ &= \Gamma_{11}^1 \Gamma_{01}^0 - \Gamma_{01}^0 \Gamma_{10}^0 - \partial_1 \Gamma_{01}^0 \\ &= -\frac{M^2}{r^2 (r-2M)^2} - \frac{M^2}{r^2 (r-2M)^2} + \frac{2M(r-M)}{r^2 (r-2M)^2} \\ &= \frac{2M}{r^2 (r-2M)}, \end{aligned}$$

$$R_{111}^1 = 0,$$

$$\begin{aligned} R_{211}^2 &= \Gamma_{11}^1 \Gamma_{21}^2 - \Gamma_{21}^0 \Gamma_{10}^2 - \Gamma_{21}^1 \Gamma_{11}^2 - \Gamma_{21}^2 \Gamma_{12}^2 - \Gamma_{21}^3 \Gamma_{13}^2 + \partial_2 \Gamma_{11}^2 - \partial_1 \Gamma_{21}^2 \\ &= \Gamma_{11}^1 \Gamma_{21}^2 - \Gamma_{21}^2 \Gamma_{12}^2 - \partial_1 \Gamma_{21}^2 \\ &= -\frac{M}{r^2 (r-2M)} - \frac{1}{r^2} + \frac{1}{r^2} \\ &= -\frac{M}{r^2 (r-2M)}, \end{aligned}$$

$$\begin{aligned} R_{311}^3 &= \Gamma_{11}^1 \Gamma_{31}^3 - \Gamma_{31}^0 \Gamma_{10}^3 - \Gamma_{31}^1 \Gamma_{11}^3 - \Gamma_{31}^2 \Gamma_{12}^3 - \Gamma_{31}^3 \Gamma_{13}^3 + \partial_3 \Gamma_{11}^3 - \partial_1 \Gamma_{31}^3 \\ &= \Gamma_{11}^1 \Gamma_{31}^3 - \Gamma_{31}^3 \Gamma_{13}^3 - \partial_1 \Gamma_{31}^3 \\ &= -\frac{M}{r^2 (r-2M)} - \frac{1}{r^2} + \frac{1}{r^2} \\ &= -\frac{M}{r^2 (r-2M)}, \end{aligned}$$

logo

$$R_{11} = 0; \quad (17.14)$$

para $ij = 12$ temos

$$\begin{aligned} R_{k12}^k &= \Gamma_{12}^m \Gamma_{km}^k - \Gamma_{k2}^m \Gamma_{1m}^k + \partial_k \Gamma_{12}^k - \partial_1 \Gamma_{k2}^k \\ &= \Gamma_{12}^2 \Gamma_{k2}^k - \Gamma_{k2}^1 \Gamma_{11}^k - \Gamma_{k2}^2 \Gamma_{12}^k - \Gamma_{k2}^3 \Gamma_{13}^k + \partial_k \Gamma_{12}^k - \partial_1 \Gamma_{k2}^k, \end{aligned}$$

donde

$$\begin{aligned} R_{012}^0 &= \Gamma_{12}^2 \Gamma_{02}^0 - \Gamma_{02}^1 \Gamma_{11}^0 - \Gamma_{02}^2 \Gamma_{12}^0 - \Gamma_{02}^3 \Gamma_{13}^0 + \partial_0 \Gamma_{12}^0 - \partial_1 \Gamma_{02}^0 = 0, \\ R_{112}^1 &= 0, \\ R_{212}^2 &= \Gamma_{12}^2 \Gamma_{22}^2 - \Gamma_{22}^1 \Gamma_{11}^2 - \Gamma_{22}^2 \Gamma_{12}^2 - \Gamma_{22}^3 \Gamma_{13}^2 + \partial_2 \Gamma_{12}^2 - \partial_1 \Gamma_{22}^2 = \partial_2 \Gamma_{12}^2 \\ &= -\frac{1}{r^2}, \\ R_{312}^3 &= \Gamma_{12}^2 \Gamma_{32}^3 - \Gamma_{32}^1 \Gamma_{11}^3 - \Gamma_{32}^2 \Gamma_{12}^3 - \Gamma_{32}^3 \Gamma_{13}^3 + \partial_3 \Gamma_{12}^3 - \partial_1 \Gamma_{32}^3 \\ &= \Gamma_{12}^2 \Gamma_{32}^3 - \Gamma_{32}^3 \Gamma_{13}^3 - \partial_1 \Gamma_{32}^3 \\ &= \frac{1}{r} \cot \theta - \frac{1}{r} \cot \theta + \frac{1}{r^2} \\ &= \frac{1}{r^2}, \end{aligned}$$

logo

$$R_{12} = R_{21} = 0; \quad (17.15)$$

para $ij = 13$ temos

$$\begin{aligned} R_{k13}^k &= \Gamma_{13}^m \Gamma_{km}^k - \Gamma_{k3}^m \Gamma_{1m}^k + \partial_k \Gamma_{13}^k - \partial_1 \Gamma_{k3}^k \\ &= \Gamma_{13}^3 \Gamma_{k3}^k - \Gamma_{k3}^1 \Gamma_{11}^k - \Gamma_{k3}^2 \Gamma_{12}^k - \Gamma_{k3}^3 \Gamma_{13}^k + \partial_k \Gamma_{13}^k - \partial_1 \Gamma_{k3}^k, \end{aligned}$$

donde

$$\begin{aligned} R_{013}^0 &= \Gamma_{13}^3 \Gamma_{03}^0 - \Gamma_{03}^1 \Gamma_{11}^0 - \Gamma_{03}^2 \Gamma_{12}^0 - \Gamma_{03}^3 \Gamma_{13}^0 + \partial_0 \Gamma_{13}^0 - \partial_1 \Gamma_{03}^0 = 0, \\ R_{113}^1 &= 0, \\ R_{213}^2 &= \Gamma_{13}^3 \Gamma_{23}^2 - \Gamma_{23}^1 \Gamma_{11}^2 - \Gamma_{23}^2 \Gamma_{12}^2 - \Gamma_{23}^3 \Gamma_{13}^2 + \partial_2 \Gamma_{13}^2 - \partial_1 \Gamma_{23}^2 = 0, \\ R_{313}^3 &= \Gamma_{13}^3 \Gamma_{33}^3 - \Gamma_{33}^1 \Gamma_{11}^3 - \Gamma_{33}^2 \Gamma_{12}^3 - \Gamma_{33}^3 \Gamma_{13}^3 + \partial_3 \Gamma_{13}^3 - \partial_1 \Gamma_{33}^3 = 0, \end{aligned}$$

logo

$$R_{13} = R_{31} = 0; \quad (17.16)$$

para $ij = 22$ temos

$$\begin{aligned} R_{k22}^k &= \Gamma_{22}^m \Gamma_{km}^k - \Gamma_{k2}^m \Gamma_{2m}^k + \partial_k \Gamma_{22}^k - \partial_2 \Gamma_{k2}^k \\ &= \Gamma_{22}^1 \Gamma_{k1}^k - \Gamma_{k2}^1 \Gamma_{21}^k - \Gamma_{k2}^2 \Gamma_{22}^k - \Gamma_{k2}^3 \Gamma_{23}^k + \partial_k \Gamma_{22}^k - \partial_2 \Gamma_{k2}^k, \end{aligned}$$

donde

$$\begin{aligned}
R_{022}^0 &= \Gamma_{22}^1 \Gamma_{01}^0 - \Gamma_{02}^1 \Gamma_{21}^0 - \Gamma_{02}^2 \Gamma_{22}^0 - \Gamma_{02}^3 \Gamma_{23}^0 + \partial_0 \Gamma_{22}^0 - \partial_2 \Gamma_{02}^0 = \Gamma_{22}^1 \Gamma_{01}^0 \\
&= -\frac{M}{r}, \\
R_{122}^1 &= \Gamma_{22}^1 \Gamma_{11}^1 - \Gamma_{12}^1 \Gamma_{21}^1 - \Gamma_{12}^2 \Gamma_{22}^1 - \Gamma_{12}^3 \Gamma_{23}^1 + \partial_1 \Gamma_{22}^1 - \partial_2 \Gamma_{12}^1 \\
&= \Gamma_{22}^1 \Gamma_{11}^1 - \Gamma_{12}^2 \Gamma_{22}^1 + \partial_1 \Gamma_{22}^1 \\
&= \frac{M}{r} + \frac{r-2M}{r} - 1 \\
&= -\frac{M}{r}, \\
R_{222}^2 &= 0, \\
R_{322}^3 &= \Gamma_{22}^1 \Gamma_{31}^3 - \Gamma_{32}^1 \Gamma_{21}^3 - \Gamma_{32}^2 \Gamma_{22}^3 - \Gamma_{32}^3 \Gamma_{23}^3 + \partial_3 \Gamma_{22}^3 - \partial_2 \Gamma_{32}^3 \\
&= \Gamma_{22}^1 \Gamma_{31}^3 - \Gamma_{32}^3 \Gamma_{23}^3 - \partial_2 \Gamma_{32}^3 \\
&= -\frac{r-2M}{r} - \cot^2 \theta + \frac{1}{\sin^2 \theta} = -\frac{r-2M}{r} - \frac{\cos^2 \theta}{\sin^2 \theta} + \frac{1}{\sin^2 \theta} \\
&= -1 + \frac{2M}{r} + \frac{1-\cos^2 \theta}{\sin^2 \theta} = -1 + \frac{2M}{r} + \frac{\sin^2 \theta}{\sin^2 \theta} \\
&= \frac{2M}{r},
\end{aligned}$$

logo

$$R_{22} = 0; \quad (17.17)$$

para $ij = 23$ temos

$$\begin{aligned}
R_{k23}^k &= \Gamma_{23}^m \Gamma_{km}^k - \Gamma_{k3}^m \Gamma_{2m}^k + \partial_k \Gamma_{23}^k - \partial_2 \Gamma_{k3}^k \\
&= \Gamma_{23}^3 \Gamma_{k3}^k - \Gamma_{k3}^1 \Gamma_{21}^k - \Gamma_{k3}^2 \Gamma_{22}^k - \Gamma_{k3}^3 \Gamma_{23}^k + \partial_k \Gamma_{23}^k - \partial_2 \Gamma_{k3}^k,
\end{aligned}$$

donde

$$\begin{aligned}
R_{023}^0 &= \Gamma_{23}^3 \Gamma_{03}^0 - \Gamma_{03}^1 \Gamma_{21}^0 - \Gamma_{03}^2 \Gamma_{22}^0 - \Gamma_{03}^3 \Gamma_{23}^0 + \partial_0 \Gamma_{23}^0 - \partial_2 \Gamma_{03}^0 = 0, \\
R_{123}^1 &= \Gamma_{23}^3 \Gamma_{13}^1 - \Gamma_{13}^1 \Gamma_{21}^1 - \Gamma_{13}^2 \Gamma_{22}^1 - \Gamma_{13}^3 \Gamma_{23}^1 + \partial_1 \Gamma_{23}^1 - \partial_2 \Gamma_{13}^1 = 0, \\
R_{223}^2 &= 0, \\
R_{323}^3 &= \Gamma_{23}^3 \Gamma_{23}^2 - \Gamma_{23}^1 \Gamma_{21}^2 - \Gamma_{23}^2 \Gamma_{22}^2 - \Gamma_{23}^3 \Gamma_{23}^2 + \partial_2 \Gamma_{23}^2 - \partial_2 \Gamma_{23}^2 = 0,
\end{aligned}$$

