Notas de aula

Relatividade Restrita

Daniel Augusto Turolla Vanzella

1 de dezembro de 2020

Parte I

O Palco

Capítulo 1

Montando o espaço-tempo de Minkowski

1.1 Linhalândia

1.2 Eventos e linhas-de-mundo

Para deixar claro quando estivermos nos referindo a pontos do espaço-tempo, ao invés de pontos do espaço euclidiano usual, os denominaremos de *eventos*. Fisicamente, um evento pode ser pensado como "algo" que, de acordo com nossa percepção fragmentada do espaço-tempo em "tempo" e "espaço", pode ser idealizado como tendo ocorrido "num ponto do espaço, num instante de tempo". A colisão de duas partículas, a emissão de um fóton por um átomo excitado, a absorção de uma partícula por um detector, a desintegração de um núcleo atômico; todos eles marcam uma localização no espaço-tempo e, assim, são exemplos de eventos. Talvez o(a) leitor(a) possa achar, num primeiro momento, que essa definição de evento faria com que o espaço-tempo fosse muito vazio deles. Porém, para se convencer do contrário, basta o(a) leitor(a) se imaginar em qualquer ponto do espaço, em qualquer instante de tempo, e perceber que dificilmente ele(a) não estaria, nesse ponto e nesse instante, recebendo alguma "informação" (raios de luz, por exemplo, ou fótons de alguma outra porção do espectro eletromagnético) vinda de (pelo menos) duas direções diferentes; ou seja, esse ponto no espaço e instante no tempo seria sinalizado pelo *evento* do cruzamento de (pelo menos) dois fótons particulares. Isso deve bastar para mostrar que o espaço-tempo é preenchido por eventos *físicos*. Ainda assim, é bastante útil adotarmos uma visão mais idealizada, na qual um evento não precisa necessariamente marcar a ocorrência *de fato* de um evento físico, mas apenas ter o potencial para tal. Isso é semelhante a como pensamos no espaço tridimensional usual: não é necessário que haja uma partícula de fato num dado ponto do espaço para nos referirmos a ele, como uma idealização. Com isso, o espaço-tempo será constituído pelo *conjunto de todos os eventos*, nessa acepção idealizada de evento. Por enquanto, denotaremos o espaço-tempo por $\mathcal{E}$.

É importante deixar claro que embora a descrição (e às vezes até a interpretação) de um evento possa depender de observador, sua existência é algo absoluto. Um observador $\mathcal{O}$ pode descrever a colisão de duas certas partículas como tendo ocorrido no ponto $P \in \mathbb{E}^3$ do espaço (euclidiano), no instante $t \in \mathbb{R}$, e, assim, achar natural descrever esse evento pelo par ordenado $(t, P) \in \mathbb{R} \times \mathbb{E}^3$. Já um outro observador, $\widetilde{\mathcal{O}}$, de acordo com sua convenção, pode descrever *a mesma* colisão como tendo ocorrido no ponto $Q \in \mathbb{E}^3$ do espaço, no instante $\tilde{t} \in \mathbb{R}$, e, assim, descrevê-lo por $(\tilde{t}, Q) \in \mathbb{R} \times \mathbb{E}^3$. Apesar das descrições serem possivelmente diferentes [pois, em geral, $(t, P) \neq (\tilde{t}, Q)$], ambas se referem ao mesmo ponto (i.e., evento) no espaço-tempo, digamos $p \in \mathcal{E}$. Isso deixa claro que não devemos confundir ou identificar rigidamente um evento (nesse caso, $p \in \mathcal{E}$) no espaço-tempo com sua descrição em termos de localização temporal (nesse caso, t ou $\tilde{t}$) e espacial (nesse caso, P ou Q) — e, de fato, boa parte das situações aparentemente paradoxais, comuns em Relatividade, advém dessa identificação dependente de observador. Colocando de maneira mais simbólica, embora possamos fazer associações $\mathcal{E} \ni p \mapsto (t, P) \in \mathbb{R} \times \mathbb{E}^3$ (de infinitas maneiras diferentes), *não* devemos fazer a identificação rígida $p = (t, P)$. Essa distinção é totalmente análoga ao que acontece com, por exemplo, o plano euclidiano $\mathbb{E}^2$: embora possamos (e seja muito conveniente!) *descrever* um dado ponto $P \in \mathbb{E}^2$ através de um par ordenado $(x, y) \in \mathbb{R}^2$ (ou seja, *coordenadas*), devemos ter em mente que $P \neq (x, y)$, pois há infinitas maneiras diferentes de expressarmos o mesmo ponto P em termos de coordenadas.

Introduzida a noção de evento, passemos agora à de *linha-de-mundo*. Assim como um objeto puntual se movendo no espaço euclidiano descreve uma *trajetória* nesse espaço — o conjunto de todos os pontos pelos quais o objeto passa em algum momento —, o mesmo acontece quando consideramos a história desse objeto no espaço-tempo. Para diferenciar trajetórias espaciais de trajetórias espaço-temporais, denominaremos estas últimas de *linhas-de-mundo*: o conjunto de todos os eventos "ocupados" pelo objeto ao longo de sua história. Note que enquanto uma trajetória espacial pode se resumir a um único ponto (representando algo "parado no espaço"), o inexorável passar do tempo obriga qualquer objeto — que exista por mais do que apenas um mero instante — a possuir uma linha-de-mundo que se estende na "direção temporal" (vide **Fig. 1.1**). Mas para dar sentido preciso a "direção temporal", entre outros conceitos, precisamos equipar $\mathcal{E}$ com algumas estruturas adicionais — o que faremos mais adiante.

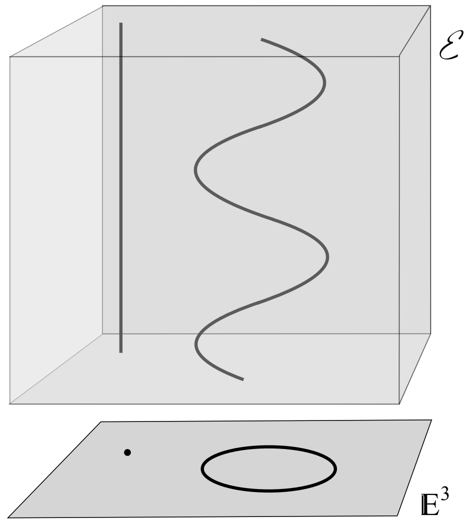

Figura 1.1: Representação de trajetórias circular e puntual no espaço e as respectivas linhas-de-mundo no espaço-tempo.

1.3 Requisitos físicos

Um espaço-tempo é mais do que apenas o mero conjunto de todos os eventos — assim como o espaço euclidiano é mais do que apenas o conjunto de seus pontos. Há estruturas adicionais definidas em $\mathcal{E}$. Se fôssemos adotar uma postura "generalista", definindo primeiro aquelas estruturas comuns a qualquer espaço-tempo e, então, acrescentando, "camada por camada", estruturas cada vez mais particulares até chegarmos ao espaço-tempo específico da Relatividade Restrita — o *espaço-tempo de Minkowski* —, teríamos que começar apresentando boa parte do arcabouço da *Geometria Diferencial*. Isso nos desviaria demais do objetivo desta disciplina e, por isso, adotaremos uma postura diferente: já muniremos $\mathcal{E}$ com as estruturas matemáticas específicas que fazem dele o espaço-tempo de Minkowski $\mathbb{M}$. Antes, porém, de munirmos $\mathcal{E}$ com as estruturas matemáticas necessárias, vamos tentar motivá-las fisicamente.

Uma característica primordial que diferencia um espaço-tempo de outro espaço qualquer é o fato que, a partir de todo e qualquer evento, tem que existir um conceito de "direção temporal" — que daria a direção, no espaço-tempo, do inexorável "passar do tempo", ao longo da qual uma *medida* de tempo deve existir —, enquanto outras direções devem representar "direções espaciais" — ao longo das quais nossa noção de distância euclidiana tem que fazer sentido. Considerando que as noções de "repouso" e "movimento" são relativas — como estabelecido desde Galileu para fenômenos mecânicos e estendido por Einstein a fenômenos eletromagnéticos; esse é o conteúdo do

Princípio da Relatividade —, qualquer que seja a linha-de-mundo de um observador, ela terá que determinar uma possível "direção temporal" do espaço-tempo em cada evento pelo qual ela passa, pois qualquer um pode se considerar "em repouso no espaço" — e, portanto, só evoluindo no tempo — em relação a si mesmo. Assim, vemos que deve haver não apenas uma direção temporal, mas infinitas, sendo que linhas-de-mundo passando por um mesmo evento $p \in \mathcal{E}$ em direções temporais diferentes representam partículas se movendo uma em relação à outra. E, pelo Princípio da Relatividade, nenhuma dessas direções temporais pode ser privilegiada em relação às demais.

Falar de "direção" naturalmente nos remete ao conceito de *vetor*, o que sugere fortemente que cada evento $p \in \mathcal{E}$ terá que ter associado a si um espaço vetorial, $\mathbb{V}_p$, cujos elementos $\mathbf{u} \in \mathbb{V}_p$ codificarão informação sobre diferentes direções do espaço-tempo naquele evento — ou seja, cada $\mathbf{u} \in \mathbb{V}_p$ deve ter alguma relação de "apontar para" outros eventos de $\mathcal{E}$ a partir de p. Assim, cada $\mathbb{V}_p$ — denominado *espaço-tangente* a $\mathcal{E}$ em p — deve ser equipado com alguma estrutura, além da de um mero espaço vetorial, para distinguir as direções temporais das espaciais. Guardemos isso em mente, por enquanto.

