Introdução à Teoria da Relatividade Geral

Pedro Henrique Dalprá

4 de dezembro de 2025

Este documento reúne notas sobre os aspectos fundamentais da formulação matemática e física da Teoria da Relatividade Geral, baseadas na obra *Sapacetime and Geometry*, de Sean M. Carroll (2004). Para um estudo mais aprofundado, recomenda-se a leitura integral da obra citada ou de livros consagrados sobre o tema. Além disso, uma familiaridade prévio com Relatividade Especial facilita substancialmente o entendimento dos tópicos aqui apresentados.

1 Relatividade especial e espaço-tempo plano

A maioria dos fatos geométricos interessantes sobre o plano são independentes da nossa escolha de coordenadas. Como um exemplo simples podemos considerar a distância entre dois pontos, dada por:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2. \tag{1}$$

Em um sistema de coordenadas x' e y' que são rotacionadas em relação ao sistema de coordenadas original, a fórmula da distância permanece inalterada:

$$(\Delta s)^2 = (\Delta x')^2 + (\Delta y')^2. \tag{2}$$

Dizemos, portanto, que a distância é invariante sob tais mudanças de coordenadas.

Um evento é definido como um único instante no espaço e no tempo, caracterizado por (t, x, y, z). Então o intervalo espaço-temporal entre dois eventos é:

$$(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2. \tag{3}$$

Se estabelecermos um novo referencial inercial (t', x', y', z'), o intervalo permanece inalterado:

$$(\Delta s)^2 = -(c\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2. \tag{4}$$

É por isso que faz sentido pensar na Relatividade Especial como uma teoria do espaço-tempo quadridimensional, conhecido como espaço de Minkowski.

Vamos introduzir uma notação conveniente para vetores:

$$x^\mu : \begin{aligned} x^0 &= ct, \\ x^1 &= x, \\ x^2 &= y, \\ x^3 &= z. \end{aligned}$$

Às vezes será útil referir-se às componentes de espaço e tempo de x^μ separadamente, portanto usaremos sobrescritos latinos para representar apenas as componentes espaciais:

$$x^i : \begin{aligned} x^1 &= x, \\ x^2 &= y, \\ x^3 &= z. \end{aligned}$$

Também é conveniente escrever o intervalo espaço-temporal de forma mais compacta, portanto introduziremos uma matriz 4×4, chamada de métrica:

$$n_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Temos então

$$(\Delta s)^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu. \tag{5}$$

Esta fórmula introduz a convenção de somatório, na qual os índices são somados. Note que esta expressão recupera o intervalo espaço-temporal do espaço de Minkowski:

$$\begin{aligned} (\Delta s)^2 &= \eta_{00}\Delta x^0 \Delta x^0 + \eta_{10}\Delta x^1 \Delta x^0 + \cdots + \eta_{33}\Delta x^3 \Delta x^3 \\ &= -1 \cdot c\Delta t \cdot c\Delta t + 0 \cdot \Delta x \cdot c\Delta t + \cdots + 1 \cdot \Delta z \cdot \Delta z \\ &= -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2. \end{aligned} \tag{6}$$

Porém, nem todas as trajetórias são suficientemente boas para serem construídas a partir de segmentos de linhas retas. Em circunstâncias mais gerais, é útil introduzir o intervalo infinitesimal, ou elemento de linha:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{7}$$

1.1 Transformações de Lorentz

Surge então uma questão fundamental: quais transformações mantêm ds^2 invariante? Um tipo simples são as translações, que apenas deslocam as coordenadas:

$$x^\mu \to x^{\mu'} = x^\mu + a^\mu, \tag{8}$$

onde a^μ é um conjunto de quatro números fixos. As translações deixam as diferenças Δx^μ inalteradas, portanto é notável que o intervalo Δs^2 permaneça inalterado.

O outro único tipo de transformação é multiplicar x^μ por uma matriz (independente do espaço-tempo):

$$x^{\mu'} = \Lambda^{\mu'}_{\mu} x^\mu \tag{9}$$

ou, convencionalmente,

$$x' = \Lambda x. \tag{10}$$

As matrizes que satisfazem essa transformação são conhecidas como transformações de Lorentz. O exemplo mais simples é uma rotação no plano xy:

$$\Lambda^{\mu'}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Exercício 1. Mostre que o intervalo ds^2 permanece invariante sob a transformação representada pela matriz de rotação acima.