logo

$$R_{23} = R_{32} = 0; \quad (17.18)$$

e, finalmente, para $ij = 33$ temos

$$\begin{aligned}
R_{k33}^k &= \Gamma_{33}^m \Gamma_{km}^k - \Gamma_{k3}^m \Gamma_{3m}^k + \partial_k \Gamma_{33}^k - \partial_3 \Gamma_{k3}^k \\
&= \Gamma_{33}^1 \Gamma_{k1}^k + \Gamma_{33}^2 \Gamma_{k2}^k - \Gamma_{k3}^1 \Gamma_{31}^k - \Gamma_{k3}^2 \Gamma_{32}^k - \Gamma_{k3}^3 \Gamma_{33}^k + \partial_k \Gamma_{33}^k
\end{aligned}$$

onde

$$\begin{aligned}
 R_{033}^0 &= \Gamma_{33}^1 \Gamma_{01}^0 + \Gamma_{33}^2 \Gamma_{02}^0 - \Gamma_{03}^1 \Gamma_{31}^0 - \Gamma_{03}^2 \Gamma_{32}^0 - \Gamma_{03}^3 \Gamma_{33}^0 + \partial_0 \Gamma_{33}^0 = \Gamma_{33}^1 \Gamma_{01}^0 \\
 &= -\frac{M}{r} \sin^2 \theta, \\
 R_{133}^1 &= \Gamma_{33}^1 \Gamma_{11}^1 + \Gamma_{33}^2 \Gamma_{12}^1 - \Gamma_{13}^1 \Gamma_{31}^1 - \Gamma_{13}^2 \Gamma_{32}^1 - \Gamma_{13}^3 \Gamma_{33}^1 + \partial_1 \Gamma_{33}^1 \\
 &= \Gamma_{33}^1 \Gamma_{11}^1 - \Gamma_{13}^3 \Gamma_{33}^1 + \partial_1 \Gamma_{33}^1 \\
 &= \frac{M}{r} \sin^2 \theta + \frac{r - 2M}{r} \sin^2 \theta - \sin^2 \theta \\
 &= -\frac{M}{r} \sin^2 \theta, \\
 R_{233}^2 &= \Gamma_{33}^1 \Gamma_{21}^2 + \Gamma_{33}^2 \Gamma_{22}^2 - \Gamma_{23}^1 \Gamma_{31}^2 - \Gamma_{23}^2 \Gamma_{32}^2 - \Gamma_{23}^3 \Gamma_{33}^2 + \partial_2 \Gamma_{33}^2 \\
 &= \Gamma_{33}^1 \Gamma_{21}^2 - \Gamma_{23}^3 \Gamma_{33}^2 + \partial_2 \Gamma_{33}^2 \\
 &= -\frac{r - 2M}{r} \sin^2 \theta + \frac{\cos \theta}{\sin \theta} \sin \theta \cos \theta - \cos^2 \theta + \sin^2 \theta \\
 &= \frac{2M}{r} \sin^2 \theta, \\
 R_{333}^3 &= 0,
 \end{aligned}$$

logo

$$R_{33} = 0. \quad (17.19)$$

Para mostrar que a métrica de Schwarzschild é uma métrica assintoticamente plana, precisamos calcular as demais componentes do tensor curvatura. Isso é deixado como exercício. ■

17.5 Proposição. *Se a massa esférica estática tem massa M em unidades geométricas e denotamos $R = 2M$, o espaço-tempo para $r > R$ e para $0 < r < R$ com a métrica de Schwarzschild satisfaz a equação de Einstein.*

Prova: Pois a métrica de Schwarzschild satisfaz $\text{Ric} = 0$, $S = 0$ e o tensor momento-energia satisfaz $T = 0$ para $r > R$. ■

Massa em unidades geométricas é definida por

$$M = \frac{G}{c^2} M_{\text{kg}},$$

17.6 Definição. O raio $R = 2M$ é chamado o **raio do horizonte de evento**.

O espaço-tempo com a métrica de Schwarzschild para $r > R$ é chamado **espaço-tempo exterior de Schwarzschild**.

O espaço-tempo com a métrica de Schwarzschild para $0 < r < R$ é chamado o **buraco negro de Schwarzschild**. □

Quando o chamado *raio de Schwarzschild* r é suficientemente grande, a métrica é aproximadamente de Minkowski e pode-se pensar t como sendo o tempo e r a distância radial. Para r pequeno, no entanto, estas interpretações passam a não valer. Quando r é menor que $R = 2M$, a função

$$f(r) = 1 - \frac{2M}{r}$$

passa a ser negativa e ∂_t torna-se um vetor do tipo espaço, enquanto que ∂_r passa a ser um vetor do tipo tempo. Neste caso, as analogias com a teoria de gravitação de Newton perdem o valor.

O raio do horizonte de evento é onde nada, nem a luz passa. Este é o valor predito também pela teoria gravitacional de Newton. De fato,

$$V(r) = -G \frac{M_{\text{kg}} m_{\text{kg}}}{r}$$

é a energia potencial gravitacional de uma partícula com massa m_{kg} (escolhendo energia potencial zero no infinito), de modo que sua energia mecânica total é dada por

$$E = K + V = \frac{1}{2}m_{\text{kg}}v_{\text{m/s}}^2 - G \frac{M_{\text{kg}}m_{\text{kg}}}{r}.$$

Portanto, sua *velocidade de escape* (isto é, a velocidade que ela deve ter para chegar pelo menos em repouso no infinito, ou seja, com energia mecânica total nula) satisfaz

$$\frac{1}{2}m_{\text{kg}}v_{\text{m/s}}^2 = G \frac{M_{\text{kg}}m_{\text{kg}}}{r},$$

isto é,

$$v_{\text{m/s}} = \sqrt{\frac{2GM_{\text{kg}}}{r}}.$$

Em unidades naturais do espaço-tempo temos

$$v = \frac{v_{\text{m/s}}}{c} = \sqrt{\frac{2GM_{\text{kg}}}{c^2 r}} = \sqrt{\frac{2M}{r}}.$$

Em particular, a velocidade de escape será a velocidade da luz, isto é, $v = 1$, quando

$$r = 2M.$$

Porém, como veremos mais tarde, enquanto que na teoria de Newton a partícula lançada do horizonte de evento diretamente para cima com velocidade menor que a da luz sobe alguma distância antes de voltar a cair (pois sobe desacelerando até que sua velocidade caia totalmente para zero para algum valor finito de r), pela teoria da relatividade geral nem a luz pode avançar qualquer distância por menor que seja a partir do horizonte de eventos (ou ultrapassá-lo, se lançada abaixo dele). Por outro lado, mesmo na teoria da relatividade geral, é possível que uma partícula atravesse o horizonte de eventos sem sentir nada, nem mesmo as forças de maré; isso aconteceria por exemplo em um buraco negro muito massivo, com um raio do horizonte de evento muito grande (como o hipotetizado buraco negro que reside no centro da Via Láctea). A singularidade em $r = 2M$ na métrica de Schwarzschild é uma singularidade matemática, proveniente da escolha de coordenadas. Como veremos depois, em outros sistemas de coordenadas a singularidade desaparece.

17.3 Conservação de Energia e Momento Angular em Geodésicas do Espaço-tempo de Schwarzschild

17.3.1 Equação Geodésica na Métrica de Schwarzschild

17.7 Lema. *Em um sistema de coordenadas ortogonais (x^0, \dots, x^n) , a equação geodésica para uma geodésica $\gamma(\tau) = (x^0(\tau), \dots, x^n(\tau))$ pode ser escrita na forma*

$$\frac{d}{d\tau} \left[g_{kk} \left(\frac{dx^k}{d\tau} \right) \right] = \frac{1}{2} \sum_{i=0}^n \partial_k g_{ii} \left(\frac{dx^i}{d\tau} \right)^2, \quad k = 0, 1, \dots, n. \quad (17.20)$$

Prova: Para $k = 1, \dots, n$ a equação geodésica é

$$\frac{d^2 x^k}{d\tau^2} + \sum_{i,j=0}^n \Gamma_{ij}^k \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} = 0.$$

Como os símbolos de Christoffel são dados por

$$\Gamma_{ij}^k = \frac{1}{2} (\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) g^{mk},$$

em um sistema de coordenadas ortogonal temos

$$\Gamma_{ij}^k = \frac{1}{2g_{kk}} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}), \quad (17.21)$$

pois g é uma matriz diagonal. Daí,

$$\Gamma_{ij}^k = \begin{cases} 0, & \text{se } i \neq j \text{ e } i, j \neq k, \\ \frac{1}{2g_{kk}} \partial_j g_{kk}, & \text{se } i \neq j \text{ e } i = k, \\ \frac{1}{2g_{kk}} \partial_i g_{kk}, & \text{se } i \neq j \text{ e } j = k, \\ \frac{1}{g_{kk}} \partial_k g_{kk} - \frac{1}{2g_{kk}} \partial_k g_{kk}, & \text{se } i = j = k, \\ -\frac{1}{2g_{kk}} \partial_k g_{ii}, & \text{se } i = j \neq k, \end{cases}$$

onde

$$\sum_{i,j=0}^n \Gamma_{ij}^k \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} = \frac{1}{g_{kk}} \sum_{i=0}^n \partial_i g_{kk} \frac{dx^i}{d\tau} \frac{dx^k}{d\tau} - \frac{1}{2g_{kk}} \sum_{i=0}^n \partial_k g_{ii} \left(\frac{dx^i}{d\tau} \right)^2.$$

Logo, a equação geodésica se torna

$$\frac{d^2 x^k}{d\tau^2} + \frac{1}{g_{kk}} \sum_{i=0}^n \partial_i g_{kk} \frac{dx^i}{d\tau} \frac{dx^k}{d\tau} - \frac{1}{2g_{kk}} \sum_{i=0}^n \partial_k g_{ii} \left(\frac{dx^i}{d\tau} \right)^2 = 0,$$

onde

$$g_{kk} \frac{d^2 x^k}{d\tau^2} + \sum_{i=0}^n \partial_i g_{kk} \frac{dx^i}{d\tau} \frac{dx^k}{d\tau} = \frac{1}{2} \sum_{i=0}^n \partial_k g_{ii} \left(\frac{dx^i}{d\tau} \right)^2.$$

Como, pela regra do produto e pela regra da cadeia,

$$\begin{aligned} \frac{d}{d\tau} \left[g_{kk} \left(\frac{dx^k}{d\tau} \right) \right] &= g_{kk} \frac{d^2 x^k}{d\tau^2} + \frac{dg_{kk}}{d\tau} \frac{dx^k}{d\tau} \\ &= g_{kk} \frac{d^2 x^k}{d\tau^2} + \sum_{i=0}^n \frac{\partial g_{kk}}{\partial x^i} \frac{dx^i}{d\tau} \frac{dx^k}{d\tau}, \end{aligned}$$

segue o resultado. ■

17.8 Proposição. A equação das geodésicas $\gamma(\tau) = (t(\tau), r(\tau), \theta(\tau), \phi(\tau))$ na métrica de Schwarzschild é

$$\frac{d}{d\tau} \left[\left(1 - \frac{2M}{r} \right) \frac{dt}{d\tau} \right] = 0,$$

$$\frac{d}{d\tau} \left[\frac{1}{1 - \frac{2M}{r}} \frac{dr}{d\tau} \right] = \frac{M}{r^2} \left(\frac{dt}{d\tau} \right)^2 - \frac{M}{(r - 2M)^2} \left(\frac{dr}{d\tau} \right)^2 + r \left(\frac{d\theta}{d\tau} \right)^2 + r \sin^2 \theta \left(\frac{d\phi}{d\tau} \right)^2,$$

$$\frac{d}{d\tau} \left[r^2 \frac{d\theta}{d\tau} \right] = r^2 \sin \theta \cos \theta \left(\frac{d\phi}{d\tau} \right)^2,$$

$$\frac{d}{d\tau} \left[r^2 \sin^2 \theta \left(\frac{d\phi}{d\tau} \right) \right] = 0.$$