Outra propriedade física que é impressa na estrutura do espaço-tempo relaciona-se ao conceito de *inércia*. Por alguma razão — para mim, misteriosa —, partículas livres de qualquer ação externa seguem trajetórias (e linhas-de-mundo) muito particulares. Esse é um fato conhecido desde Galileu, que Newton codificou na sua *1ª lei do movimento* e que continua válido no contexto relativístico. A *lei da inércia*, como é conhecida a 1ª lei de Newton, normalmente é colocada em termos da afirmação que um corpo livre de qualquer ação externa mantém seu "estado de movimento", prosseguindo segundo um *movimento retilíneo e uniforme* (ou permanecendo em repouso, caso este fosse seu "estado de movimento" inicial). Evidentemente, para se fazer essa afirmação, deve-se adotar um tipo especial de *referencial* em relação ao qual a posição da partícula é medida, como função do tempo, para se testar sua uniformidade e retidão. Uma escolha "ruim" de referencial tornará essa formulação da lei da inércia inútil, pois mesmo uma partícula livre pode ter um movimento complicado quando visto de um referencial arbitrário — tente pensar no movimento de uma partícula livre visto a partir de um referencial girante[1]. Mas podemos despir a lei da inércia desses aspectos "relativos" e enxergar seu conteúdo *absoluto* (i.e., independente de

[1]É inevitável a pergunta: qual é o tipo "certo" de referencial em relação ao qual o movimento de partículas livres é retilíneo e uniforme? A resposta pode parecer decepcionante: o tipo "certo" é aquele no qual partículas livres têm movimento retilíneo e uniforme! — por definição, um *referencial inercial*. Apesar de aparentemente tautológico, a *existência* de um tal referencial é algo não-trivial e *esse* é o conteúdo físico dessa formulação da lei da inércia. Essa dose de circularidade nas definições de "partículas livres" e "referenciais inerciais" nos iludiu por mais de dois séculos, até que sua compreensão acabou por levar a um entendimento mais profundo sobre a *gravidade*, na chamada *Teoria da Relatividade Geral*.

referencial):

(i) A linha-de-mundo de uma partícula livre que passa pelo evento $p \in \mathcal{E}$ depende apenas do "estado de movimento" da partícula em p (ou seja, da informação de um trecho arbitrariamente pequeno da linha-de-mundo em torno de p). Logo, duas partículas livres que tenham suas linhas-de-mundo coincidindo em algum trecho em torno de $p \in \mathcal{E}$ terão linhas-de-mundo coincidentes em toda sua extensão;

(ii) O "estado de movimento" em $p \in \mathcal{E}$ resume-se à direção, no espaço-tempo, dessa linha-de-mundo em p;

(iii) Partículas livres inicialmente em repouso uma em relação à outra permanecerão em repouso relativo.

Os três pontos acima estão em ordem crescente de especificidade. O ponto (i) simplesmente diz que partículas livres possuem trajetórias/linhas-de-mundo particulares, independentes das características intrínsecas das partículas, e que podem ser determinadas com "informações locais". Já o ponto (ii) especifica qual o tipo de "informação local" que caracteriza o "estado de movimento" da partícula: a direção da linha-de-mundo no espaço-tempo. Podemos usar esses dois pontos, em conjunto, para *definir* uma identificação de direções em eventos diferentes: se os eventos $p, q \in \mathcal{E}$ são conectados por uma linha-de-mundo de uma partícula livre (por brevidade, uma linha-de-mundo *inercial*), então a direção da linha-de-mundo inercial em p é a mesma que em q. Note que, com essa identificação, linhas-de-mundo inerciais são, por definição, "linhas retas" no espaço-tempo (no sentido que são curvas que têm a mesma direção em todos os seus pontos).

Por fim, o ponto (iii) é o que exige — e, portanto, impõe — mais restrições sobre a estrutura do espaço-tempo. Uma leitura geométrica desse ponto é que a identificação entre direções do espaço-tempo em eventos diferentes — que foi estabelecida acima apenas para direções particulares de eventos conectados por uma linha-de-mundo inercial — pode ser estendida, de maneira consistente, para eventos espacialmente separados: direções em p e q serão identificadas entre si se, e somente se, estiverem associadas a linhas-de-mundo de partículas livres que estejam uma em repouso em relação à outra. Logo, partículas em repouso uma em relação a outra possuem, por definição, linhas-de-mundo *paralelas*.

Como dissemos acima, os pontos (i), (ii) e (iii) estão em ordem crescente de especificidade. Isso significa que o ponto (ii) se constrói sobre o ponto (i) e que o ponto (iii) se constrói sobre os pontos (i) e (ii) em conjunto. O ponto (ii) está intimamente relacionado ao fato que a conexão entre *cinemática* (i.e., as *características* do movimento) e *dinâmica* (i.e., as *causas* do movimento) se dá no nível da *aceleração* (*2ª lei do movimento* de Newton). Não é difícil pensar num hipotético universo em que a conexão se desse em outro nível (velocidade, taxa de variação da aceleração, etc.), caso em que (i) poderia continuar sendo válido mas o significado de "estado de movimento" seria outro. Também podemos imaginar um universo em que (i) e (ii) são válidos mas não (iii); um espaço-tempo no qual linhas-de-mundo inerciais ainda sejam "retas" mas que não seja possível estabelcer um "paralelismo" à distância[2]. Em $\mathbb{M}$ adotaremos a validade de todos esses pontos.

Mas ainda falta um ponto chave. Os requisitos discutidos até agora — distinção entre direções temporais e espaciais e validade da lei da inércia, como conhecida desde Galileu — se fazem presentes tanto no contexto da Relatividade Restrita quanto no da Mecânica Newtoniana (quando esta última é colocada na linguagem de espaço-tempo; vide Exercício ⑨). O que diferencia esses contextos é que a Relatividade Restrita adota como *postulado* o fato — muito bem estabelecido experimentalmente — que a *velocidade da luz no vácuo* é uma constante universal: não importa qual seja o movimento relativo entre fonte de luz e observador, a velocidade da luz (no vácuo), medida *localmente*[3] pelo observador, tem sempre o mesmo valor (denotado por c).[4] Visto na linguagem de espaço-tempo, isso significa duas coisas:

(a) Considere que um *flash* de luz seja emitido num evento $p \in \mathcal{E}$ qualquer. Os eventos pelos quais a luz desse *flash* passa independem da linha-de-mundo da fonte que emitiu o flash. Logo, as linhas-de-mundo da luz emanando de um evento $p \in \mathcal{E}$ denunciam e demarcam uma estrutura *absoluta* do próprio espaço-tempo — que denominaremos *cone-de-luz futuro* de p, representado por $\mathcal{C}_p^+$ (vide **Fig. 1.2**; o *cone-de-luz passado* de p, $\mathcal{C}_p^-$, é definido de maneira análoga, demarcado pelas linhas-de-mundo de luz que podem chegar ao evento p);

(b) Considere o evento $q \in \mathcal{E}$ determinado pelo cruzamento de uma linha-de-mundo de luz com a linha-de-mundo de um observador qualquer. O valor da velocidade da luz que esse observador mede em q independe da linha-de-mundo do observador (e vale c; vide **Fig. 1.2**).

De (a) vemos que, além de haver direções temporais e espaciais em $\mathcal{E}$, há também direções especiais dadas pelas linhas-de-mundo da luz. E de (b) temos que, embora a inclinação relativa entre linhas-de-mundo de diferentes observadores caracterize o movimento relativo entre eles, a velocidade que

[2]Demoramos mais de dois séculos para perceber que, na verdade, *esse* é o *nosso* universo — mas isso seria uma discussão para a disciplina de Relatividade Geral.

[3]*Localmente* significa que o observador faz uso apenas de uma porção suficientemente pequena, em torno de sua posição, da trajetória do raio de luz para medir sua velocidade.

[4]Nota de "preciosismo" histórico: a rigor, o postulado adotado por Einstein em 1905, acerca da velocidade da luz, dizia respeito a sua *independência* em relação ao estado de movimento da fonte emissora. No entanto, combinado com o outro postulado adotado, o *Princípio da Relatividade de Einstein* — que, em resumo, dizia que nenhum fenômeno de mecânica ou eletromagnetismo poderia ser usado para se discernir entre dois referenciais inerciais —, conclui-se que a velocidade da luz também não pode depender do estado de movimento do observador. Logo, do ponto de vista lógico, tal distinção é irrelevante.

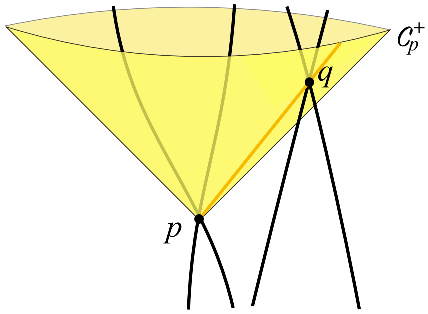

Figura 1.2: Diagrama espaço-tempo mostrando que a estrutura de cone-de-luz a partir de um evento p é independente do estado de movimento da fonte que emitiu a luz. Além disso, diferentes observadores medindo a velocidade da mesma linha-de-mundo de luz em q medem o mesmo valor c, embora os observadores tenham movimento relativo entre eles.

eles atribuem para a luz independe da inclinação relativa entre suas linhas-de-mundo e a da luz.

Agora que sabemos os requisitos físicos que desejamos impor em nosso espaço-tempo de Minkowski — todos motivados por fatos muito bem estabelecidos experimentalmente —, vamos às estruturas matemáticas que os implementam.

1.4 Estruturas matemáticas

Como dissemos anteriormente, muniremos $\mathcal{E}$ com as estruturas necessárias para fazer dele o espaço-tempo específico da Relatividade Restrita: o espaço-tempo de Minkowski $\mathcal{E} = \mathbb{M}$. Assim, já começaremos pela estrutura mais restritiva que, ao mesmo tempo, formalizará o conceito de "direção" em cada evento (os espaços tangentes $\mathbb{V}_p$ mencionados na seção anterior, $p \in \mathbb{M}$) e, ainda, proverá uma identificação consistente dessas direções em eventos diferentes. Tudo isso (e mais) é feito impondo-se que $\mathbb{M}$ seja um *espaço afim*.