Resolução: dx^μ se transforma pela regra

$$dx^{\mu'} = \Lambda^{\mu'}_{\mu} dx^\mu, \tag{11}$$

escrevendo na forma matricial, temos

$$\begin{pmatrix} cdt' \\ dx' \\ dy' \\ dz' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix}, \tag{12}$$

na forma de sistemas, escrevemos

$$\begin{cases} cdt' = 1 \cdot cdt + 0 \cdot dx + 0 \cdot dy + 0 \cdot dz \\ dx' = 0 \cdot cdt + \cos\theta \cdot dx + \sin\theta \cdot dy + 0 \cdot dz \\ dy' = 0 \cdot cdt - \sin\theta \cdot dx + \cos\theta \cdot dy + 0 \cdot dz \\ dz' = 0 \cdot cdt + 0 \cdot dx + 0 \cdot dy + 1 \cdot dz \end{cases} \tag{13}$$

$$\therefore \begin{cases} dt' = dt \\ dx' = \cos\theta dx + \sin\theta dy \\ dy' = -\sin\theta dx + \cos\theta dy \\ dz' = dz \end{cases} \tag{14}$$

por fim, note que

$$dt'^2 = dt^2, \qquad dx'^2 = dx^2, \qquad dy'^2 = dy^2 \qquad \text{e} \qquad dz'^2 = dz^2. \tag{15}$$

Logo,

$$ds'^2 = ds^2. \tag{16}$$

Quod erat demonstrandum!
(Como queríamos demonstrar!)

1.2 Vetores

Os vetores no espaço de Minkowski são frequentemente chamados de quadrivetores. Nesse espaço, além de não existir produto vetorial entre dois quadrivetores, perde-se o sentido em pensar na translação desses quadrivetores no espaço-tempo. Esses não são conceitos úteis em relatividade. Em vez disso, a cada ponto p no espaço-tempo associamos o conjunto de todos os vetores possíveis localizados nesse ponto. Esse conjunto é conhecido como espaço tangente T_p.

Muitas vezes é útil para propósitos concretos decompor vetores em componentes em relação a algum conjunto de vetores de base. Uma base é qualquer conjunto de vetores que tanto gera o espaço vetorial quanto é linearmente independente. Qualquer vetor é uma combinação linear de vetores da base. Para qualquer espaço vetorial dado, haverá um número infinito de bases legítimas. Para um espaço tangente associado a um ponto no espaço de Minkowski, a dimensão é obviamente 4.

Qualquer vetor A pode ser escrito como uma combinação linear de vetores da base:

$$A = A^{\mu}(\hat{e}_{(\mu)}). \tag{17}$$

Uma curva ou trajetória parametrizada através do espaço-tempo é escrita como uma função de um parâmetro, $x^{\mu}(\lambda)$. O vetor tangente a essa curva possui componentes

$$V^{\mu} = \frac{dx^{\mu}}{d\lambda}. \tag{18}$$

O vetor completo é portanto

$$V = V^{\mu}(\hat{e}_{(\mu)}). \tag{19}$$

Sob uma transformação de Lorentz o parâmetro λ permanece inalterado, assim V^{μ} deve se transformar como:

$$V^{\mu} \rightarrow V^{\mu'} = \Lambda^{\mu'}_{\nu} V^{\nu}. \tag{20}$$

Logo

$$\begin{aligned} V = V^{\mu}(\hat{e}_{(\mu)}) &= V^{\nu'}(\hat{e}_{(\nu')}) \\ &= \Lambda^{\nu'}_{\mu} V^{\mu}(\hat{e}_{(\nu')}) \end{aligned} \tag{21}$$

$$\therefore \hat{e}_{(\mu)} = \Lambda^{\nu'}_{\mu} \hat{e}_{(\nu')}. \tag{22}$$

Usaremos a seguinte notação:

$$(\Lambda^{-1})^{\nu'}_{\mu} = \Lambda^{\mu}_{\nu'}. \tag{23}$$

ou

$$\Lambda^{\mu}_{\nu'} \Lambda^{\nu'}_{\rho} = \delta^{\mu}_{\rho}, \tag{24}$$

onde δ^{μ}_{ρ} é o delta de Kronecker

$$\delta^{\mu}_{\rho} = \begin{cases} 1, \text{ se } \mu = \rho \\ 0, \text{ se } \mu \neq \rho \end{cases}$$

1.3 Vetores duais

Uma vez que definimos um espaço vetorial, podemos definir um espaço vetorial associado, chamado de espaço vetorial dual ou espaço cotangente. O espaço vetorial é denotado por T_p e o espaço cotangente por T_p^*.

Se $w \in T_p^*$, ele é chamado de vetor dual e age como uma transformação linear tal que

$$w(aV + bW) = aw(V) + bw(W) \in \Re, \tag{25}$$

onde V e W são vetores e a e b são números reais.

Se $n \in T_p^*$,

$$(an + bw)(V) = an(V) + bw(V). \tag{26}$$

T_p^* é o espaço de todas as transformações lineares do espaço vetorial para os números reais.