Prova: Segue do lema anterior, substituindo

$$\begin{aligned}g_{00} &= -\left(1 - \frac{2M}{r}\right), \\g_{11} &= \frac{1}{1 - \frac{2M}{r}}, \\g_{22} &= r^2, \\g_{33} &= r^2 \sin^2 \theta,\end{aligned}$$

(note que os coeficientes da métrica não dependem de t e ϕ) e

$$\begin{aligned}\partial_1 g_{00} &= \frac{2M}{r^2}, \\ \partial_1 g_{11} &= -\frac{2M}{(r - 2M)^2}, \\ \partial_1 g_{22} &= 2r, \\ \partial_1 g_{33} &= 2r \sin^2 \theta, \\ \partial_2 g_{33} &= 2r^2 \sin \theta \cos \theta.\end{aligned}$$

■

17.3.2 Energia e Momento Angular

Da proposição anterior segue que existem constantes reais E, L tais que

$$\begin{aligned}\left(1 - \frac{2M}{r}\right) \frac{dt}{d\tau} &= E, \\r^2 \sin^2 \theta \left(\frac{d\phi}{d\tau}\right) &= L.\end{aligned}$$

E é a energia e L o momento angular.

17.4 Queda Livre

17.5 Cálculo da Deflexão da Luz

17.6 Precessão do Perihélio de Mercúrio

17.7 Buracos Negros de Schwarzschild

17.8 Produtos Dobrados

O espaço-tempo de Schwarzschild é um exemplo particular de uma variedade métrica do tipo *produto dobrado*.

17.8.1 Espaçotempos Estáticos

17.9 Definição. Sejam M^3 uma variedade riemanniana de dimensão 3 e $f \in C^\infty(M)$ uma função positiva. Se I é um intervalo aberto, definimos o **espaçotempo estático padrão** como sendo a variedade de Lorentz

$I \times M$ com a métrica

$$-f dt^2 + ds^2.$$

□

Capítulo 18

Campos de Jacobi

Neste capítulo estudaremos como a curvatura afeta as geodésicas. Em um ponto p , as geodésicas radiais partem deste ponto e se irradiam. Veremos que em uma região de curvatura seccional positiva as geodésicas se separam menos, em comparação com as geodésicas (retas) no espaço euclidiano (ou seja, os raios em $T_p M$), enquanto que em uma região de curvatura seccional negativa as geodésicas se separam mais.

18.1 A Equação de Jacobi

Dados $p \in M$ e $v \in T_p M$ tais que $\exp_p(v)$ está definido, considere a variação geodésica $F : (-\varepsilon, \varepsilon) \times [0, 1] \rightarrow M$ da geodésica radial

$$\gamma(t) = \exp_p(tv)$$

vista na demonstração do Lema de Gauss, ou seja

$$F(s, t) = \exp_p(tv(s))$$

onde $v(s)$ é uma curva de $T_p M$ partindo de $v(0) = v$ com velocidade $v'(0) =: w$. Para cada s fixado, $F_s(t) = F(s, t)$ é a geodésica radial que parte de p com velocidade $v(s)$. A velocidade de afastamento destas geodésicas radiais no ponto $\gamma(t)$ é exatamente o vetor

$$\frac{\partial F}{\partial s}(0, t) = d(\exp_p)_{tv} tw, \quad (18.1)$$

que é o vetor velocidade da curva transversal partindo de $\gamma(t)$. Neste capítulo estudaremos este campo em detalhes. Veremos em primeiro lugar nesta seção que ele satisfaz uma equação diferencial, chamada a *equação de Jacobi* (também conhecida como *equação do desvio geodésico* em Relatividade Geral). [Para conveniência do leitor, repetimos o Lema 13.34 que foi visto no apêndice do Capítulo 13, juntamente com a sua demonstração.]

18.1 Lema. *Seja $F : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ uma variação e V um campo vetorial ao longo de F . Então*

$$\frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} = R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) V.$$

Prova: Escolha um sistema de coordenadas para uma vizinhança de $p \in M$ e escreva

$$V = \sum_{i=1}^n V^i(s, t) \partial_i.$$

Então

$$\frac{DV}{ds} = \frac{D}{ds} \left(\sum_{i=1}^n V^i \partial_i \right) = \sum_{i=1}^n \frac{\partial V^i}{\partial s} \partial_i + \sum_{i=1}^n V^i \frac{D}{ds} \partial_i,$$

onde

$$\frac{D}{dt} \frac{DV}{ds} = \sum_{i=1}^n \frac{\partial^2 V^i}{\partial s \partial t} \partial_i + \sum_{i=1}^n \frac{\partial V^i}{\partial s} \frac{D}{dt} \partial_i + \sum_{i=1}^n \frac{\partial V^i}{\partial t} \frac{D}{ds} \partial_i + \sum_{i=1}^n V^i \frac{D}{dt} \frac{D}{ds} \partial_i.$$

Trocando t por s ,

$$\frac{D}{ds} \frac{DV}{dt} = \sum_{i=1}^n \frac{\partial^2 V^i}{\partial s \partial t} \partial_i + \sum_{i=1}^n \frac{\partial V^i}{\partial t} \frac{D}{ds} \partial_i + \sum_{i=1}^n \frac{\partial V^i}{\partial s} \frac{D}{dt} \partial_i + \sum_{i=1}^n V^i \frac{D}{ds} \frac{D}{dt} \partial_i.$$

Logo,

$$\frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} = \sum_{i=1}^n V^i \left(\frac{D}{dt} \frac{D}{ds} - \frac{D}{ds} \frac{D}{dt} \right) \partial_i. \quad (18.2)$$

Escrevendo

$$F(s, t) = (x^1(s, t), \dots, x^n(s, t)),$$

de modo que

$$\begin{aligned} \frac{\partial F}{\partial t} &= \sum_{j=1}^n \frac{\partial x^j}{\partial t} \partial_j, \\ \frac{\partial F}{\partial s} &= \sum_{k=1}^n \frac{\partial x^k}{\partial s} \partial_k, \end{aligned}$$

temos

$$\frac{D}{ds} \partial_i = \nabla_{\sum \frac{\partial x^k}{\partial s} \partial_k} \partial_i = \sum_{k=1}^n \frac{\partial x^k}{\partial s} \nabla_{\partial_k} \partial_i.$$

Daí,

$$\begin{aligned} \frac{D}{dt} \frac{D}{ds} \partial_i &= \frac{D}{dt} \left(\sum_{k=1}^n \frac{\partial x^k}{\partial s} \nabla_{\partial_k} \partial_i \right) \\ &= \sum_{k=1}^n \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \sum_{k=1}^n \frac{\partial x^k}{\partial s} \frac{D}{dt} (\nabla_{\partial_k} \partial_i) \\ &= \sum_{k=1}^n \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \sum_{k=1}^n \frac{\partial x^k}{\partial s} \nabla_{\sum \frac{\partial x^j}{\partial t} \partial_j} (\nabla_{\partial_k} \partial_i) \\ &= \sum_{k=1}^n \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \sum_{k,j=1}^n \frac{\partial x^k}{\partial s} \frac{\partial x^j}{\partial t} \nabla_{\partial_j} \nabla_{\partial_k} \partial_i. \end{aligned}$$

Da mesma forma, trocando t por s e depois j por k ,

$$\begin{aligned} \frac{D}{ds} \frac{D}{dt} \partial_i &= \sum_{k=1}^n \frac{\partial^2 x^k}{\partial t \partial s} \nabla_{\partial_k} \partial_i + \sum_{k,j=1}^n \frac{\partial x^k}{\partial t} \frac{\partial x^j}{\partial s} \nabla_{\partial_j} \nabla_{\partial_k} \partial_i \\ &= \sum_{k=1}^n \frac{\partial^2 x^k}{\partial s \partial t} \nabla_{\partial_k} \partial_i + \sum_{k,j=1}^n \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} \nabla_{\partial_k} \nabla_{\partial_j} \partial_i. \end{aligned}$$

Portanto,

$$\begin{aligned}
\frac{D}{dt} \frac{DV}{ds} - \frac{D}{ds} \frac{DV}{dt} &= \sum_{i=1}^n V^i \sum_{j,k=1}^n \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} (\nabla_{\partial_j} \nabla_{\partial_k} \partial_i - \nabla_{\partial_k} \nabla_{\partial_j} \partial_i) \\
&= \sum_{i=1}^n V^i \sum_{j,k=1}^n \frac{\partial x^j}{\partial t} \frac{\partial x^k}{\partial s} R(\partial_j, \partial_k) \partial_i \\
&= \sum_{i=1}^n V^i R \left(\sum_{j=1}^n \frac{\partial x^j}{\partial t} \partial_j, \sum_{k=1}^n \frac{\partial x^k}{\partial s} \partial_k \right) \partial_i \\
&= R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) \sum_{i=1}^n V^i \partial_i \\
&= R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) V.
\end{aligned}$$

■

18.2 Proposição (Equação de Jacobi). *Seja $F : (-\varepsilon, \varepsilon) \times [0, 1] \rightarrow M$ uma variação geodésica e denote*

$$J(t) = \frac{\partial F}{\partial s}(0, t).$$

Então J satisfaz a equação diferencial linear

$$\frac{D^2 J}{dt^2} + R(J, \gamma') \gamma' = 0. \quad (18.3)$$

Prova: Como as curvas principais $t \mapsto F(s, t)$ da variação são geodésicas, temos

$$\frac{D}{dt} \frac{\partial F}{\partial t} = 0.$$

Segue do lema anterior e do Lema de Simetria 12.25 que

$$\begin{aligned}
0 &= \frac{D}{ds} \left(\frac{D}{dt} \frac{\partial F}{\partial t} \right) = \frac{D}{dt} \left(\frac{D}{ds} \frac{\partial F}{\partial t} \right) - R \left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s} \right) \frac{\partial F}{\partial t} \\
&= \frac{D}{dt} \left(\frac{D}{dt} \frac{\partial F}{\partial s} \right) + R \left(\frac{\partial F}{\partial s}, \frac{\partial F}{\partial t} \right) \frac{\partial F}{\partial t} \\
&= \frac{D^2}{dt^2} \frac{\partial F}{\partial s} + R \left(\frac{\partial F}{\partial s}, \frac{\partial F}{\partial t} \right) \frac{\partial F}{\partial t}.
\end{aligned}$$

■

18.2 Campos de Jacobi

18.3 Definição. Seja $\gamma : I \rightarrow M$ uma geodésica. Qualquer campo vetorial ao longo de γ que satisfaz a equação de Jacobi é chamado um **campo de Jacobi** ao longo de γ . □

Segue da Proposição 18.2 que o campo variacional de uma variação geodésica é um campo de Jacobi.