Munir o conjunto de eventos de uma estrutura de espaço afim significa que podemos falar na separação entre eventos como sendo um *vetor* — exatamente como no caso de "segmentos orientados" no espaço euclidiano. Mais precisamente, consideraremos que existe um espaço vetorial (real) de dimensão d (no nosso caso de interesse, $d = 4$), que denotaremos por $\mathbb{V}$, e um mapeamento $\psi : \mathbb{M} \times \mathbb{M} \to \mathbb{V}$ que a cada par de eventos $p, q \in \mathbb{M}$ associa um vetor $\overrightarrow{pq} := \psi(p, q) \in \mathbb{V}$, satisfazendo as seguintes propriedades[5]:

[5]Utilizaremos a notação de vetor com uma "flecha" em cima *apenas* no contexto de segmentos orientados, por similaridade com o caso euclidiano.

(i) Fixado um evento qualquer $o \in \mathbb{M}$, o mapeamento $\psi(o, \cdot) : \mathbb{M} \to \mathbb{V}$ é bijetor (ou seja, além de existir o vetor $\overrightarrow{op} \in \mathbb{V}$ para qualquer $p \in \mathbb{M}$, qualquer vetor de $\mathbb{V}$ é da forma $\overrightarrow{op}$ para um, e apenas um, $p \in \mathbb{M}$);

(ii) Para quaisquer eventos $p, q, r \in \mathbb{M}$, $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$.

- **Exercício:** Mostre que o mapeamento ψ é antissimétrico; ou seja, que $\overrightarrow{pq} := \psi(p, q) = -\psi(q, p) =: -\overrightarrow{qp}$ para quaisquer $p, q \in \mathbb{M}$. Em particular, mostre que $\overrightarrow{pp} := \psi(p, p) = \mathbf{0}$ (o vetor nulo).

A propriedade (i) significa que, escolhido um evento qualquer como referência, podemos "colar" $\mathbb{V}$ sobre $\mathbb{M}$ — "colando" o vetor nulo $\mathbf{0}$ sobre o evento escolhido como referência —, conferindo ao próprio espaço-tempo uma estrutura de espaço vetorial. Já a propriedade (ii) fixa a relação entre "colagens" feitas em eventos diferentes adotados como referência. Essas propriedades nos permitem estabelecer uma noção bem específica de "igualdade de vetores em eventos distintos", pois todos os espaços tangentes são identificados como sendo o mesmo espaço vetorial: $\mathbb{V}_p = \mathbb{V}$, $p \in \mathbb{M}$. Assim, o conceito de "retas" — e, em particular, o de "linhas-de-mundo inerciais" — fica bem estabelecido: dado um evento $p \in \mathbb{M}$ e um vetor (não-nulo) $\mathbf{u} \in \mathbb{V}$, a reta que passa por p na direção de $\mathbf{u}$ é dada pelos eventos $q(\lambda) \in \mathbb{M}$ satisfazendo $\psi(p, q(\lambda)) = \lambda\mathbf{u}$, com $\lambda \in \mathbb{R}$. Por simplicidade de notação, representaremos essa reta por $q(\lambda) := p + \lambda\mathbf{u}$. Também a noção de paralelismo — e, em particular, de "repouso relativo" — fica bem definida: duas retas $q(\lambda) := p + \lambda\mathbf{u}$ e $s(\lambda) := r + \lambda\mathbf{v}$ são paralelas se existir $\alpha \in \mathbb{R}$ tal que $\mathbf{v} = \alpha\mathbf{u}$.

Elementos de $\mathbb{V}$ serão chamados de *quadri-vetores* (ou simplesmente 4-vetores) apenas para reforçar a dimensionalidade de $\mathbb{V}$ e para distingui-los de outros "tipos" de vetores com os quais trabalharemos. Mais adiante, teremos que adotar uma notação mais conveniente para fazer essa distinção.

Mas apenas a estrutura de espaço afim ainda não nos permite distinguir o espaço-tempo de Minkowski $\mathbb{M}$ de, por exemplo, um espaço euclidiano de 4 dimensões, $\mathbb{E}^4$ — pois ambos possuem essa mesma estrutura. Ainda temos que prover $\mathbb{M}$ — ou , mais especificamente, $\mathbb{V}$ — com uma estrutura que distingua "direções temporais" de "direções espaciais", como discutido no início da seção anterior, mas que, ao mesmo tempo, não privilegie uma direção temporal em relação às demais direções temporais (Princípio da Relatividade). Além disso, queremos introduzir noções de "distância espacial" (ao longo das direções espaciais) e "intervalo de tempo" (ao longo das direções temporais), de modo que se p e $q = p + \mathbf{u}$ são eventos separados por uma distância espacial D (ou por um intervalo de tempo T), então p e $r = p + \lambda\mathbf{u}$ ($\lambda \in \mathbb{R}$) estão separados por uma distância espacial $|\lambda|D$ (ou por um intervalo de tempo λT). Em suma, precisamos munir $\mathbb{V}$ com um tipo de "norma" — para atribuir "tamanho" aos vetores que caracterizam a separação entre eventos —, mas, ao mesmo tempo, discernindo "tamanhos" nas direções temporais de "tamanhos" nas direções espaciais.

Veremos que todas essas propriedades desejadas são providas por uma *forma simétrica, bilinear, não-degenerada* $\mathcal{G} : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ com *assinatura lorentziana* — que significa que a cada par de 4-vetores $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, $\mathcal{G}$ associa um número real $\mathcal{G}(\mathbf{u}, \mathbf{v})$ satisfazendo as seguintes propriedades:

(i) $\mathcal{G}(\mathbf{u}, \mathbf{v}) = \mathcal{G}(\mathbf{v}, \mathbf{u})$; (Simetria)

(ii) Sendo $\lambda \in \mathbb{R}$, $\mathcal{G}(\mathbf{u}, \mathbf{v} + \lambda\mathbf{w}) = \mathcal{G}(\mathbf{u}, \mathbf{v}) + \lambda\mathcal{G}(\mathbf{u}, \mathbf{w})$; (Linearidade)

(iii) Se $\mathcal{G}(\mathbf{u}, \mathbf{v}) = 0$ para todo $\mathbf{v} \in \mathbb{V}$, então $\mathbf{u} = \mathbf{0}$; (Não-degenerescência)

(iv) Existe um 4-vetor $\mathbf{u}_0 \in \mathbb{V}$ tal que $\mathcal{G}(\mathbf{u}_0, \mathbf{u}_0) < 0$. E todos os 4-vetores $\mathbf{e} \neq \mathbf{0}$ satisfazendo $\mathcal{G}(\mathbf{u}_0, \mathbf{e}) = 0$ também satisfazem $\mathcal{G}(\mathbf{e}, \mathbf{e}) > 0$. (Assinatura lorentziana)

As propriedades de (i) a (iii) são, na verdade, comuns a qualquer *produto interno* — inclusive o definido no espaço euclidiano (mais precisamente, na sua estrutura afim). A distinção aparece mesmo na propriedade (iv), mostrando que existe uma diferenciação entre as possíveis direções no espaço-tempo: aquelas dadas por 4-vetores $\mathbf{u}$ satisfazendo $\mathcal{G}(\mathbf{u}, \mathbf{u}) < 0$ e aquelas dadas por 4-vetores $\mathbf{u}$ satisfazendo $\mathcal{G}(\mathbf{u}, \mathbf{u}) > 0$.

- **Exercício:** Mostre que as propriedades (i) a (iv) implicam que:

  (a) Há infinitas direções ao longo das quais $\mathcal{G}$ é *negativo-definido* (ou seja, direções dadas por $\mathbf{u}$ com $\mathcal{G}(\mathbf{u}, \mathbf{u}) < 0$);

  (b) Há 4-vetores $\ell \neq \mathbf{0}$ tais que $\mathcal{G}(\ell, \ell) = 0$.

Sugestivamente, os 4-vetores $\mathbf{u}$ satisfazendo $\mathcal{G}(\mathbf{u}, \mathbf{u}) < 0$ são classificados como *tipo-tempo* — pois representarão as direções temporais —, enquanto aqueles satisfazendo $\mathcal{G}(\mathbf{u}, \mathbf{u}) > 0$ são classificados como *tipo-espaço* — por representarem as direções espaciais.

Embora $\mathcal{G}$ não seja um legítimo produto interno no sentido usual [por conta da propriedade (iv)], usaremo-lo como se fosse, definindo normas, projeções e ortogonalidade (com o devido cuidado). Assim, para um 4-vetor $\mathbf{u} \in \mathbb{V}$ qualquer, sua norma é definida por $\|\mathbf{u}\| := \sqrt{|\mathcal{G}(\mathbf{u}, \mathbf{u})|}$. Note que, ao contrário do que acontece com normas usuais, existem 4-vetores não-nulos cuja norma é zero (vide item (b) do Exercício anterior) — e esses 4-vetores desempenharão um papel de destaque em $\mathbb{M}$. Além disso, dado um 4-vetor $\mathbf{u}$ com $\mathcal{G}(\mathbf{u}, \mathbf{u}) \neq 0$, define-se o operador de *projeção na direção de* $\mathbf{u}$, $\mathcal{P}_{\mathbf{u}} : \mathbb{V} \to \mathbb{V}$, através de

$$\mathcal{P}_{\mathbf{u}}(\mathbf{v}) := \frac{\mathcal{G}(\mathbf{u}, \mathbf{v})}{\mathcal{G}(\mathbf{u}, \mathbf{u})}\,\mathbf{u}$$

- **Exercício:** Mostre que o operador definido acima de fato é uma projeção (ou seja, é *idempotente*, $\mathcal{P}_{\mathbf{u}}^2 = \mathcal{P}_{\mathbf{u}}$) e que, mais que isso, é uma *projeção ortogonal* (ou seja, é *auto-adjunto*, $\mathcal{G}(\mathbf{v}, \mathcal{P}_{\mathbf{u}}(\mathbf{w})) = \mathcal{G}(\mathcal{P}_{\mathbf{u}}(\mathbf{v}), \mathbf{w})$ para quaisquer $\mathbf{v}, \mathbf{w} \in \mathbb{V}$).