Vamos definir um conjunto de vetores base $\hat{\theta}^{(\nu)}$, de forma que:

$$\hat{\theta}^{(\nu)}(\hat{e}_{(\mu)}) = \delta^{\nu}_{\mu}. \tag{27}$$

Então cada vetor dual pode ser escrito como:

$$w = w_\mu \hat{\theta}^{(\mu)}. \tag{28}$$

Representamos $w \in T_p^*$ com w_μ e $V \in T_p$ com V^μ. Além disso vetores de T_p^* são chamados de vetores covariantes e vetores de T_p são chamados de contravariantes.

Escrevemos a ação de um vetor covariante em um vetor contravariante como:

$$\begin{aligned} w(V) &= w_\mu V^\nu \hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) \\ &= w_\mu V^\nu \delta^{\mu}_{\nu} \\ &= w_\mu V^\nu \in \Re. \end{aligned} \tag{29}$$

E a ação de um vetor contravariante em um vetor covariante:

$$V(w) = w_\mu V^\nu. \tag{30}$$

O espaço dual do próprio espaço dual é o espaço vetorial original. E a ação de um campo vetorial dual sobre um campo vetorial não é um único número, mas um escalar no espaço-tempo. Um escalar é uma grandeza que permanece inalterada sob transformações de Lorentz. Imagine o espaço de vetores-coluna com n componentes, então o espaço dual é o conjunto de vetores-linha com n componenetes, e a ação entre eles é dada pela multiplicação matricial:

$$w(V) = (w_1\, w_2\, \cdots\, w_n) \begin{pmatrix} V^1 \\ V^2 \\ \vdots \\ V^n \end{pmatrix} = w_i V^i.$$

No espaço-tempo, o exemplo mais simples de um vetor dual é o gradiente de uma função escalar, que denotamos por "d":

$$d\phi = \frac{\partial \phi}{\partial x^\mu} \hat{\theta}^{(\mu)}. \tag{31}$$

A regra da cadeia convencional usada para transformar derivadas parciais equivale, neste caso, à regra de transformação de componentes de vetores duais:

$$\begin{aligned} \frac{\partial \phi}{\partial x^\mu} &= \frac{\partial \phi}{\partial x^{\mu'}} \frac{\partial x^{\mu'}}{\partial x^\mu} \\ &= \frac{\partial \phi}{\partial x^{\mu'}} \Lambda^{\mu'}_{\mu}. \end{aligned} \tag{32}$$

O fato do gradiente ser um vetor dual leva à seguinte notação abreviada para derivadas parciais:

$$\frac{\partial \phi}{\partial x^\mu} = \partial_\mu \phi. \tag{33}$$

1.4 Tensores

Uma generalização direta de vetores e vetores duais é a noção de tensor. Assim como um vetor dual é uma aplicação linear de vetores para $\Re$, um tensor de tipo (ou rank) $(k,\, l)$ é uma aplicação multilinear de uma coleção de vetores duais e vetores para $\Re$.

$$T : T_p^* \times \cdots \times T_p^* \times T_p \times \cdots \times T_p \to \Re \tag{34}$$

$$(k \text{ vezes}) \qquad\qquad (l \text{ vezes})$$

Aqui "$\times$" denota produto cartesiano. Multilinearidade significa que o tensor age linearmente em cada um dos seus argumentos; por exemplo, para um tensor do tipo (1, 1), temos

$$T(aw + b\eta, cV + dW) = acT(w, V) + adT(w, W) + bcT(\eta, V) + bdT(\eta, W). \tag{35}$$

Desse ponto de vista, um escalar é um tensor do tipo (0, 0), um vetor é um tensor do tipo (1, 0) e um vetor dual é um tensor do tipo (0, 1).

O espaço de todos os tensores de um tipo fixo forma um espaço vetorial; eles podem ser somados ou multiplicados por números reais. Para construir uma base para este espaço, precisamos definir uma nova operação conhecida como produto tensorial, denotada por $\otimes$. Se T é um tensor $(k,\, l)$ e S é um tensor $(m,\, n)$, definimos um tensor $(k+m,\, l+n)$ $T \otimes S$ por

$$\begin{aligned} T \otimes S(w^{(1)}, \ldots, w^{(k)}, \ldots, w^{(k+m)}, V^{(1)}, \ldots, V^{(n)}, \ldots, V^{(n+l)}) \\ = T(w^{(1)}, \ldots, w^{(k)}, V^{(1)}, \ldots, V^{(n)}) S(w^{(k+1)}, \ldots, k^{(k+m)}, V^{(n+1)}, \ldots, V^{(n+l)}), \end{aligned} \tag{36}$$

em outras palavras, primeiro aja T no conjunto apropriado de vetores duais e vetores, e então aja S no restante, e então multiplique as respostas. Observe que em geral $T \otimes S \neq S \otimes T$.