Um campo de Jacobi é determinado por suas condições iniciais:

18.4 Proposição (Existência e Unicidade de Campos de Jacobi). *Seja $\gamma : I \rightarrow M$ uma geodésica. Dado $t_0 \in I$ e $V, W \in T_{\gamma(t_0)}M$, existe um único campo de Jacobi ao longo de γ tal que*

$$\begin{aligned} J(t_0) &= V, \\ \frac{DJ}{dt}(t_0) &= W. \end{aligned}$$

Em particular, os campos de Jacobi ao longo de uma geodésica formam um espaço vetorial de dimensão $2n$.

Prova: Seja $\{E_1(t), \dots, E_n(t)\}$ campos ortonormais paralelos ao longo de γ . Escrevendo

$$J = \sum_{i=1}^n J^i E_i,$$

temos

$$\frac{D^2 J}{dt^2}(t) = \sum_{i=1}^n \frac{d^2 J^i}{dt^2} E_i,$$

e

$$\begin{aligned} R(J, \gamma') \gamma' &= R \left(\sum_{k=1}^n J^k E_k, \sum_{j=1}^n \frac{d\gamma^j}{dt} E_j \right) \sum_{l=1}^n \frac{d\gamma^l}{dt} E_l \\ &= \sum_{j,k,l=1}^n \frac{d\gamma^j}{dt} \frac{d\gamma^l}{dt} J^k R(E_k, E_j) E_l \\ &= \sum_{j,k,l=1}^n R_{kjl}^i \frac{d\gamma^j}{dt} \frac{d\gamma^l}{dt} J^k E_i, \end{aligned}$$

de modo que a equação de Jacobi é equivalente ao sistema de equações diferenciais lineares de segunda ordem:

$$\frac{d^2 J^i}{dt^2} + \sum_{j,k,l=1}^n R_{kjl}^i \frac{d\gamma^j}{dt} \frac{d\gamma^l}{dt} J^k = 0$$

para $i = 1, \dots, n$. Como o sistema é linear, a existência e unicidade de soluções está garantida em todo o intervalo I , dadas condições iniciais. ■

Um campo vetorial ao longo de uma curva que não seja do tipo luz sempre pode ser decomposto nas suas componentes tangencial e normal:

$$V = V^\top \oplus V^\perp.$$

No caso de campos de Jacobi, ambas as componentes são também campos de Jacobi:

18.5 Proposição. *Seja $\gamma : I \rightarrow M$ uma geodésica que não é do tipo luz e J um campo vetorial ao longo de γ . Decomponha*

$$J = J^\top \oplus J^\perp$$

nas suas componentes tangencial e normal, respectivamente. Então J é um campo de Jacobi se e somente se J^\top e J^\perp são campos de Jacobi.

Prova: Pela linearidade da equação de Jacobi, se J^\top, J^\perp são campos de Jacobi então J é um campo de Jacobi.

Reciprocamente, seja J um campo de Jacobi. Como J^\top é um múltiplo escalar de γ' , temos

$$R(J^\top, \gamma') = 0,$$

onde

$$R(J, \gamma') = R(J^\perp, \gamma').$$

Além disso, como

$$\langle R(J, \gamma') \gamma', \gamma' \rangle = R(J, \gamma' \gamma', \gamma') = 0,$$

temos que $R(J, \gamma') \gamma'$ é normal a γ .

Porque γ é uma geodésica, temos

$$\begin{aligned} \left(\frac{DV}{dt} \right)^\top &= \frac{DV^\top}{dt}, \\ \left(\frac{DV}{dt} \right)^\perp &= \frac{DV^\perp}{dt}, \end{aligned}$$

para qualquer campo V . De fato, para qualquer campo W temos

$$\frac{D}{dt} \langle W, \gamma' \rangle = \left\langle \frac{DW}{dt}, \gamma' \right\rangle + \left\langle W, \frac{D\gamma'}{dt} \right\rangle = \left\langle \frac{DW}{dt}, \gamma' \right\rangle,$$

de modo que $\langle W, \gamma' \rangle = 0$ implica $\left\langle \frac{DW}{dt}, \gamma' \right\rangle = 0$, isto é, se W é normal à geodésica γ , então sua derivada covariante $\frac{DW}{dt}$ também é normal a γ ; da mesma forma, se W é um campo tangente a γ , então sua derivada covariante $\frac{DW}{dt}$ também é tangente a γ , pois se $W = f\gamma'$, então

$$\frac{DW}{dt} = \frac{df}{dt} \gamma' + f \frac{D\gamma'}{dt} = \frac{df}{dt} \gamma'.$$

Logo,

$$\begin{aligned} \left(\frac{D^2 J}{dt^2} \right)^\top &= \frac{D^2 J^\top}{dt^2}, \\ \left(\frac{D^2 J}{dt^2} \right)^\perp &= \frac{D^2 J^\perp}{dt^2}. \end{aligned}$$

Portanto, a equação de Jacobi

$$\frac{D^2 J}{dt^2} + R(J, \gamma') \gamma' = 0$$

dá origem às duas equações de Jacobi

$$\begin{aligned} \frac{D^2 J^\top}{dt^2} + R(J^\top, \gamma') \gamma' &= 0, \\ \frac{D^2 J^\perp}{dt^2} + R(J^\perp, \gamma') \gamma' &= 0. \end{aligned}$$

■

18.6 Exemplo. Existem sempre dois campos de Jacobi *tangenciais* triviais ao longo de uma geodésica γ :

$$J_0(t) = \gamma'(t) \tag{18.4}$$

que satisfaz as condições iniciais

$$J_0(0) = \gamma'(0) \quad \text{e} \quad \frac{DJ_0}{dt}(0) = 0$$

e

$$J_1(t) = t\gamma'(t) \quad (18.5)$$

que satisfaz as condições iniciais

$$J_1(0) = 0 \quad \text{e} \quad \frac{D J_1}{dt}(0) = \gamma'(0).$$

O primeiro é um campo de Jacobi porque

$$\begin{aligned} \frac{D^2 J_0}{dt^2} &= \frac{D}{dt} \frac{D\gamma'}{dt} = \frac{D}{dt} 0 = 0, \\ R(J_0, \gamma') \gamma' &= R(\gamma', \gamma') \gamma = 0. \end{aligned}$$

O segundo é um campo de Jacobi porque

$$\begin{aligned} \frac{D^2 J_1}{dt^2} &= \frac{D}{dt} \frac{D J_1}{dt} = \frac{D}{dt} \left(\gamma' + t \frac{D\gamma'}{dt} \right) = \frac{D\gamma'}{dt} = 0, \\ R(J_1, \gamma') \gamma' &= R(t\gamma', \gamma') \gamma' = tR(\gamma', \gamma') \gamma' = 0. \end{aligned}$$

De fato, J_0 é o campo variacional da variação $F(s, t) = \gamma(t+s)$, enquanto que J_1 é o campo variacional da variação $F(s, t) = \gamma(te^s)$. Como estas variações são apenas reparametrizações das geodésicas γ , elas não podem dizer mais nada do que a própria γ . \square

Devido ao visto no exemplo anterior, apenas campos de Jacobi normais ao longo de γ são interessantes, capazes de fornecer informação geométrica nova.

18.7 Proposição. *Seja $\gamma : I \rightarrow M$ uma geodésica, $t_0 \in I$ e J um campo de Jacobi ao longo de γ . Então*

$$\langle J(t), \gamma'(t) \rangle = At + B,$$

onde

$$\begin{aligned} A &= \left\langle \frac{D J}{dt}(t_0), \gamma'(t_0) \right\rangle, \\ B &= \langle J(t_0), \gamma'(t_0) \rangle - \left\langle \frac{D J}{dt}(t_0), \gamma'(t_0) \right\rangle t_0. \end{aligned}$$

Em particular, um campo de Jacobi ao longo de γ é normal a γ' se e somente se

$$J(t_0), \frac{D J}{dt}(t_0) \perp \gamma'(t_0).$$

Além disso, qualquer campo de Jacobi ortogonal a γ' em dois pontos é um campo normal.

18.8 Lema. *Em particular, o subespaço dos campos de Jacobi tangenciais tem dimensão 2, enquanto que o subespaço dos campos de Jacobi normais tem dimensão $2n - 2$.*

Prova: Pela compatibilidade da métrica,

$$\begin{aligned} \frac{d^2}{dt^2} \langle J, \gamma' \rangle &= \frac{d}{dt} \left(\left\langle \frac{D J}{dt}, \gamma' \right\rangle + \left\langle J, \frac{D\gamma'}{dt} \right\rangle \right) \\ &= \frac{d}{dt} \left\langle \frac{D J}{dt}, \gamma' \right\rangle = \left\langle \frac{D^2 J}{dt^2}, \gamma' \right\rangle + \left\langle \frac{D J}{dt}, \frac{D\gamma'}{dt} \right\rangle \\ &= \left\langle \frac{D^2 J}{dt^2}, \gamma' \right\rangle \\ &= -\langle R(J, \gamma') \gamma', \gamma' \rangle \\ &= -R(J, \gamma', \gamma', \gamma') \\ &= 0. \end{aligned}$$

Logo,

$$f(t) = \langle J(t), \gamma'(t) \rangle = At + B$$

para alguns $A, B \in \mathbb{R}$. Como

$$\begin{aligned} f(t_0) &= \langle J(t_0), \gamma'(t_0) \rangle, \\ f'(t_0) &= \left\langle \frac{D}{dt}(t_0), \gamma'(t_0) \right\rangle, \end{aligned}$$

segue a identidade enunciada. Consequentemente, $f \equiv 0$ se e somente se $J(t_0), \frac{D}{dt}(t_0) \perp \gamma'(t_0)$. Da mesma forma, se J se anula em dois pontos, então f é a função linear nula. ■

18.9 Proposição. *Dados $p \in M$, $v \in T_p M$ tal que $\exp_p(v)$ está definido e $w \in T_v T_p M = T_p M$, considere a variação geodésica*

$$F(s, t) = \exp_p(tv(s)),$$

onde $v(s)$ é uma curva em $T_p M$ com $v(0) = v$ e $v'(0) = w$. Então o campo de Jacobi

$$J(t) = \frac{\partial F}{\partial s}(0, t) = d(\exp_p)_{tv} tw$$

ao longo da geodésica radial $\gamma : [0, 1] \rightarrow M$ dada por $\gamma(t) = \exp_p(tv)$ satisfaz

$$\begin{aligned} J(0) &= 0, \\ \frac{D}{dt}(0) &= w. \end{aligned}$$

Prova: É óbvio que $J(0) = 0$. Para mostrar que

$$\frac{D}{dt}(0) = w,$$

calculamos

$$\begin{aligned} \frac{D}{dt}(t) &= \frac{D}{dt} \left[d(\exp_p)_{tv} tw \right] = \frac{D}{dt} \left[td(\exp_p)_{tv} w \right] \\ &= d(\exp_p)_{tv} w + t \frac{D}{dt} \left[d(\exp_p)_{tv} w \right]. \end{aligned}$$

Logo,

$$\frac{D}{dt}(0) = d(\exp_p)_0 w = w.$$

■

Em coordenadas normais, é fácil calcular campos de Jacobi:

18.10 Proposição. *Sejam $p \in M$, (x^1, \dots, x^n) coordenadas normais centradas em p e $\gamma : I \rightarrow M$ uma geodésica radial partindo de p . Então, para cada vetor $V \in T_p M$, o campo de Jacobi ao longo de γ tal que*

$$\begin{aligned} J(0) &= 0, \\ \frac{D}{dt}(0) &= V, \end{aligned}$$

é dado por

$$J(t) = \sum_{i=1}^n t V^i \left. \frac{\partial}{\partial x^i} \right|_{\gamma(t)}.$$