Por fim, define-se o *espaço ortogonal a* $\mathbf{u}$ por

$$\mathbb{V}^{\perp}(\mathbf{u}) := \{\mathbf{v} \in \mathbb{V};\ \mathcal{G}(\mathbf{u}, \mathbf{v}) = 0\}\,.$$

- **Exercício:** Dado um 4-vetor $\mathbf{u} \neq \mathbf{0}$ qualquer, mostre que $\mathbb{V}^{\perp}(\mathbf{u}) \subset \mathbb{V}$ é um subespaço vetorial de dimensão 3. Além disso, mostre que:

  (a) Se $\mathcal{G}(\mathbf{u}, \mathbf{u}) < 0$, então $\mathbf{u} \notin \mathbb{V}^{\perp}(\mathbf{u})$ — portanto, $\{\mathbf{u}\} \cup \mathbb{V}^{\perp}(\mathbf{u})$ gera todo $\mathbb{V}$ — e a norma em $\mathbb{V}^{\perp}(\mathbf{u})$ é positivo-definida;

  (b) Se $\mathcal{G}(\mathbf{u}, \mathbf{u}) = 0$, então $\mathbf{u} \in \mathbb{V}^{\perp}(\mathbf{u})$ — portanto, $\{\mathbf{u}\} \cup \mathbb{V}^{\perp}(\mathbf{u})$ *não* gera todo $\mathbb{V}$ — e a forma $\mathcal{G}$ restrita a $\mathbb{V}^{\perp}(\mathbf{u})$ é degenerada [ou seja, não satisfaz a propriedade (iii)] ;

  (c) Se $\mathcal{G}(\mathbf{u}, \mathbf{u}) > 0$, então $\mathbf{u} \notin \mathbb{V}^{\perp}(\mathbf{u})$ — portanto, $\{\mathbf{u}\} \cup \mathbb{V}^{\perp}(\mathbf{u})$ gera todo $\mathbb{V}$ — e a forma $\mathcal{G}$ restrita a $\mathbb{V}^{\perp}(\mathbf{u})$ satisfaz todas as propriedades (i)-(iv).

- **Exercício:** Dado um 4-vetor $\mathbf{u}$ com $\|\mathbf{u}\| \neq 0$, mostre que $\mathbf{v}^{\perp} := \mathbf{v} - \mathcal{P}_{\mathbf{u}}(\mathbf{v}) \in \mathbb{V}^{\perp}(\mathbf{u})$ para qualquer 4-vetor $\mathbf{v}$.

Como com qualquer espaço vetorial, é útil introduzirmos bases para expressar os 4-vetores em termos de componentes. Qualquer conjunto linearmente independente de quatro 4-vetores faria esse papel, mas a presença da forma $\mathcal{G}$ privilegia algumas dessas escolhas. Se estivéssemos lidando com um espaço vetorial munido de um legítimo produto interno, as *bases ortonormais* seriam essas escolhas privilegiadas. No nosso caso, esse conceito é substituído pelo de *base tetrada.* Uma base tetrada é um conjunto de quatro 4-vetores $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ satisfazendo a condição de ortonormalidade (adaptada)

$$\mathcal{G}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \eta_{\mu\nu}, \tag{1.1}$$

onde os índices μ, ν tomam valores em $\{0, 1, 2, 3\}$ e $\eta_{\mu\nu}$ são as entradas da matriz diagonal

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.2}$$

(subentendendo que a primeira linha e primeira coluna são indexadas por $\mu = 0$ e $\nu = 0$, respectivamente, e assim por diante). Ou seja, $\mathbf{e}_0$ é um 4-vetor normalizado tipo-tempo, enquanto os outros elementos formam um legítimo conjunto ortonormal de 4-vetores tipo-espaço que pertencem a $\mathbb{V}^{\perp}(\mathbf{e}_0)$; i.e., $\{\mathbf{e}_j\}_{j=1,2,3}$ é uma base ortonormal de $\mathbb{V}^{\perp}(\mathbf{e}_0)$.

- **Exercício:** Considere dois 4-vetores $\mathbf{u}$ e $\mathbf{v}$ quaisquer e sejam u^μ e v^μ, respectivamente, suas componentes na *mesma* base tetrada $\{\mathbf{e}_\mu\}_{\mu=0,1,2,3}$. Ou seja,

$$\mathbf{u} = u^0\mathbf{e}_0 + u^1\mathbf{e}_1 + u^2\mathbf{e}_2 + u^3\mathbf{e}_3 = \sum_{\mu=0}^{3} u^\mu \mathbf{e}_\mu,$$

$$\mathbf{v} = v^0\mathbf{e}_0 + v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + v^3\mathbf{e}_3 = \sum_{\mu=0}^{3} v^\mu \mathbf{e}_\mu,$$

onde cada $u^\mu, v^\mu \in \mathbb{R}$ $(\mu = 0,1,2,3)$ representa a componente dos vetores $\mathbf{u}$ e $\mathbf{v}$ na direção de $\mathbf{e}_\mu$ — e *não* números u e v elevados à potência μ. Mostre que

$$\begin{aligned} \mathcal{G}(\mathbf{u},\mathbf{v}) &= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu} u^\mu v^\nu = -u^0v^0 + u^1v^1 + u^2v^2 + u^3v^3 \\ &= [\mathbf{u}]^\top \eta\, [\mathbf{v}] = [\mathbf{v}]^\top \eta\, [\mathbf{u}]\,, \end{aligned}$$

onde $[\mathbf{u}]$ representa a matriz coluna das componentes de $\mathbf{u}$, o sobrescrito $\top$ indica a transposição da matriz e η é a matriz definida na Eq. (1.2).

Veremos, mais adiante, que (campos de) bases tetradas serão usadas(os) para caracterizar o que chamamos de "(famílias de) observadores".

1.4.1 Interlúdio: Espaço dual, tensores, convenção de Einstein e notação de índices abstratos

As estruturas definidas acima — de espaço afim munido de uma forma $\mathcal{G}$ com as propriedades listadas — são tudo que precisamos para caracterizar nosso espaço-tempo como sendo o de Minkowski $\mathbb{M}$. No entanto, elas trazem consigo outras estruturas que serão muito úteis e necessárias quando quisermos descrever fenômenos ou formular leis físicas sobre o espaço-tempo. Essas estruturas são normalmente definidas e exploradas num curso de *Álgebra Linear*, mas, por completeza, faremos uma breve exposição aqui das definições e resultados mais importantes para os propósitos desta disciplina.

Espaços duais

Comecemos pelo conceito de *espaço dual*. Dado um espaço vetorial $\mathbb{V}$ qualquer, de dimensão d (finita, por simplicidade), considere os *funcionais lineares* sobre $\mathbb{V}$; ou seja, as funções $\mathbf{U} : \mathbb{V} \to \mathbb{R}$ satisfazendo

$$\mathbf{U}(\mathbf{u} + \lambda\mathbf{v}) = \mathbf{U}(\mathbf{u}) + \lambda\mathbf{U}(\mathbf{v}), \text{ para quaisquer } \mathbf{u},\mathbf{v} \in \mathbb{V},\ \lambda \in \mathbb{R}.$$

Não é difícil ver que o conjunto de *todos* os funcionais lineares sobre $\mathbb{V}$ também forma um espaço vetorial — denominado *espaço dual a* $\mathbb{V}$ e denotado por $\mathbb{V}^*$ — que possui a mesma dimensão de $\mathbb{V}$.[6]

- **Exercício:** Seja $\{\mathbf{x}_\mu\}_{\mu=0,1,2,3}$ uma base *qualquer* (ou seja, não necessariamente uma tetrada) de $\mathbb{V}$. Mostre que os quatro funcionais lineares $\mathbf{X}^\mu \in \mathbb{V}^*$, $\mu = 0, 1, 2, 3$, definidos por $\mathbf{X}^\mu(\mathbf{x}_\nu) := \delta^\mu_\nu$, formam uma base de $\mathbb{V}^*$ — dita ser a *base dual a* $\{\mathbf{x}_\mu\}_{\mu=0,1,2,3}$.

Assim como convencionamos chamar os elementos de $\mathbb{V}$ de 4-vetores, chamaremos os elementos de $\mathbb{V}^*$ de *covetores*, apenas para diferenciá-los dos 4-vetores, e é frequente nos referirmos a $\mathbb{V}^*_p = \mathbb{V}^*$ como o *espaço cotangente* a $\mathbb{M}$ em p.

- **Exercício:** Seja $\{\mathbf{x}_\mu\}_{\mu=0,1,2,3}$ uma base qualquer de $\mathbb{V}$ e $\{\mathbf{X}^\mu\}_{\mu=0,1,2,3}$ a respectiva base dual de $\mathbb{V}^*$. Seja $\mathbf{U} \in \mathbb{V}^*$ um covetor qualquer, expresso em termos da base dual como

$$\mathbf{U} = \sum_{\mu=0}^{3} U_\mu \mathbf{X}^\mu$$

(ou seja, U_μ são as componentes do covetor $\mathbf{U}$), e seja $\mathbf{u} \in \mathbb{V}$ um 4-vetor qualquer, com componentes u^μ na base dada:

$$\mathbf{u} = \sum_{\mu=0}^{3} u^\mu \mathbf{x}_\mu .$$

(a) Mostre que as componentes de $\mathbf{u}$ (u^μ) e as componentes de $\mathbf{U}$ (U_μ) são dadas por

$$u^\mu = \mathbf{X}^\mu(\mathbf{u}) \ , \quad U_\mu = \mathbf{U}(\mathbf{x}_\mu);$$

(b) Mostre que

$$\mathbf{U}(\mathbf{u}) = \sum_{\mu=0}^{3} U_\mu u^\mu = [\mathbf{U}]^\top [\mathbf{u}] = [\mathbf{u}]^\top [\mathbf{U}] \, .$$