Agora é simples construir uma base para o espaço de todos os tensores $(k,\, l)$, tomando produtos tensoriais de vetores da base e vetores duais; essa base consistirá em todos os tensores da forma

$$\hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_l)}. \tag{37}$$

Em um espaço quadridimensional, haverá 4^{k+l} tensores de base no total. Na notação de componentes, escrevemos então nosso tensor arbitrário como

$$T = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} \hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_l)}. \tag{38}$$

A ação dos tensores em um conjunto de vetores e vetores duais segue o padrão:

$$T(w^{(1)}, \ldots, w^{(k)}, V^{(1)}, \ldots, V^{(l)}) = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_k} w^{(1)}_{\mu_1} \cdots w^{(k)}_{\mu_k} V^{(1)\nu_1} \cdots V^{(l)\nu_l}. \tag{39}$$

A transformação das componentes do tensor sob transformações de Lorentz pode ser derivada aplicando o que já sabemos sobre a transformação de vetores de base e vetores duais

$$T^{\mu'_1 \cdots \mu'_k}_{\nu'_1 \cdots \nu'_l} = \Lambda^{\mu'_1}_{\mu_1} \cdots \Lambda^{\mu'_k}_{\mu_k} \Lambda^{\nu'_1}_{\nu_1} \cdots \Lambda^{\nu'_l}_{\nu_l} T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l}. \tag{40}$$

Podemos aplicar um tensor a (todo ou parte de) outro tensor para obter um terceiro tensor. Por exemplo,

$$U^{\mu}_{\nu} = T^{\mu\rho}_{\sigma} S^{\sigma}_{\rho\nu} \tag{41}$$

é um tensor (1, 1).

A noção de tensores não exige muito esforço para ser dominada; trata-se apenas de manter os índices em ordem, e as regras para manipulá-los são muito naturais.

Vejamos alguns exemplos de tensores. Primeiro consideremos o exemplo anterior de vetores colunas e seus duais, vetores linhas. Sua ação sobre um par $(w,\, V)$ é dada por:

$$M(w,V) = (w_1, \dots, w_n) \begin{pmatrix} M^1_1 & M^1_2 & \cdots & M^1_n \\ M^2_1 & M^2_2 & \cdots & M^2_n \\ \vdots & \vdots & \cdots & \vdots \\ M^n_1 & M^n_2 & \cdots & M^n_n \end{pmatrix} \begin{pmatrix} V^1 \\ V^2 \\ \vdots \\ V^n \end{pmatrix} \tag{42}$$

$$= w_i M^i_j V^j. \tag{43}$$

Se preferir, pode pensar em tensores como matrizes com um número arbitrário de índices.

No espaço-tempo já vimos alguns exemplos de tensores sem os chamarmos assim. O exemplo mais familiar de um tensor (0, 2) é a métrica $\eta_{\mu\nu}$. A ação da métrica em dois vetores é tão útil que recebe seu próprio nome, o produto interno (ou produto escalar):

$$\eta(V,W) = \eta_{\mu\nu} V^\mu W^\nu = V \cdot W. \tag{44}$$

A norma de um vetor é definida como o produto interno do vetor por si mesmo:

$$\text{se } \eta V^\mu V^\nu \text{ é} \begin{cases} < 0, \ V^\mu \text{ é temporal} \\ = 0, \ V^\mu \text{ é nulo ou tipo-luz} \\ > 0, \ V^\mu \text{ é espacial} \end{cases} \tag{45}$$

Outro tensor é o delta de Kronecker δ^μ_ν, do tipo (1, 1).

Existem também a métrica inversa $\eta^{\mu\nu}$, um tensor do tipo (2, 0) definido como o inverno da métrica:

$$\eta^{\mu\nu}\eta_{\nu\rho} = \eta_{\rho\nu}\eta^{\nu\mu} = \delta^\mu_\rho. \tag{46}$$

Há também o tensor de Levi-Civita, um tensor (0, 4):

$$\epsilon_{\mu\nu\rho\sigma} = \begin{cases} +1 \text{ se } \mu\nu\rho\sigma \text{ é uma permutação par de } 0123 \\ -1 \text{ se } \mu\nu\rho\sigma \text{ é uma permutação ímpar de } 0123 \\ 0 \text{ caso contrário} \end{cases} \tag{47}$$

Todos os tensores apresentados, exceto o delta de Kronecker, dependem da métrica, isto é, dependem do sistema de coordenadas. Porém, no espaço-tempo plano, suas componentes permanecem inalteradas.

Um exemplo mais típico de tensor é o tensor de intensidade do campo eletromagnético. Todos sabemos que os campos eletromagnéticos são compostos pelo vetor de campo elétrico E^i e pelo vetor de campo magnético B^i. Na verdade, eles são componentes de um tensor (0, 2) $F_{\mu\nu}$, definido por

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = -F_{\nu\mu}. \tag{48}$$