Prova: Se $J(t)$ é um campo ao longo de γ definido pela fórmula acima, temos

$$J(0) = 0$$

e (usando a fórmula da derivada covariante em coordenadas normais)

$$\begin{aligned} \frac{D J}{dt}(0) &= \sum_{k=1}^n \left[\frac{d J^k}{dt}(0) + \sum_{i,j=1}^n \frac{d \gamma^i}{dt}(0) \Gamma_{ij}^k(0) J^j(0) \right] \frac{\partial}{\partial x^k}(0) \\ &= \sum_{k=1}^n \frac{d J^k}{dt}(0) \frac{\partial}{\partial x^k}(0) \\ &= \sum_{k=1}^n V^k \frac{\partial}{\partial x^k}(0) \\ &= V. \end{aligned}$$

Pelo teorema de existência e unicidade, basta provar que J é um campo de Jacobi; mostraremos que J é o campo variacional de uma variação geodésica. De fato, se $W = \gamma'(0)$, como γ é uma geodésica radial, ela é dada por

$$\gamma(t) = \exp_p(tW);$$

em coordenadas, podemos escrever simplesmente $\gamma(t) = tW$. Defina a variação geodésica

$$F(s, t) = \exp_p(t(W + sV));$$

em coordenadas, podemos escrever simplesmente $F(s, t) = t(W + sV)$. Observe que $v(s) = W + sV$ satisfaz $v(0) = W$ e $v'(0) = V$. Como por definição

$$\frac{\partial}{\partial x^i} \Big|_{\gamma(t)} = d(\exp_p)_{tW} e_i,$$

segue que o campo variacional de F é

$$\frac{\partial F}{\partial s}(0, t) = d(\exp_p)_{tW}(tV) = \sum_{i=1}^n tV^i \frac{\partial}{\partial x^i} \Big|_{\gamma(t)} = J(t).$$

■

18.3 Campos de Jacobi em Variedades de Curvatura Seccional Constante

Para métricas com curvatura seccional constante, temos uma fórmula explícita para campos de Jacobi de um tipo diferente, expressando um campo de Jacobi como o múltiplo escalar de um campo vetorial paralelo.

18.11 Proposição. *Seja M uma variedade riemanniana com curvatura seccional constante K e $\gamma : I \rightarrow M$ uma geodésica unitária. Os campos de Jacobi normais ao longo de γ que se anulam em $t = 0$ são dados por*

$$J(t) = \rho(t) E(t),$$

onde E é um campo paralelo normal ao longo de γ e

$$\rho(t) = \begin{cases} t & \text{se } K = 0, \\ R \operatorname{sen} \frac{t}{R} & \text{se } K = \frac{1}{R^2} > 0, \\ R \operatorname{senh} \frac{t}{R} & \text{se } K = -\frac{1}{R^2} < 0. \end{cases}$$

Prova: Como M tem curvatura seccional constante K , seu endomorfismo curvatura é dado por

$$R(X, Y)Z = K[\langle Y, Z \rangle X - \langle X, Z \rangle Y],$$

conforme vimos no capítulo anterior. Logo, como $\|\gamma'\| = 1$, se J é um campo de Jacobi normal temos que

$$R(J, \gamma')\gamma' = K[\langle \gamma', \gamma' \rangle J - \langle J, \gamma' \rangle \gamma'] = KJ.$$

Portanto, a equação de Jacobi para um campo normal de Jacobi ao longo de uma geodésica unitária em uma variedade com curvatura seccional constante K é

$$\frac{D^2 J}{dt^2} + KJ = 0.$$

Escolhendo um campo vetorial paralelo normal $E(t)$ ao longo de γ e substituindo $J(t) = \rho(t)E(t)$ na equação de Jacobi podemos obter uma solução para esta equação para quaisquer condições iniciais; pelo teorema de existência e unicidade ela é a única solução para a equação. De fato, temos

$$\begin{aligned} \frac{D^2(\rho E)}{dt^2} &= \frac{D}{dt} \frac{D(\rho E)}{dt} = \frac{D}{dt} \left(\frac{d\rho}{dt} E + \rho \frac{DE}{dt} \right) = \frac{D}{dt} \left(\frac{d\rho}{dt} E \right) \\ &= \frac{d^2\rho}{dt^2} E + \frac{d\rho}{dt} \frac{DE}{dt} \\ &= \frac{d^2\rho}{dt^2} E, \end{aligned}$$

logo, substituindo $J(t) = \rho(t)E(t)$ na equação de Jacobi produz

$$\left(\frac{d^2\rho}{dt^2} + K\rho \right) E = 0$$

e a equação linear de segunda ordem

$$\frac{d^2\rho}{dt^2} + K\rho = 0$$

tem as soluções dadas no enunciado para as condições iniciais $\rho(0) = 0$ e $\rho'(0) = 1$. ■

Combinando as fórmulas destes dois lemas, obteremos nossa primeira aplicação geométrica dos campos de Jacobi: uma fórmula explícita para métricas de curvatura constante em coordenadas normais; desta fórmula seguirá que variedades de curvatura seccional constante são localmente isométricas.

18.12 Proposição. *Seja (M, g) uma variedade riemanniana com curvatura seccional constante K . Sejam $p \in M$, (x^1, \dots, x^n) coordenadas normais centradas em p e r a função distância radial. Denote por $\|\cdot\|_g$ a norma de um vetor tangente na métrica g e por $\|\cdot\|_e$ a norma euclidiana de um vetor tangente nestas coordenadas. Para cada q nesta vizinhança normal de p e para cada $V \in T_q M$ escreva*

$$V = V_r + V_S,$$

onde V_r está na direção de $\partial/\partial r$ e V_S é tangente à esfera $S_r(p)$ que passa por q . Então a métrica g pode ser escrita

$$\|V\|_g = \begin{cases} \|V_r\|_e^2 + \|V_S\|_e^2 & \text{se } K = 0, \\ \|V_r\|_e^2 + \frac{R^2}{r^2} \left(\operatorname{sen}^2 \frac{r}{R} \right) \|V_S\|_e^2 & \text{se } K = \frac{1}{R^2} > 0, \\ \|V_r\|_e^2 + \frac{R^2}{r^2} \left(\operatorname{senh}^2 \frac{r}{R} \right) \|V_S\|_e^2 & \text{se } K = -\frac{1}{R^2} < 0. \end{cases}$$

Prova: Pelo lema de Gauss, a decomposição $V = V_r + V_S$ é ortogonal, logo

$$\|V\|_g^2 = \|V_r\|_g^2 + \|V_S\|_g^2.$$

Como a distância radial é a distância euclideana (também consequência do lema de Gauss) temos que

$$\|V_r\|_g = \|V_r\|_e.$$

Falta apenas calcular $\|V_S\|_g$.

Considere a geodésica radial unitária de p a q . Pela Proposição 18.10, o campo

$$J(t) = \sum_{i=1}^n \frac{t}{r} V_S^i \frac{\partial}{\partial x^i} \Big|_{\gamma(t)}$$

é um campo de Jacobi $J(t)$ que se anula em p e tal que $J(r) = V_S$. Em particular, J é normal a γ' em p e q e portanto J é normal ao longo de γ pela Proposição 18.7. Logo, J possui a representação dada na Proposição 18.11. Segue que

$$\|V_S\|_g^2 = \|J(r)\|^2 = |\rho(r)|^2 \|E(r)\|^2 = |\rho(r)|^2 \|E(0)\|^2 = |\rho(r)|^2 \left\| \frac{D}{dt}(0) \right\|^2,$$

já que $\rho'(0) = 1$. Mas

$$\frac{D}{dt}(0) = \sum_{i=1}^n \frac{1}{r} V_S^i \frac{\partial}{\partial x^i} \Big|_p.$$

Como $g = e$ em p , segue o resultado. ■

18.13 Proposição (Unicidade Local de Métricas de Curvatura Constante). *Sejam (M, g) e $(\widetilde{M}, \widetilde{g})$ variedades riemannianas com curvatura seccional constante K . Então (M, g) e $(\widetilde{M}, \widetilde{g})$ são localmente isométricas.*

Prova: Sejam $p \in M$ e $\tilde{p} \in \widetilde{M}$ pontos quaisquer e $B_\varepsilon(p), B_\varepsilon(\tilde{p})$ bolas geodésicas, em particular vizinhanças normais parametrizadas por

$$\begin{aligned} \varphi &= \exp_p : B_\varepsilon(0) \longrightarrow B_\varepsilon(p), \\ \tilde{\varphi} &= \exp_{\tilde{p}} : B_\varepsilon(0) \longrightarrow B_\varepsilon(\tilde{p}). \end{aligned}$$

Como as métricas na bola $B_\varepsilon(0) \subset \mathbb{R}^n$, onde identificamos \mathbb{R}^n com ambos os espaços tangentes $T_p M$ e $T_{\tilde{p}} \widetilde{M}$ são idênticas pela Proposição 18.12, segue que $\tilde{\varphi} \circ \varphi^{-1}$ é uma isometria. ■

18.4 Velocidade de Afastamento das Geodésicas e Curvatura Secional

Se $r(t)$ é tal que

$$\lim_{t \rightarrow 0} \frac{r(t)}{t^k} = 0,$$

denotaremos este fato por

$$r(t) = o(t^k).$$

18.14 Lema. Sejam $p \in M$ e $\gamma : I \rightarrow M$ a geodésica radial partindo de p com velocidade inicial $\gamma'(0) = V$. Seja $W \in T_p M$ com $\|W\| = 1$ e considere o campo de Jacobi J tal que

$$\begin{aligned} J(0) &= 0, \\ \frac{DJ}{dt}(0) &= W. \end{aligned}$$

Então

$$\|J(t)\|^2 = t^2 - \frac{1}{3}R(V, W, W, V)t^4 + o(t^4).$$

Em particular, se γ é uma geodésica unitária,

$$\|J(t)\|^2 = t^2 - \frac{1}{3}K(V, W)t^4 + o(t^4)$$

e

$$\|J(t)\| = t - \frac{1}{6}K(V, W)t^3 + o(t^3).$$

Prova: Para simplificar a notaçāo, denotaremos

$$\begin{aligned} J'(t) &= \frac{DJ}{dt}(t), \\ J''(t) &= \frac{D^2J}{dt^2}(t) \end{aligned}$$

e, em geral,

$$J^{(k)}(t) = \frac{D^k J}{dt^k}(t).$$

Para provar o resultado, aplicaremos a fórmula de Taylor à função real $f : I \rightarrow \mathbb{R}$ definida por

$$f(t) = \|J(t)\|^2 = \langle J(t), J(t) \rangle.$$

Temos

$$\begin{aligned} f' &= 2\langle J, J' \rangle, \\ f'' &= 2\langle J', J' \rangle + 2\langle J, J'' \rangle \\ &= 2\|J'\|^2 + 2\langle J, J'' \rangle, \end{aligned}$$

e

$$\begin{aligned} f(0) &= \|J(0)\|^2 = 0, \\ f'(0) &= 2\langle J(0), J'(0) \rangle = 0, \\ f''(0) &= 2\|J'(0)\|^2 + 2\langle J(0), J''(0) \rangle = 2\|J'(0)\|^2 = 2. \end{aligned}$$