Naturalmente, do mesmo jeito que podemos falar do espaço dual a $\mathbb{V}$, podemos também definir o espaço dual a $\mathbb{V}^*$: $(\mathbb{V}^*)^*$. Esse espaço é constituído de funcionais lineares sobre $\mathbb{V}^*$; ou seja, por *todas* as funções $\underline{\mathbf{u}} : \mathbb{V}^* \to \mathbb{R}$ satisfazendo $\underline{\mathbf{u}}(\mathbf{U}+\lambda\mathbf{V}) = \underline{\mathbf{u}}(\mathbf{U})+\lambda\underline{\mathbf{u}}(\mathbf{V})$, para quaisquer $\mathbf{U}, \mathbf{V} \in \mathbb{V}^*$, $\lambda \in \mathbb{R}$. Podemos ir além e falar do dual a $(\mathbb{V}^*)^*$ — $[(\mathbb{V}^*)^*]^*$ —, do dual a $[(\mathbb{V}^*)^*]^*$ — $\left\{[(\mathbb{V}^*)^*]^*\right\}^*$ —, e assim por diante, construindo uma cadeia infinita de

[6]Subentendendo que a combinação linear de dois funcionais quaisquer $\mathbf{U}, \mathbf{V} \in \mathbb{V}^*$ é definida por $(\mathbf{U} + \lambda\mathbf{V})(\mathbf{u}) := \mathbf{U}(\mathbf{u}) + \lambda\mathbf{V}(\mathbf{u})$, para *todo* $\mathbf{u} \in \mathbb{V}$ e $\lambda \in \mathbb{R}$.

espaços vetoriais a partir de $\mathbb{V}$. E pela mesma razão que $\mathbb{V}$ e $\mathbb{V}^*$ possuem a mesma dimensão d, todos os espaços dessa cadeia infinita também têm dimensão d. Logo, todos eles são *isomorfos* — ou seja, podem ser identificados entre eles, elemento a elemento, de modo a preservar a estrutura de espaço linear de cada um deles. Na verdade, há *infinitas* maneiras *diferentes* de se fazer essas identificações e a questão que se coloca é: haveria, dentre as infinitas maneiras de se identificar dois espaços vetorias dessa cadeia, *uma* maneira privilegiada?

Apenas com a estrutura de espaço vetorial dada em $\mathbb{V}$, a resposta é "*não*" para os "duais imediatos" — $\mathbb{V}$ e $\mathbb{V}^*$, $\mathbb{V}^*$ e $(\mathbb{V}^*)^*$, $(\mathbb{V}^*)^*$ e $[(\mathbb{V}^*)^*]^*$, e assim por diante.[7] No entanto, há *sim* um isomorfismo "natural" — denominado de *isomorfismo canônico* — entre os "duplo-duais": $\mathbb{V}$ e $(\mathbb{V}^*)^*$, $\mathbb{V}^*$ e $[(\mathbb{V}^*)^*]^*$, e assim por diante. Explicitamente, esse isomorfismo canônico entre $\mathbb{V}$ e $(\mathbb{V}^*)^*$ é dado pela identificação

$$(\mathbb{V}^*)^* \ni \underline{\mathbf{u}} \leftrightarrow \mathbf{u} \in \mathbb{V} \iff \underline{\mathbf{u}}(\mathbf{U}) = \mathbf{U}(\mathbf{u}) \text{ para } \textit{todo } \mathbf{U} \in \mathbb{V}^*.$$

- **Exercício:** Mostre que $\underline{\mathbf{u}} \in (\mathbb{V}^*)^*$ é canonicamente identificado com $\mathbf{u} \in \mathbb{V}$ se, e somente se, suas *componentes* $\underline{u}^\mu$ e u^μ, nas respectivas bases (duplo-)duais, são iguais: $\underline{u}^\mu = u^\mu$, $\mu = 0, 1, 2, 3$. (Sugestão: combine a definição do isomorfismo canônico dada acima com as últimas igualdades do Exercício anterior.)

Com essa identificação, a cadeia infinita de duais se reduz a apenas dois elementos: $\mathbb{V}$ e $\mathbb{V}^*$. Além disso, vemos que há uma simetria entre eles: $\mathbb{V}^*$ é o dual de $\mathbb{V}$, assim como $\mathbb{V}$ pode ser visto como o dual de $\mathbb{V}^*$ [pois $\mathbb{V} \equiv (\mathbb{V}^*)^*$]. Esses dois espaços servirão de base para a construção a seguir.

Tensores

O conceito de *tensor* surge, em física, da necessidade de se definir grandezas que, além de elas próprias poderem ser somadas entre elas e multiplicadas por números reais — o que uma simples estrutura de espaço vetorial já provê —, também precisam ser multiplicadas entre elas ou por elementos de outro espaço vetorial. Um exemplo simples é o *momentum* de uma partícula. Evidentemente, podemos somar os momenta de partículas que constituem um sistema para calcular o momentum total do sistema (e, então, ver se este está sujeito a alguma lei de conservação, por exemplo). Os próprios momenta são definidos pelo produto de números reais — as massas das partículas numa dada unidade — por vetores — as velocidades. Porém, sabemos, de nossa experiência com mecânica newtoniama, que também é útil se definir o *momentum angular* de uma partícula, através de um produto

[7] No entanto, no nosso caso, há uma estrutura adicional: a forma bilinear $\mathcal{G}$ — que mais adiante denominaremos de *métrica*. Essa estrutura permite, sim, estabelecer *um* isomorfismo preferencial entre os duais diretos. Vide Exercício ②.

"apropriado" entre momentum e posição, ambos sendo grandezas vetoriais — que vivem em "cópias" de $\mathbb{V}$ com unidades diferentes. Assim, precisamos dar sentido a, por exemplo, produtos de elementos de $\mathbb{V}$.

Antes de prosseguirmos, vale a pena deixar claro o espírito por trás da definição que buscaremos. Note que ao falarmos de elementos de $\mathbb{V}$ e de $\mathbb{V}^*$, fizemos questão de apresentá-los como objetos cuja definição e existência *independem* do conceito de "componentes". É claro que, *dada* uma base de $\mathbb{V}$, existe uma associação um a um (mais precisamente, um isomorfismo) entre $\mathbb{V}$ e $\mathbb{R}^4$: $\mathbf{u} \leftrightarrow (u^\mu)_{\mu=0,1,2,3}$. O mesmo vale para $\mathbb{V}^*$: $\mathbf{U} \leftrightarrow (U_\mu)_{\mu=0,1,2,3}$. Mas desde o começo, ainda quando falávamos apenas do conceito de espaço-tempo, quisemos deixar clara a distinção entre o "objeto" em si ($p \in \mathbb{M}$, $\mathbf{u} \in \mathbb{V}$, $\mathbf{U} \in \mathbb{V}^*$) e sua *representação* (respectivamente, $(t, \mathbf{P}) \in \mathbb{R} \times \mathbb{E}^3$, $(u^\mu) \in \mathbb{R}^4$, $(U_\mu) \in \mathbb{R}^4$): o "objeto" — que descreverá alguma propriedade do sistema físico de interesse — tem existência absoluta, enquanto que sua representação depende de escolhas (como sistemas de coordenadas, bases, etc.) que são "externas" ao sistema físico de interesse — e que, portanto, deveriam ser desnecessárias para se *definir* as grandezas que caracterizam ou descrevem o sistema. Em resumo, queremos, por exemplo, definir o produto de elementos de $\mathbb{V}$ sem termos que nos referir às componentes desses elementos. Mais genericamente, queremos definir, em cada ponto do espaço-tempo, grandezas que são combinações de produtos de outras grandezas, mas sem a necessidade de se adotar nenhum sistema de coordenadas ou base para isso — para garantir que o "objeto" que codifica essa grandeza tenha existência absoluta e não seja refém de, ou fique preso a, escolhas arbitrárias.

Tensores são os objetos matemáticos que implementam o que queremos. A idéia é generalizar o procedimento usado para se definir os covetores e definir um *tensor de posto* (m, n) (com $m, n \in \mathbb{N}$) como sendo um *funcional multilinear* atuando sobre m cópias de $\mathbb{V}^*$ e n cópias de $\mathbb{V}$; ou seja, uma função $T : \mathbb{V}^* \times \cdots \times \mathbb{V}^* \times \mathbb{V} \times \cdots \times \mathbb{V} \to \mathbb{R}$ que é linear em *cada um* de seus $m + n$ argumentos. Vale notar que, com essa definição, um vetor pode ser visto como um tensor de posto $(1, 0)$, enquanto que um covetor é um tensor de posto $(0, 1)$. A própria forma simétrica bilinear $\mathcal{G}$ definida anteriormente é um exemplo de tensor de posto $(0, 2)$. Por definição (e consistência de notação), tensores de posto $(0, 0)$ são os *escalares*: grandezas que já assumem valores (reais) sem necessidade de se atuar em nada.

O conjunto de todos os tensores de posto (m, n) será denotado por $\mathcal{T}(m, n)$.[8] Esse conjunto possui estrutura natural de espaço vetorial, pois pode-se facilmente definir combinações lineares de funcionais multilineares, assim como fizemos anteriormente para funcionais lineares (vide nota de rodapé 6). Mas, além disso, dados dois tensores quaisquer, $T \in \mathcal{T}(m, n)$ e

[8]Note, então, que $\mathcal{T}(0, 0) = \mathbb{R}$, $\mathcal{T}(1, 0) = \mathbb{V}$ e $\mathcal{T}(0, 1) = \mathbb{V}^*$.

$Q \in \mathcal{T}(r,s)$, podemos *definir* o tensor *produto*, $T \otimes Q \in \mathcal{T}(m+r, n+s)$ por

$$\begin{aligned}(T \otimes Q)(\mathbf{U}^1, ..., \mathbf{U}^{m+r}, \mathbf{u}_1, ..., \mathbf{u}_{n+s}) := {} & T(\mathbf{U}^1, ..., \mathbf{U}^m, \mathbf{u}_1, ..., \mathbf{u}_n) \\ & \times Q(\mathbf{U}^{m+1}, ..., \mathbf{U}^{m+r}, \mathbf{u}_{n+1}, ..., \mathbf{u}_{n+s}),\end{aligned}$$

para qualquer $(\mathbf{U}^1, ..., \mathbf{U}^{m+r}, \mathbf{u}_1, ..., \mathbf{u}_{n+s}) \in \mathbb{V}^* \times ... \times \mathbb{V}^* \times \mathbb{V} \times ... \times \mathbb{V}$.