Como

$$\begin{aligned} f''' &= 2\langle J'', J' \rangle + 2\langle J', J'' \rangle + 2\langle J', J'' \rangle + 2\langle J, J''' \rangle \\ &= 6\langle J', J'' \rangle + 2\langle J, J''' \rangle \end{aligned}$$

e

$$J''(0) = -R(J(0), \gamma'(0))\gamma'(0) = -R(0, V)V = 0,$$

segue que

$$f'''(0) = 6\langle J'(0), J''(0) \rangle + 2\langle J(0), J'''(0) \rangle = 0.$$

Falta apenas calcular a última derivada. Temos

$$\begin{aligned} f^{(4)} &= 6 \langle J'', J'' \rangle + 6 \langle J', J''' \rangle + 2 \langle J', J''' \rangle + 2 \langle J, J^{(4)} \rangle \\ &= 6 \|J''\|^2 + 8 \langle J', J''' \rangle + 2 \langle J, J^{(4)} \rangle, \end{aligned}$$

de modo que

$$f^{(4)}(0) = 8 \langle J'(0), J'''(0) \rangle = 8 \langle W, J'''(0) \rangle.$$

Mas

$$\begin{aligned} \langle J'''(0), W \rangle &= \left\langle \frac{D}{dt} J''(t) \Big|_{t=0}, W \right\rangle \\ &= - \left\langle \frac{D}{dt} R(J(t), \gamma'(t)) \gamma'(t) \Big|_{t=0}, W \right\rangle. \end{aligned}$$

Afirmamos que

$$\left\langle \frac{D}{dt} R(J(t), \gamma'(t)) \gamma'(t) \Big|_{t=0}, W \right\rangle = R(V, W, W, V).$$

De fato, como

$$\begin{aligned} \langle R(J, \gamma') \gamma', J' \rangle &= R(J, \gamma', \gamma', J') \\ &= R(\gamma', J', J, \gamma') \\ &= \langle R(J', \gamma') \gamma', J \rangle, \end{aligned}$$

segue que

$$\frac{d}{dt} \langle R(J, \gamma') \gamma', J' \rangle = \frac{d}{dt} \langle R(J', \gamma') \gamma', J \rangle,$$

onde

$$\left\langle \frac{D}{dt} R(J, \gamma') \gamma', J' \right\rangle + \langle R(J, \gamma') \gamma', J'' \rangle = \left\langle \frac{D}{dt} R(J', \gamma') \gamma', J \right\rangle + \langle R(J', \gamma') \gamma', J' \rangle.$$

Calculando em $t = 0$, como $J(0) = J''(0) = 0$ e $J'(0) = W$, temos que

$$\left\langle \frac{D}{dt} R(\gamma'(t), J(t)) \gamma'(t) \Big|_{t=0}, W \right\rangle = \langle R(W, V) V, W \rangle = R(W, V, V, W).$$

Consequentemente,

$$f^{(4)}(0) = -8R(V, W, W, V).$$

Segue portanto da fórmula de Taylor que

$$f(t) = t^2 - \frac{1}{3}R(V, W, W, V)t^4 + o(t^4).$$

Quando, $\|V\| = \|W\| = 1$, temos $R(V, W, W, V) = K(V, W)$ e portanto

$$f(t) = t^2 - \frac{1}{3}K(V, W)t^4 + o(t^4).$$

Para provar a última expressão do enunciado, seja

$$g(t) = \|J(t)\| = \sqrt{f(t)}.$$

Então $f(t) = [g(t)]^2$ e

$$\begin{aligned} f'(t) &= 2g(t)g'(t), \\ f''(t) &= 2[g'(t)]^2 + 2g(t)g''(t), \\ f'''(t) &= 6g'(t)g''(t) + 2g(t)g'''(t), \\ f^{(4)}(t) &= 6[g''(t)]^2 + 8g'(t)g'''(t) + 2g(t)g^{(4)}(t). \end{aligned}$$

Daí

$$\begin{aligned} g(0) &= 0, \\ f''(0) &= 2[g'(0)]^2 + 2g(0)g''(0) \implies 2 = 2[g'(0)]^2 + 0 \\ &\implies g'(0) = 1, \\ f'''(0) &= 6g'(0)g''(0) + 2g(0)g'''(0) \implies 0 = 6g''(0) + 0 \\ &\implies g''(0) = 0, \\ f^{(4)}(0) &= 6[g''(0)]^2 + 8g'(0)g'''(0) + 2g(0)g^{(4)}(0) \implies -8K(V, W) = 0 + 8g'''(0) + 0 \\ &\implies g'''(0) = -K(V, W). \end{aligned}$$

Segue da fórmula de Taylor que

$$g(t) = t - \frac{1}{6}K(V, W)t^3 + o(t^3).$$

■

Este resultado produz a relação entre a aproximação ou espalhamento das geodésicas em um ponto de acordo com a curvatura seccional neste ponto. De fato, considere

$$F(s, t) = \exp_p(tv(s)),$$

onde $v(s)$ é uma curva de $T_p M$ com $v(0) = v$, $v'(0) = w$ e $\|v(s)\| = \|w\| = 1$. Considere os raios partindo da origem, eles se afastam do raio tv com velocidade absoluta

$$\left\| \frac{\partial}{\partial s} tv(s) \Big|_{s=0} \right\| = \|tv'(0)\| = t\|w\| = t.$$

Agora considere as geodésicas radiais $F_s(t)$ partindo de p . Como

$$J(t) = d(\exp_p)_{tv} tw = \frac{\partial F}{\partial s}(0, t)$$

é um campo de Jacobi que satisfaz

$$\begin{aligned} J(0) &= 0, \\ \frac{DJ}{dt}(0) &= w, \end{aligned}$$

ele satisfaz as hipóteses do enunciado do Lema 18.14 e portanto

$$\left\| \frac{\partial F}{\partial s}(0, t) \right\| = t - \frac{1}{6}K(V, W)t^3 + o(t^3).$$

Isso significa que as geodésicas radiais $F_s(t)$ se afastam da geodésica radial $F_0(t) = \exp_p(tv)$ com velocidade que difere de t por uma termo de terceira ordem cujo sinal é o oposto do sinal da curvatura seccional da variedade em p associada ao plano σ gerado por v e w . Isso significa que se $K_p(\sigma) > 0$ as geodésicas se afastam menos que os raios do espaço tangente $T_p M$, enquanto que se $K_p(\sigma) < 0$ as geodésicas se afastam mais que os raios do espaço tangente $T_p M$.

18.5 Pontos Conjugados

Agora vamos estudar as singularidades da aplicação exponencial através dos campos de Jacobi. A existência de uma tal relação é sugerida pela Proposição 18.9 e pode ser vista de maneira concreta na esfera \mathbb{S}_R^n . Todas as geodésicas que saem de um ponto p se encontram no seu ponto antipodal $-p$, que está a uma distância πR de p ao longo de qualquer uma destas geodésicas. A aplicação exponencial é um difeomorfismo sobre a bola geodésica $B_{\pi R}(0) \subset T_p M$ mas deixa de ser um difeomorfismo na fronteira da bola. E pela Proposição 18.11, os campos de Jacobi normais ao longo destas geodésicas que se anulam em p tem o seu próximo zero exatamente em $-p$. Por outro lado, a Proposição 18.10 mostra que em uma vizinhança normal (ou seja uma vizinhança que é a imagem de um conjunto onde \exp_p é um difeomorfismo) nenhum campo de Jacobi que se anula em p pode se anular em qualquer ponto da vizinhança. Assim deve se esperar existir uma relação entre as singularidades da aplicação exponencial e os zeros de um campo de Jacobi.

18.15 Definição. Seja $\gamma : [a, b] \rightarrow M$ uma geodésica. Dizemos que o ponto $q = \gamma(b)$ é **conjugado** ao ponto $p = \gamma(a)$ ao longo de γ se existe um campo de Jacobi J ao longo de γ não identicamente nulo que se anula em p e q , isto é, tal que

$$J(a) = J(b) = 0.$$

A **multiplicidade** do ponto conjugado q é a dimensão do subespaço dos campos de Jacobi que se anulam em p e q . \square

Pelo teorema de existência e unicidade, a dimensão do espaço dos campos de Jacobi que se anulam em p é n ; como o campo de Jacobi tangencial $J(t) = t\gamma'(t)$ só se anula em p (de fato, qualquer campo de Jacobi tangencial que se anula em dois pontos é necessariamente nulo pela Proposição 18.7), a multiplicidade de um ponto conjugado é no máximo $n - 1$. Em particular, na definição de pontos conjugados, o ponto q é conjugado a p se e somente se existir um campo de Jacobi **normal** ao longo de γ não identicamente nulo que se anula em p e q . O número $n - 1$ é atingido na esfera pela Proposição 18.11, já que para pontos antipodais existe um campo de Jacobi que anula neles para cada campo paralelo normal.

18.16 Proposição. Sejam $p \in M$, $v \in T_p M$ e $q = \exp_p(v)$. Então \exp_p é um difeomorfismo local em uma vizinhança de v se e somente se q não é conjugado a p ao longo da geodésica $\gamma(t) = \exp_p(tv)$.

Além disso, a multiplicidade de q é igual a $\dim \ker d(\exp_p)_v$.

Prova: Como vimos na Proposição 18.9, todo campo de Jacobi ao longo de γ tal que $J(0) = 0$ é da forma

$$J(t) = d(\exp_p)_{tv} tw.$$

para algum vetor $w \in T_p M$ e $J'(0) = w$. Portanto, q é conjugado a p se e somente se

$$J(1) = d(\exp_p)_v w = 0,$$

isto é, se e somente se v é ponto crítico de \exp_p . Como os campos de Jacobi J_1, \dots, J_k que se anulam em p são linearmente independentes se e somente se os vetores

$$J'_1(0) = w_1, \dots, J'_k(0) = w_k$$

são linearmente independentes em $T_p M$, isso conclui a demonstração. \blacksquare

18.17 Proposição. Seja $\gamma : [a, b] \rightarrow M$ uma geodésica tal que $\gamma(b)$ não é conjugado a $\gamma(a)$.

Dados $V \in T_{\gamma(a)} M$ e $W \in T_{\gamma(b)} M$, existe um único campo de Jacobi J ao longo de γ tal que

$$J(a) = V,$$

$$J(b) = W.$$

Prova: Seja \mathcal{J} o espaço vetorial dos campos de Jacobi ao longo de γ tais que $J(a) = 0$. Este espaço tem dimensão n , correspondente às n escolhas linearmente independentes de $\frac{DJ}{dt}(a)$. Defina uma aplicação linear $\Phi : \mathcal{J} \rightarrow T_{\gamma(b)}M$ por

$$\Phi(J) = J(b).$$

Como $\gamma(b)$ não é conjugado a $\gamma(a)$, Φ é injetiva. Já que $\dim \mathcal{J} = n = \dim T_{\gamma(b)}M$, segue que Φ é um isomorfismo. Logo existe $J_1 \in \mathcal{J}$ tal que

$$J_1(b) = W,$$

ou seja, existe um único campo de Jacobi J_1 ao longo de γ tal que

$$J_1(a) = 0,$$

$$J_1(b) = W.$$

Considerando o espaço vetorial $\tilde{\mathcal{J}}$ dos campos de Jacobi ao longo de γ tais que $J(b) = 0$ e a aplicação linear $\Psi : \tilde{\mathcal{J}} \rightarrow T_{\gamma(a)}M$ por

$$\Psi(J) = J(a),$$

concluímos de maneira análoga que existe um único campo de Jacobi J_2 ao longo de γ tal que

$$J_2(a) = V,$$

$$J_2(b) = 0.$$

O campo $J = J_1 + J_2$ é o campo de Jacobi que satisfaz as condições do enunciado. A unicidade segue imediatamente do fato de $\gamma(b)$ não ser conjugado a $\gamma(a)$, pois a diferença de dois campos de Jacobi diferentes que satisfazem as condições do enunciado seria um campo de Jacobi não trivial se anulando em $\gamma(a)$ e $\gamma(b)$.

■

Capítulo 19

Causalidade

19.1 Orientação Temporal

Em uma variedade de Lorentz arbitrária, em cada espaço tangente $T_p M$, como em todo espaço de Minkowski, existem dois cones temporais. Escolher um destes cones temporais equivale a orientar o espaço tangente $T_p M$. A questão é se é possível orientar todos os espaços tangentes de M de uma maneira contínua. Lembre-se que escolher um cone temporal equivale na prática a escolher um vetor do tipo tempo v_0 em $T_p M$; então, um vetor causal (isto é, um vetor que não é do tipo espaço) v está dirigido para o futuro se $\langle v_0, v \rangle < 0$ e para o passado se $\langle v_0, v \rangle > 0$.

19.1 Definição. Seja M uma variedade de Lorentz. Dizemos que um campo vetorial $X \in \mathcal{T}(M)$ é do **tipo tempo** se X_p é do tipo tempo para todo $p \in M$. \square

Em geral, uma variedade de Lorentz não possui campos vetoriais do tipo tempo globalmente definidos.

19.2 Definição. Dizemos que uma variedade de Lorentz M é **temporalmemente orientável** se existe uma função τ que atribui a cada ponto p um cone temporal τ_p em $T_p M$ e um campo vetorial do tipo tempo tais que $X_p \in \tau_p$ para todo $p \in M$.

Neste caso, dizemos que τ é uma **orientação temporal** para M .

Se $v \in T_p M$ é um vetor do tipo causal, dizemos que v é dirigido para o **futuro** se $v \in \bar{\tau}_p$ e para o **passado** se $-v \in \bar{\tau}_p$. \square

Em outras palavras, v é dirigido para o futuro se

$$\langle v, X_p \rangle < 0$$

e para o passado se

$$\langle v, X_p \rangle > 0.$$

Veja a demonstração do Teorema 19.12 para uma definição alternativa usando transporte paralelo.

19.3 Definição. Uma variedade de Lorentz conexa, temporalmente orientada, é chamada um **espaçotempo**. \square

Em Relatividade Geral se trabalha com espaçotempos, já que é altamente improvável que dois observadores, mesmo não iniciais, se encontrem em um ponto do espaçotempo e o tempo para um segue o seu ritmo natural em direção ao futuro, enquanto que para o outro observador o tempo corre em direção ao passado.

19.2 Teoria de Causalidade para Espaçotempos

De agora em diante, M denotará sempre um espaço-tempo. Em um espaço-tempo, uma curva diferenciável causal é sempre orientada para o futuro ou orientada para o passado, já que uma curva não pode mudar continuamente sua orientação temporal do futuro para o passado ou vice-versa.

19.2.1 Cronologia

19.4 Definição. Dados pontos $p, q \in M$, dizemos que p **precede** q **cronologicamente** e escrevemos

$$p \ll q$$

se existe uma curva diferenciável do tipo tempo orientada para o futuro ligando p a q .

Dizemos que p **precede** q **causalmente** e escrevemos

$$p \leq q$$

se $p = q$ ou se existe uma curva diferenciável causal orientada para o futuro ligando p a q .

Dizemos que p **precede** q **causalmente estritamente** e escrevemos

$$p < q$$

se $p \leq q$ e $p \neq q$. \square

As relações \ll e \leq são claramente transitivas.

19.5 Proposição. Se $p \ll q$ e $q \leq r$ ou se $p \leq q$ e $q \ll r$, então $p \ll r$.

Prova: D. ■

Estendemos a noção de orientação temporal para curvas contínuas:

19.6 Definição. Uma curva contínua $\alpha : (a, b) \rightarrow M$ é uma **curva causal orientada para o futuro** se para cada $t_0 \in (a, b)$ existe $\varepsilon > 0$ e uma vizinhança normal convexa V de $\alpha(t_0)$ contendo $\alpha(t_0 - \varepsilon, t_0 + \varepsilon)$ tal que para todos t_1, t_2 com $t_0 - \varepsilon < t_1 < t_2 < t_0 + \varepsilon$ existe uma curva causal orientada para o futuro em V de $\alpha(t_1)$ a $\alpha(t_2)$. \square

A necessidade de considerar vizinhanças normais convexas na Definição 18.6 se justifica pelo fato de existirem espaço-templos em que $p \ll q$ para todos os pontos $p, q \in M$, logo nestes espaço-tempos qualquer curva contínua α satisfaz $\alpha(t_1) \ll \alpha(t_2)$ e $\alpha(t_2) \ll \alpha(t_1)$ para todos t_1, t_2 no domínio de α , portanto apenas a condição $\alpha(t_1) \ll \alpha(t_2)$ para todos $t_1 < t_2$ implicaria que toda curva contínua é uma curva causal orientada para o futuro.

19.7 Definição. Dado $p \in M$, definimos:

$$\begin{aligned} \text{futuro cronológico de } p : \quad I^+(p) &= \{q \in M : p \ll q\}, \\ \text{passado cronológico de } p : \quad I^-(p) &= \{q \in M : q \ll p\}, \\ \text{futuro causal de } p : \quad J^+(p) &= \{q \in M : p \leq q\}, \\ \text{passado causal de } p : \quad J^-(p) &= \{q \in M : q \leq p\}. \end{aligned}$$

\square

19.3 Apêndice: O Recobrimento Duplo Temporamente Orientável

Toda variedade riemanniana não orientável possui um recobrimento duplo orientável (veja, por exemplo, [Carmo], último exercício do Capítulo 0). Um resultado análogo vale para variedades de Lorentz, como veremos. A construção do recobrimento duplo temporalmente orientável de uma variedade de Lorentz dada abaixo segue [BEE], pp. 51–52. É uma construção bem mais abstrata que a contida, por exemplo, em [ONeill], p. 194. E há alguns aspectos que este autor não conseguiu compreender (especificamente, Lema 19.10, que parece ser o resultado usado no Lema 19.11, que é o Lema 3.2 de [BEE]).

19.8 Definição. Seja M uma variedade de Lorentz, fixe um ponto $p_0 \in M$ e escolha uma orientação temporal $v_0 \in T_{p_0}M$. Para qualquer $p \in M$, definimos uma relação de equivalência no conjunto das curvas diferenciáveis por partes ligando p_0 a p

$$\Omega_{p_0,p} = \{\alpha \in [0, 1] \rightarrow M : \alpha \text{ é diferenciável por partes, } \alpha(0) = p_0 \text{ e } \alpha(1) = p\}$$

da seguinte forma: duas curvas $\alpha, \beta \in \Omega_{p_0,p}$ são equivalentes se os campos paralelos V_α, V_β ao longo de respectivamente α, β definidos por $V_\alpha(0) = V_\beta(0) = v_0$ satisfazem

$$\langle V_\alpha(1), V_\beta(1) \rangle < 0.$$

□

Esta relação de equivalência divide $\Omega_{p_0,p}$ em duas classes: como os campos V_α e V_β são paralelos, segue que $V_1(1)$ e $V_\beta(1)$ são vetores do tipo tempo em $T_p M$, pois

$$\begin{aligned} \langle V_\alpha(1), V_\alpha(1) \rangle &= \langle V_\alpha(0), V_\alpha(0) \rangle = \langle v_0, v_0 \rangle < 0, \\ \langle V_\beta(1), V_\beta(1) \rangle &= \langle V_\beta(0), V_\beta(0) \rangle = \langle v_0, v_0 \rangle < 0, \end{aligned}$$

logo ou eles estão no mesmo cone temporal ou estão em cones temporais opostos. Note que duas curvas diferenciavelmente homotópicas pertencem à mesma classe de equivalência (pois podemos definir um campo paralelo ao longo da imagem da homotopia, já que a equação diferencial do transporte paralelo é linear e a imagem da homotopia é compacta), embora curvas equivalentes não precisem ser necessariamente homotópicas.

Dada uma curva diferenciável $\alpha : [0, 1] \rightarrow M$ com $\alpha(0) = p_0$, denotaremos por $[\alpha]$ a classe de equivalência de α e

$$\widetilde{M} = \{[\alpha] : \alpha : [0, 1] \rightarrow M \text{ com } \alpha(0) = p_0\}.$$

19.9 Proposição. \widetilde{M} é uma variedade de Lorentz e $\pi : \widetilde{M} \rightarrow M$ definida por

$$\pi([\alpha]) = \alpha(1)$$

é um recobrimento duplo.

Prova: Dotamos \widetilde{M} da topologia quociente através desta relação de equivalência, o que torna \widetilde{M} um espaço de recobrimento topológico com $\pi : \widetilde{M} \rightarrow M$ a aplicação de recobrimento. Lembramos então que o espaço de recobrimento topológico de uma variedade diferenciável possui uma única estrutura de variedade diferenciável que torna a aplicação de recobrimento diferenciável (veja [Lee 1], Proposição 4.40, p. 92, para uma demonstração). Finalmente, definimos uma métrica de Lorentz em \widetilde{M} através do pullback da projeção, isto é,

$$\tilde{g} = \pi^* g,$$

ou seja,

$$\tilde{g}_{[\alpha]}(\tilde{v}, \tilde{w}) = g_{\alpha(1)}(d\pi_{[\alpha]}\tilde{v}, d\pi_{[\alpha]}\tilde{w}).$$

Em particular, π é uma isometria local. ■

Note que se M é temporalmente orientável, então $\widetilde{M} = M$. Caso contrário, mostraremos que \widetilde{M} é. Para isso, precisaremos dos lemas a seguir.

19.10 Lema. Sejam $\tilde{p}_0 \in \pi^{-1}(p_0)$ e $\tilde{\alpha} : [0, 1] \rightarrow \widetilde{M}$ é uma curva diferenciável por partes tal que

$$\tilde{\alpha}(0) = \tilde{p}_0.$$

Se $\alpha = \pi \circ \tilde{\alpha}$, então

$$\tilde{\alpha}(1) = [\alpha].$$

Prova: Este resultado parece ser incorreto pelo seguinte argumento. Suponha que

$$\pi^{-1}(p_0) = \{\tilde{p}_0^+, \tilde{p}_0^-\}.$$

Como ambos M e \widetilde{M} são variedades de Lorentz e toda variedade diferenciável é localmente contrátil, podemos assumir vizinhanças contráteis \tilde{V}^\pm de \tilde{p}_0^\pm tais que $\pi(\tilde{V}^\pm) = V$, V sendo uma vizinhança contrátil de p_0 .

Dentro das vizinhanças \tilde{V}^\pm podemos tomar laços $\tilde{\alpha}^\pm$ com pontos inicial e final em \tilde{p}_0^\pm . Se o enunciado é correto, então se $\alpha^\pm = \pi \circ \tilde{\alpha}^\pm$, segue que

$$\tilde{p}_0^\pm = \tilde{\alpha}^\pm(1) = [\alpha^\pm].$$

Em particular,

$$[\alpha^+] \neq [\alpha^-],$$

já que $\tilde{p}_0^+ \neq \tilde{p}_0^-$. Mas α^+ e α^- , sendo laços com pontos inicial e final em p_0 , dentro da vizinhança contrátil V , são (diferenciavelmente) homotópicos e portanto

$$[\alpha^+] = [\alpha^-],$$

uma contradição.