- **Exercício:** Mostre que $\mathcal{T}(m,n)$ é um espaço vetorial de dimensão d^{m+n}. [Sugestão: Mostre que se $\{\mathbf{x}_\mu\}_{\mu=0,...,d-1}$ é uma base qualquer de $\mathbb{V}$ e $\{\mathbf{X}^\mu\}_{\mu=0,...,d-1}$ a respectiva base dual de $\mathbb{V}^*$, então os tensores

  $$\{\mathbf{x}_{\mu_1} \otimes ... \otimes \mathbf{x}_{\mu_m} \otimes \mathbf{X}^{\nu_1} \otimes ... \otimes \mathbf{X}^{\nu_n}\}_{\mu_i, \nu_j = 0,...,d-1}$$

  formam uma base de $\mathcal{T}(m,n)$ — denominada *base induzida*.]

- **Exercício:** Seja $T \in \mathcal{T}(m,n)$ um tensor qualquer e considere sua expansão em termos da base introduzida no exercício anterior:

  $$T = \sum_{\mu_1=0}^{d-1} ... \sum_{\mu_m=0}^{d-1} \sum_{\nu_1=0}^{d-1} ... \sum_{\nu_n=0}^{d-1} T^{\mu_1 ... \mu_m}_{\nu_1 ... \nu_n} \, \mathbf{x}_{\mu_1} \otimes ... \otimes \mathbf{x}_{\mu_m} \otimes \mathbf{X}^{\nu_1} \otimes ... \otimes \mathbf{X}^{\nu_n}.$$

  Mostre que suas componentes são dadas por

  $$T^{\mu_1 ... \mu_m}_{\nu_1 ... \nu_n} = T(\mathbf{X}^{\mu_1}, ..., \mathbf{X}^{\mu_m}, \mathbf{x}_{\nu_1}, ..., \mathbf{x}_{\nu_n}).$$

- **Exercício:** Sejam $\mathbf{u}$ e $\mathbf{v}$ dois 4-vetores quaisquer, cujas componentes numa base arbitrária $\{\mathbf{x}_\mu\}_{\mu=0,1,2,3}$ são $(u^\mu)_{\mu=0,1,2,3}$ e $(v^\mu)_{\mu=0,1,2,3}$, respectivamente.

  (a) Mostre que as componentes de $\mathbf{u} \otimes \mathbf{v}$ na base induzida de $\mathcal{T}(2,0)$ são dadas por $(u^\mu v^\nu)_{\mu,\nu=0,1,2,3}$; ou seja,

  $$\mathbf{u} \otimes \mathbf{v} = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} u^\mu v^\nu \, \mathbf{x}_\mu \otimes \mathbf{x}_\nu;$$

  (b) Com o resultado do item anterior, mostre, nesse caso simples, que o produto de tensores *não* é, em geral, comutativo;

  (c) Repita o item (a) para $\mathbf{u} \otimes \mathbf{V}$, onde $\mathbf{V} \in \mathbb{V}^*$, e mostre que, nesse caso, $\mathbf{u} \otimes \mathbf{V} = \mathbf{V} \otimes \mathbf{u}$.

- **Exercício:** Calcule as componentes do tensor $\mathcal{G}$, definido anteriormente, na base de $\mathcal{T}(0,2)$ induzida pela base tetrada $\{\mathbf{e}_\mu\}_{\mu=0,1,2,3}$. Em seguida, escreva $\mathcal{G}$ *explicitamente* em termos dessa base induzida. (Use $\{\mathbf{E}^\mu\}_{\mu=0,1,2,3}$ para a base dual à base tetrada.)

- **Exercício:** Sendo $T \in \mathcal{T}(m,n)$ um tensor qualquer, mostre que

$$T\left(\mathbf{U},...,\mathbf{V},\mathbf{u},...,\mathbf{v}\right) = \sum_{\mu_1=0}^{d-1} ... \sum_{\mu_m=0}^{d-1} \sum_{\nu_1=0}^{d-1} ... \sum_{\nu_n=0}^{d-1} T^{\mu_1...\mu_m}_{\nu_1...\nu_n} U_{\mu_1}...V_{\mu_m} u^{\nu_1}...v^{\nu_n},$$

onde se subentende a quantidade de covetores e vetores no argumento de T.

Outra operação útil de se definir para tensores é a chamada *contração*. Seja $T \in \mathcal{T}(m,n)$ um tensor qualquer com $m, n \geq 1$. Então, a contração do i-ésimo argumento covetorial com o j-ésimo argumento vetorial ($1 \leq i \leq m$, $1 \leq j \leq n$) é definida pelo mapeamento $\mathcal{C}_{(i,j)} : \mathcal{T}(m,n) \to \mathcal{T}(m-1,n-1)$ dado por

$$\mathcal{C}_{(i,j)}\left(T\right)\left(...\right) := \sum_{\mu=0}^{d-1} T(..., \mathbf{X}^{\mu}, ..., \mathbf{x}_{\mu}, ...),$$

onde *cada* elemento da base $\{\mathbf{x}_\mu\}_{\mu=0,...,d-1}$ é inserido no j-ésimo argumento vetorial de T, ao mesmo tempo que o *correspondente* elemento da base dual $\{\mathbf{X}^\mu\}_{\mu=0,...,d-1}$ é inserido no i-ésimo argumento covetorial e então soma-se sobre todos os valores de $\mu = 0, ..., d-1$. (Subentende-se a maneira como os outros $m+n-2$ argumentos aparecem na expressão acima.)

- **Exercício:** Sendo $\mathbf{u} \in \mathcal{T}(1,0)$ e $\mathbf{U} \in \mathcal{T}(0,1)$, calcule $\mathcal{C}_{(1,1)}\left(\mathbf{u} \otimes \mathbf{U}\right)$ e mostre que

$$\mathcal{C}_{(1,1)}\left(\mathbf{u} \otimes \mathbf{U}\right) = \mathbf{U}(\mathbf{u}) = \sum_{\mu=0}^{d-1} U_\mu u^\mu.$$

É importante frisar que, embora tenha sido feita menção a uma base de $\mathbb{V}$ e à respectiva base dual de $\mathbb{V}^*$ na definição acima de contração, o resultado da contração é um legítimo tensor de posto $(m-1, n-1)$; logo, independente de base. O Exercício abaixo ilustra bem, num caso particular, como isso ocorre.

- **Exercício:** Considere um tensor $T \in \mathcal{T}(2,3)$ com componentes $T^{\mu\nu}_{\alpha\beta\gamma}$ numa base arbitrária; ou seja,

$$T = \sum_{\mu=0}^{d-1} \sum_{\nu=0}^{d-1} \sum_{\alpha=0}^{d-1} \sum_{\beta=0}^{d-1} \sum_{\gamma=0}^{d-1} T^{\mu\nu}_{\alpha\beta\gamma} \, \mathbf{x}_\mu \otimes \mathbf{x}_\nu \otimes \mathbf{X}^\alpha \otimes \mathbf{X}^\beta \otimes \mathbf{X}^\gamma.$$

Mostre que

$$\mathcal{C}_{(1,3)}(T) = \sum_{\nu=0}^{d-1} \sum_{\alpha=0}^{d-1} \sum_{\beta=0}^{d-1} \left(\sum_{\lambda=0}^{d-1} T^{\lambda\nu}_{\alpha\beta\lambda} \right) \mathbf{x}_\nu \otimes \mathbf{X}^\alpha \otimes \mathbf{X}^\beta;$$

ou seja, que $\mathcal{C}_{(1,3)}(T)$ é um legítimo tensor de posto $(1,2)$ cujas componentes são dadas por $\sum_{\lambda=0}^{d-1} T^{\lambda\nu}_{\alpha\beta\lambda}$.

Convenção de Einstein e notação de índices abstratos

Os Exercícios anteriores ilustram bem que somatórias são (muito!) frequentes quando trabalhamos com tensores, em particular quando lidamos com expressões dadas em termos de suas componentes. Note, no entanto, que da maneira como introduzimos a notação de bases e componentes, sempre que um índice é somado sobre todos seus possíveis valores (de 0 a $d-1$), esse *mesmo* índice aparece *exatamente duas vezes* numa mesma parcela: uma de forma sobrescrita, outra subescrita. Isso permite que identifiquemos os índices a serem somados *sem* a necessidade do símbolo de somatória, o que, por sua vez, possibilita que adotemos a convenção — denominada *convenção de soma de Einstein*[9] — na qual os símbolos de somatória são *subentendidos e omitidos* sempre que o *mesmo índice estiver repetido em cima e embaixo* numa mesma parcela de uma expressão. Exemplificando o uso dessa convenção, note como ficam algumas das expressões que encontramos até agora:

$$\mathbf{u} = u^\mu \mathbf{x}_\mu, \quad \mathbf{U} = U_\mu \mathbf{X}^\mu, \quad \mathbf{U}(\mathbf{u}) = U_\mu u^\mu,$$

$$T = T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} \, \mathbf{x}_{\mu_1} \otimes \dots \otimes \mathbf{x}_{\mu_m} \otimes \mathbf{X}^{\nu_1} \otimes \dots \otimes \mathbf{X}^{\nu_n},$$

$$T\left(\mathbf{U}, \dots, \mathbf{V}, \mathbf{u}, \dots, \mathbf{v}\right) = T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} \, U_{\mu_1} \dots V_{\mu_m} u^{\nu_1} \dots v^{\nu_n},$$

$$\mathcal{C}_{(1,3)} : T^{\mu\nu}_{\alpha\beta\gamma} \, \mathbf{x}_\mu \otimes \mathbf{x}_\nu \otimes \mathbf{X}^\alpha \otimes \mathbf{X}^\beta \otimes \mathbf{X}^\gamma \mapsto T^{\lambda\nu}_{\alpha\beta\lambda} \, \mathbf{x}_\nu \otimes \mathbf{X}^\alpha \otimes \mathbf{X}^\beta.$$

Além disso, note que o nome usado para o índice repetido é irrelevante, sendo um índice “mudo” (como acontece com variáveis de integração em Cálculo). O único cuidado que se deve tomar é o de que o mesmo índice não pode aparecer mais de *duas* vezes (uma embaixo, uma em cima) numa mesma parcela.