Se o resultado for correto, um possível argumento seria o seguinte. Para fixar idéias, assuma que \tilde{p}_0 é a classe de equivalência da curva constante $\alpha_0(t) \equiv p_0$. Neste caso, o campo paralelo à curva constante α_0 com vetor inicial v_0 é simplesmente $V(t) \equiv v_0$.

Defina $\tilde{\alpha}_t : [0, 1] \rightarrow \widetilde{M}$ por

$$\tilde{\alpha}_t(s) = \tilde{\alpha}(st),$$

de modo que $\tilde{\alpha}_t$ é simplesmente o segmento da curva $\tilde{\alpha}$ que liga \tilde{p}_0 a $\tilde{\alpha}(t)$, seu domínio reparametrizado para $[0, 1]$. Observe que $\tilde{\alpha}_0$ é a curva constante $\tilde{\alpha}(s) \equiv \tilde{p}_0$ e que $\tilde{\alpha}_1 = \tilde{\alpha}$. Seja $\alpha_t = \pi \circ \tilde{\alpha}_t$. Mostraremos que o conjunto

$$A = \{t \in [0, 1] : \tilde{\alpha}_t(1) = [\alpha_t]\}$$

é aberto e fechado em $[0, 1]$, provando o resultado. Por escolha, $0 \in A$. (...) ■

19.11 Lema. Sejam $\tilde{p}_0 \in \pi^{-1}(p_0)$ e $\tilde{v}_0 \in T_{\tilde{p}_0}\widetilde{M}$ o único vetor do tipo tempo tal que $d\pi_{\tilde{p}_0}\tilde{v}_0 = v_0$.

Sejam $\tilde{p} \in \widetilde{M}$ e $\tilde{\alpha}, \tilde{\beta} : [0, 1] \rightarrow \widetilde{M}$ curvas diferenciáveis por partes tais que

$$\begin{aligned} \tilde{\alpha}(0) &= \tilde{\beta}(0) = \tilde{p}_0, \\ \tilde{\alpha}(1) &= \tilde{\beta}(1) = \tilde{p}. \end{aligned}$$

Se $\tilde{V}_\alpha, \tilde{V}_\beta$ são os campos paralelos ao longo de respectivamente $\tilde{\alpha}, \tilde{\beta}$ definidos por $\tilde{V}_\alpha(0) = \tilde{V}_\beta(0) = \tilde{v}_0$, então

$$\langle \tilde{V}_\alpha(1), \tilde{V}_\beta(1) \rangle < 0.$$

Prova: A unicidade de \tilde{v}_0 segue do fato de toda aplicação de recobrimento diferenciável ser um difeomorfismo local. Sejam

$$\begin{aligned}\alpha &= \pi \circ \tilde{\alpha}, \\ \beta &= \pi \circ \tilde{\beta}.\end{aligned}$$

Como $\pi : \widetilde{M} \rightarrow M$ é uma isometria local, os campos

$$\begin{aligned}V_\alpha &= \pi_* (\tilde{V}_\alpha), \\ V_\beta &= \pi_* (\tilde{V}_\beta),\end{aligned}$$

são paralelos ao longo de α_1, α_2 , respectivamente, com

$$V_\alpha(0) = V_\beta(0) = \pi_*(\tilde{v}_0) = v_0$$

e

$$\langle V_\alpha(1), V_\beta(1) \rangle = \left\langle \pi_* (\tilde{V}_\alpha(1)), \pi_* (\tilde{V}_\beta(1)) \right\rangle = \langle \tilde{V}_\alpha(1), \tilde{V}_\beta(1) \rangle.$$

Suponha por absurdo que

$$\langle \tilde{V}_\alpha(1), \tilde{V}_\beta(1) \rangle \not< 0.$$

Como $\tilde{V}_\alpha(1), \tilde{V}_\beta(1)$ são vetores do tipo tempo, necessariamente $\langle \tilde{V}_\alpha(1), \tilde{V}_\beta(1) \rangle > 0$. Logo $\langle V_\alpha(1), V_\beta(1) \rangle > 0$ em p e por definição $[\alpha_1] \neq [\alpha_2]$. Por outro lado, pelo Lema 19.10, temos também que

$$\begin{aligned}\tilde{\alpha}(1) &= [\alpha], \\ \tilde{\beta}(1) &= [\beta],\end{aligned}$$

contradizendo $\tilde{\alpha}_1(1) = \tilde{\alpha}_2(1) = \tilde{p}$. ■

19.12 Teorema. *Toda variedade de Lorentz não temporalmente orientável possui um recobrimento duplo temporalmente orientável.*

Prova: Veja Teorema 3.3 em [BEE], p. 52. Implícita na demonstração deste teorema é uma definição alternativa para a orientabilidade temporal de uma variedade de Lorentz M : M é temporalmente orientável se, fixados um ponto base qualquer $p_0 \in M$ e um vetor do tipo tempo qualquer $v_0 \in T_{p_0}M$, para todo $p \in M$ e para todas as curvas diferenciáveis α, β ligando p_0 a p , os campos paralelos V_α, V_β ao longo de respectivamente α, β definidos por $V_\alpha(0) = V_\beta(0) = v_0$ satisfazem

$$\langle V_\alpha(1), V_\beta(1) \rangle < 0.$$

Esta condição implica que o transporte paralelo do cone futuro determinado por v_0 em p_0 para qualquer ponto p é independente do caminho escolhido de p_0 para p . Logo, uma escolha consistente de vetores do tipo tempo futuros pode ser feita por transporte paralelo a partir de p_0 . ■

Referências Bibliográficas

- [BEE] John K. BEEM, Paul E. EHRLICH e Kevin L. EASLEY, *Global Lorentzian Geometry*, 2nd. Ed., Marcel Dekker, 1996.
- [Benn-Tucker] Ian M. BENN e Robin W. TUCKER, *An Introduction to Spinors and Geometry with Applications in Physics*, Adam Hilger, 1987.
- [Berger] Marcel BERGER, *A Panoramic View of Riemannian Geometry*, Springer, 2002.
- [Carmo] Manfredo Perdigão do CARMO, *Geometria Riemanniana*, 2a. Edição, 1988.
- [Carmo2] Manfredo Perdigão do CARMO, *Differential Geometry of Curves and Surfaces*, Prentice Hall, 1976.
- [Dodson-Poston] C. T. J. DODSON e T. POSTON, *Tensor Geometry: the geometric viewpoint and its uses*, 2nd Edition, Graduate Texts in Mathematics 130, Springer, 1991.
- [Ferraro] Rafael FERRARO, *Einstein's Space-Time: An Introduction to Special and General relativity*, Springer, 2007.
- [Fleisch1] Daniel FLEISCH, *A Student's Guide to Maxwell's Equations*, Cambridge, 2008.
- [Fleisch2] Daniel FLEISCH, *A Student's Guide to Vectors and Tensors*, Cambridge University Press, 2012.
- [Gron-Naess] Oyvind GRON e Arne NAESS, *Einstein's Theory: A Rigorous Introduction for the Mathematically Untrained*, Springer, 2011.
- [Hoffman-Kunze] Kenneth HOFFMAN e Ray KUNZE, *Linear Algebra*, 2nd. Edition, Prentice-Hall, 1971.
- [Kobayashi] Shoshichi KOBAYASHI, *Transformation Groups in Differential Geometry*, Classics in Mathematics 70 (Reprint of the 1972 Edition), Springer, 1995.
- [Lang] Serge LANG, *Linear Algebra*, 3rd. Edition, Springer, 1987.
- [Lee 1] John M. LEE, *Introduction to Smooth Manifolds*, 2nd Edition, Graduate Texts in Mathematics 218, Springer, 2012.
- [Lee 2] John M. LEE, *Riemannian Manifolds: an Introduction to Curvature*, Graduate Texts in Mathematics 176, Springer, 1997.
- [Lima 1] Elon Lages de LIMA, *Curso de Análise vol. 2*, IMPA, 1981.
- [Lima 2] Elon Lages de LIMA, *Variedades Diferenciáveis*, IMPA, 2011.
- [McMahon] David McMAHON, *Relativity demystified*, McGraw-Hill, 2006.

- [MTW] Charles W. MISNER, Kip S. THORNE e John Archibald WHEELER, *Gravitation*, Freeman, 1973.
- [Naber] Gregory L. NABER, *The Geometry of Minkowski Spacetime: an introduction to the mathematics of the Special Theory of Relativity*, 2nd. Edition, Springer, 2012.
- [Okun1] Lev B. OKUN, *The concept of mass*, Physics Today, **42** (6), 31–36, 1989.
- [Okun2] Lev B. OKUN, *Energy and Mass in Relativity Theory*, World Scientific, 2009.
- [ONeill] Barrett O’NEILL, *Semi-Riemannian Geometry, with applications to Relativity*, Academic Press, 1983.
- [Pais] Abraham PAIS, *Subtle is the Lord: The Science and Life of Albert Einstein*, Oxford University Press, 1982.
- [Penrose] Roger PENROSE, *Techniques of Differential Topology in Relativity*, SIAM, 1972.
- [Petkov] Vesselin PETKOV, *Inertia and Gravitation: From Aristotle’s Natural Motion to Geodesics Worldlines in Curved Spacetime*, Minkowski Institute Press, 2012.
- [Rahaman] Farook RAHAMAN, *The Special Theory of Relativity: a mathematical approach*, Springer, 2014.
- [Rossmann] Wulf ROSSMANN, *Lectures on Differential Geometry*, homepage do autor, 2003.
- [Rowe] E. G. Peter ROWE, *Geometrical Physics in Minkowski Spacetime*, Springer, 2000.
- [Shutz] Bernard SHUTZ, *A First Course in General Relativity*, 2nd. Edition, Cambridge University Press, 2009.
- [Spivak] Michael SPIVAK, *A Comprehensive Introduction to Differential Geometry*, Vol. I, 2nd Edition, Publish or Perish, 1979.
- [Springer] *Springer Handbook of Spacetime*, Abhay Ashtekar e Vesselin Petkov, Editores, Springer, 2014.
- [Steenrod] Norman STEENROD, *The Topology of Fibre Bundles*, Princeton University Press, 1951.
- [Stephani] Hans STEPHANI, *Relativity, an Introduction to Special and General Relativity*, 3rd. Edition, Cambridge University Press, 2004.
- [Takeuchi] Tatsu TAKEUCHI, *An Illustrated Guide to Relativity*, Cambridge University Press, 2010.
- [Taylor-Wheeler] Edwin TAYLOR e John Archibald WHEELER, *Spacetime Physics: Introduction to Special Relativity*, 2nd. Edition, Freeman, 1992.
- [Topper] David R. TOPPER, *How Einstein Created Relativity out of Physics and Astronomy*, Springer, 2013.
- [Tsamparlis] Michael TSAMPARLIS, *Special Relativity: An Introduction with 200 Problems and Solutions*, Springer, 2010.
- [Warner] Frank W. WARNER, *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Mathematics 94, Springer, 1983.
- [Woodhouse] Nicholas M. J. WOODHOUSE, *Special Relativity*, Springer, 2002.