A convenção de soma de Einstein simplifica bastante expressões tensoriais que teremos que manipular ao longo do curso. No entanto, ainda há um fato inconveniente com nossa notação de tensores. Até agora, utilizamos letras romanas minúsculas em negrito para denotar vetores e letras romanas maiúsculas em negrito para denotar covetores. Mas agora que temos infinitos tipos diferentes de tensores, dos quais vetores e covetores são apenas dois exemplos, temos que pensar numa notação mais versátil, de modo que ao ver uma quantidade tensorial numa expressão, saibamos qual seu tipo (i.e., posto) só pela maneira como ela foi denotada.

Felizmente, uma notação assim versátil (e escandalosamente simples!) já foi pensada: a notação de *índices abstratos*.[10] A idéia é explorar o fato que, com nossa convenção para a posição de índices nas componentes e bases de tensores, é imediato identificar o posto de um tensor T se dissermos que suas

[9] Jocosamente, Einstein teria dito que essa foi sua maior contribuição à matemática.

[10] Essa notação foi introduzida pelo físico e matemático Roger Penrose, na década de 1960.

componentes são dadas, por exemplo, por $T^{\mu}_{\alpha\beta\gamma}$; nesse caso, evidentemente $T \in \mathcal{T}(1,3)$. Ou seja, a posição e a quantidade de índices (diferentes) nas *componentes* — chamados de *índices concretos*, por assumirem valores, de 0 a $d-1$ — denuncia o posto do tensor. Por outro lado, lembre-se que, dentro do espírito de se distinguir o objeto em si (absoluto) de sua representação (dependente de escolhas arbitrárias), não queremos identificar o tensor T com suas componentes. A saída para esse impasse é bem simples: ao invés de se denotar o tensor apenas por T, por exemplo, incluiremos, como parte da notação, *índices abstratos*, ou seja, índices que *não* devem assumir nenhum valor, apenas fazem parte do "nome" dado ao tensor e demarcam as posições em que os índices concretos aparecem nas componentes do tensor. No exemplo dado, ao invés de T, denotaríamos o tensor por T^{a}_{bcd}, o que imediatamente deixaria claro seu posto. É claro que temos que arrumar uma convenção para diferenciar, à vista, índices abstratos de índices concretos — para sabermos se estamos nos referindo ao tensor ou às suas componentes. Para isso, adotaremos as oito primeiras letras minúsculas do alfabeto romano $(a, b, ..., h)$ para denotar índices abstratos. Índices concretos que podem assumir valores de 0 a $d-1$ serão representados por letras gregas minúsculas $(\alpha, \beta, ...)$ — como já vínhamos fazendo — e índices concretos que assumem apenas os valores de 1 a $d-1$ serão denotados pelas letras minúsculas romanas do final do alfabeto $(i, j, ...)$. Se precisarmos de mais índices do que os providos por essa "regra", podemos anexar índices numéricos (ou concretos) a esses índices (e.g., $a_1, \mu_m, i_n, ...$).

À primeira vista, a notação de índices abstratos pode parecer um preciosismo em nível patológico. Porém, a clareza que essa notação nos trará, principalmente quando nos depararmos com objetos descritos por componentes que possuem índices mas *não* são tensores[11], mais do que justifica sua adoção — na minha opinião. Além disso, as operações tensoriais introduzidas anteriormente são mais diretamente expressas usando essa notação. Por exemplo, dados tensores $T^{a_1...a_m}_{b_1...b_n}$ e $Q^{a_1...a_r}_{b_1...b_s}$, o produto é simplesmente denotado por $T^{a_1...a_m}_{b_1...b_n} Q^{c_1...c_r}_{d_1...d_s}$, enquanto que a operação de contração fica particularmente simples:

$$\mathcal{C}_{(i,j)} : T^{a_1...a_m}_{b_1...b_n} \mapsto T^{a_1...c...a_m}_{b_1...c...b_n},$$

onde o índice abstrato "mudo" c foi colocado nos lugares de a_i e b_j. (Sendo c um índice abstrato, a convenção de Einstein não se aplica concretamente a ele. Mas o fato de ele aparecer repetido indica que, quando se calcular as *componentes* de $T^{a_1...c...a_m}_{b_1...c...b_n}$, deve-se aplicar a convenção de Einstein ao índice concreto que entrar no lugar de c.) Note que índices abstratos "mudos" não devem ser contados para se inferir o posto do tensor. Assim, por exemplo, o produto de tensores T^{a}_{bc} e Q^{ab}_{c}, seguido das contrações $\mathcal{C}_{(1,3)}$ e $\mathcal{C}_{(2,1)}$ (nessa

[11]Não, nem todo objeto com índice é um tensor.

ordem) tem como resultado final o tensor $S^e_c := T^a_{bc}Q^{eb}_a$ cujo posto é $(1,1)$ (certifique-se de que você entende isso).

Com essa notação, os elementos da base adquirem formato peculiar. Uma base arbitrária de $\mathbb{V}$ será denotada por $\{\mathbf{x}^a_\mu\}$, enquanto que uma base tetrada será dada por $\{\mathbf{e}^a_\mu\}$. O índice concreto μ, aqui, continua indexando elementos diferentes da base, enquanto que o índice abstrato aparece apenas para deixar claro o caráter vetorial de cada elemento. Analogamente, as respectivas bases duais serão denotadas por $\{\mathbf{X}^\mu_a\}$ e $\{\mathbf{E}^\mu_a\}$, onde a condição de dualidade, agora, é escrita como $\mathbf{X}^\mu_a\mathbf{x}^a_\nu = \delta^\mu_\nu = \mathbf{E}^\mu_a\mathbf{e}^a_\nu$. Manteremos o negrito para nos referirmos a elementos de bases para deixar claro que o índice concreto, aqui, *não* se refere a componentes, mas a elementos diferentes. Assim, um tensor é expresso em termos de suas componentes como

$$T^{a_1...a_m}_{b_1...b_n} = T^{\mu_1...\mu_m}_{\nu_1...\nu_n}\,\mathbf{x}^{a_1}_{\mu_1}...\mathbf{x}^{a_m}_{\mu_m}\mathbf{X}^{\nu_1}_{b_1}...\mathbf{X}^{\nu_n}_{b_n},$$

enquanto que suas componentes são dadas por

$$T^{\mu_1...\mu_m}_{\nu_1...\nu_n} = T^{a_1...a_m}_{b_1...b_n}\mathbf{X}^{\mu_1}_{a_1}...\mathbf{X}^{\mu_m}_{a_m}\mathbf{x}^{b_1}_{\nu_1}...\mathbf{x}^{b_n}_{\nu_n}.$$

O(a) leitor(a) deve se certificar de que entende a sutil diferença entre as duas expressões acima: a primeira representa uma combinação linear de elementos da base de $\mathcal{T}(m,n)$ — portanto, tendo como resultado um elemento desse mesmo espaço —, enquanto a segunda representa a aplicação (equivalentemente, a contração) de um elemento de $\mathcal{T}(m,n)$ em (com) m elementos da base de $\mathbb{V}^*$ e n elementos da base de $\mathbb{V}$ — tendo como resultado números reais. Note, ainda, a perfeita correspondência entre os índices "livres" (i.e., "não-mudos") de ambos os lados de cada expressão. Isso é uma característica obrigatória de *qualquer* legítima equação tensorial.

Munidos desse arsenal de estruturas, convenções e notações, podemos (finalmente!) voltar para a física.

1.5 Métrica e Intervalo invariante

Recapitulando, definimos nosso espaço-tempo de Minkowski $\mathbb{M}$ como sendo um espaço afim de dimensão $d=4$, munido de uma forma bilinear simétrica, não-degenerada, com assinatura lorentziana, $\mathcal{G}$. Na linguagem tensorial, vimos que $\mathcal{G}$ nada mais é do que um tensor de posto $(0,2)$ (com certas propriedades), que a partir de agora denotaremos simplesmente por g_{ab}, de acordo com a notação de índices abstratos introduzida. Traduzindo as propriedades de $\mathcal{G}$ para essa notação, temos que:

(i) $g_{ab} = g_{ba}$; (Simetria)

(ii) $g_{ab}u^b = 0$ se, se somente se, $u^a = 0$; (Não-degenerescência)

(iii) Existe u_0^a tal que $g_{ab}u_0^a u_0^b < 0$. E todo $e^a \neq 0$ satisfazendo $g_{ab}u_0^a e^b = 0$ também satisfaz $g_{ab}e^a e^b > 0$. (Assinatura lorentziana)

Note que a definição de base tetrada $\{\mathbf{e}_\mu^a\}$ passa a ser expressa como $g_{ab}\mathbf{e}_\mu^a \mathbf{e}_\nu^b = \eta_{\mu\nu}$, onde $\eta_{\mu\nu}$ são as entradas da matriz dada na Eq. (1.2). As componentes de g_{ab} serão denotadas por $\eta_{\mu\nu}$ *apenas* quando expressas numa base tetrada. Numa base qualquer $\{\mathbf{x}_\mu^a\}$ suas componentes serão comumente denotadas por $g_{\mu\nu}$ $(= g_{ab}\mathbf{x}_\mu^a \mathbf{x}_\nu^b)$.

A estrutura afim de $\mathbb{M}$ permite que g_{ab} induza uma "medida" entre pares de eventos. Cada par $p, q \in \mathbb{M}$ determina um 4-vetor separação $s^a := \overrightarrow{pq}$ que, por sua vez, tem uma norma devida a g_{ab}. Assim, define-se o *intervalo invariante* $\mathcal{I} : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ entre pares de evetos p, q como sendo a quantidade

$$\mathcal{I}(p, q) := \mathcal{G}(\overrightarrow{pq}, \overrightarrow{pq}) \equiv g_{ab}s^a s^b.$$

O adjetivo "invariante" se refere ao fato que o valor de $\mathcal{I}(p, q)$ depende apenas dos eventos escolhidos (que são entes absolutos), não da maneira como escolhemos representá-los em termos, por exemplo, de coordenadas (que podem ser arbitrárias). Assim, $\mathcal{I}$ é uma propriedade geométrica intrínseca de $\mathbb{M}$, análoga ao conceito de distância (ao quadrado) no espaço euclidiano — com a diferença importante que $\mathcal{I}$ *não* é positivo-definido. Por essa razão, g_{ab} é chamada de *métrica* do espaço-tempo.

Um dos papéis de g_{ab}, como dissemos quando a introduzimos, é o de distinguir direções temporais de espaciais em $\mathbb{V}_p = \mathbb{V}$, em cada evento $p \in \mathbb{M}$. Além disso, também vimos que g_{ab} provê uma noção de "ortogonalidade" no espaço-tempo, que, em particular, a cada direção temporal — dada, digamos, por u^a —, determina as direções que são "puramente espaciais" em relação a ela — aquelas dadas por elementos de $\mathbb{V}^\perp(u^a)$. Assim, para um observador com linha-de-mundo passando por $p \in \mathbb{M}$, na direção de u^a, sua *seção espacial* instantânea em p é definida por

$$\Sigma(p, u^a) := \left\{ q \in \mathbb{M};\ \ \overrightarrow{pq} \in \mathbb{V}^\perp(u^a) \right\}.$$

(Vide Exercício ⑥ para uma justificativa física para essa interpretação de $\Sigma(p, u^a)$.) Lembrando que g_{ab} é positivo-definido em $\mathbb{V}^\perp(u^a)$ e notando que $\Sigma(p, u^a)$ herda, de $\mathbb{M}$, uma estrutura de (sub)espaço afim, tem-se que $\Sigma(p, u^a)$ é um legítimo espaço euclidiano $\mathbb{E}^3$. Logo, de acordo com o observador em questão, um evento $q \in \Sigma(p, u^a)$ será simultâneo a p e a distância espacial entre eles será dada por $D(p, q) = \|\overrightarrow{pq}\| = \sqrt{g_{ab}s^a s^b} = \sqrt{\mathcal{I}(p, q)}$, onde $s^a = \overrightarrow{pq}$. Inversamente, quaisquer que sejam os eventos $p, q \in \mathbb{M}$ separados na direção espacial (ou seja, com $s^a = \overrightarrow{pq}$ tipo-espaço; $g_{ab}s^a s^b > 0$), sempre podemos encontrar um 4-vetor u^a tipo-tempo tal que $g_{ab}u^a s^b = 0$. Em outras palavras, se a separação entre p e q for tipo-espaço, sempre há um observador para quem esses eventos são simultâneos e espacialmente separados pela distância $D(p, q) = \sqrt{\mathcal{I}(p, q)}$. Isso nos fornece uma interpretação para o intervalo invariante nesse caso: *se* $\mathcal{I}(p,q) > 0$*, então* $\sqrt{\mathcal{I}(p,q)}$ *é a distância espacial entre os eventos* p *e* q *de acordo com um observador para quem eles são simultâneos.*

Suponha, agora, que a separação entre p e q seja numa direção temporal; ou seja, $\mathcal{I}(p,q) = g_{ab}s^a s^b < 0$. Nesse caso, existe um observador inercial cuja linha-de-mundo (que é uma reta no espaço-tempo) passa por esses dois eventos. Para esse observador particular, ambos os eventos se localizam em sua própria posição espacial, de modo que a distância entre eles é nula. Logo, $\mathcal{I}(p,q)$ depende apenas do intervalo de tempo que esse observador mede entre os eventos p e q; denotaremos esse intervalo de tempo por $\Delta\tau(p,q)$ (para diferenciar esse intervalo de tempo, medido por esse observador específico, de intervalos de tempo medidos por outros observadores). Pelas propriedades que queríamos para a dependência de Δt com a separação dos eventos — linearidade com o vetor separação; vide discussão na Seção 1.4 —, temos que $\mathcal{I}(p,q) = -\kappa^2[\Delta\tau(p,q)]^2$, onde κ^2 é uma constante de proporcionalidade a ser determinada. Logo: *se* $\mathcal{I}(p,q) < 0$*, então* $\kappa^{-1}\sqrt{-\mathcal{I}(p,q)}$ *é o intervalo de tempo entre os eventos* p *e* q *de acordo com o observador inercial para quem eles acontecem no mesmo ponto do espaço.*

O valor de κ é fixado usando-se o ingrediente que falta: linhas-de-mundo da luz. Vimos que se $\mathcal{I}(p,q) > 0$, então existe um observador para o qual p e q são simultâneos. Logo, eles não podem ser conectados por uma linha-de-mundo de luz, pois, se o fossem, a velocidade da luz medida por esse observador seria infinita (já que $\Delta t(p,q) = 0$). Por outro lado, se $\mathcal{I}(p,q) < 0$, então existe um observador para o qual ambos os eventos ocorrem no mesmo ponto do espaço. Mais uma vez, então, eles não podem ser conectados por uma linha-de-mundo de luz, pois, se o fossem, a luz teria velocidade zero para esse observador. Sendo assim, só resta o caso em que $\mathcal{I}(p,q) = 0$. Considere eventos p e q com $\mathcal{I}(p,q) = 0$ e um observador inercial $\mathcal{O}$ com linha-de-mundo passando por p, na direção dada por um 4-vetor tipo-tempo u^a *qualquer*. Agora, decomponha o 4-vetor separação $s^a = \overrightarrow{pq}$ num 4-vetor na direção de u^a, $s^a_\parallel := \mathcal{P}_{\mathbf{u}}(s^a)$, e num 4-vetor ortogonal a u^a, $s^a_\perp := s^a - s^a_\parallel \in \mathbb{V}^\perp(u^a)$. O 4-vetor $s^a_\parallel$ determina um evento $q_\parallel$ — através de $\overrightarrow{pq}_\parallel := s^a_\parallel$ — que, de acordo com o observador $\mathcal{O}$, é simultâneo a q — pois $\overrightarrow{q_\parallel q} = \overrightarrow{pq} - \overrightarrow{pq}_\parallel = s^a - s^a_\parallel \in \mathbb{V}^\perp(u^a)$ e, então, $q \in \Sigma(q_\parallel, u^a)$. Portanto, para esse observador, $\Delta t(p,q) = \Delta\tau(p,q_\parallel) = \kappa^{-1}\sqrt{-\mathcal{I}(p,q_\parallel)} = \kappa^{-1}\sqrt{-g_{ab}s^a_\parallel s^b_\parallel}$. Além disso, para esse mesmo observador, a localização espacial do evento q é a mesma do evento $q_\perp$ determinado por $\overrightarrow{pq}_\perp := s^a_\perp$. Logo, considerando que $q_\perp \in \Sigma(p,u^a)$, tem-se que $D(p,q) = D(p,q_\perp) = \sqrt{\mathcal{I}(p,q_\perp)} = \sqrt{g_{ab}s^a_\perp s^b_\perp}$. Combinando esses resultados, podemos expressar o intervalo invariante entre

p e q como

$$\begin{aligned} 0 &= \mathcal{I}(p,q) := g_{ab}s^a s^b = g_{ab}s^a_{\|}s^b_{\|} + 2g_{ab}s^a_{\|}s^b_{\perp} + g_{ab}s^a_{\perp}s^b_{\perp} \\ &= -\kappa^2 \Delta t(p,q)^2 + D(p,q)^2. \end{aligned}$$

Como a velocidade da luz, $D(p,q)/\Delta t(p,q)$, para qualquer observador é dada por c, essa condição obriga que $\kappa^2 = c^2$. Com isso, vemos que: *se $\mathcal{I}(p,q) = 0$, então existe uma linha-de-mundo de luz ligando os eventos p e q e a velocidade dessa linha-de-mundo medida por qualquer observador em p vale c.* Por isso, as direções dadas por 4-vetores ℓ^a satisfazendo $g_{ab}\ell^a\ell^b = 0$ são chamadas de *tipo-luz*. Lembrando do conceito de cones-de-luz que introduzimos no final da Seção 1.3, tem-se que

$$\mathcal{C}_p = \{q \in \mathbb{M};\ \mathcal{I}(p,q) = 0\}.$$

Note que, na argumentação acima, o fato que $\mathcal{I}(p,q) = 0$ foi utilizado apenas na manipulação final para fixar o valor de κ. Logo, a expressão

$$\mathcal{I}(p,q) = -c^2\Delta t(p,q)^2 + D(p,q)^2, \tag{1.3}$$

onde $\Delta t(p,q)$ é o intervalo de tempo e $D(p,q)$ é a distância espacial entre os eventos p e q, medidos por um observador inercial qualquer, é válida para quaisquer eventos p e q (mesmo que $\mathcal{I}(p,q) \neq 0$). É importantíssimo lembrar que, embora $\Delta t(p,q)$ e $D(p,q)$ dependam de observador, $\mathcal{I}(p,q)$ depende apenas dos eventos escolhidos. Isso nos permitirá deduzir, de maneira geométrica, os principais efeitos cinemáticos da Relatividade referentes a observadores inerciais.

• <u>Exercícios</u>

① Condidere o tensor $\delta \in \mathcal{T}(1,1)$ definido por

$$\delta(\mathbf{U},\mathbf{u}) := \mathbf{U}(\mathbf{u}),$$

para todos $\mathbf{u} \in \mathbb{V}$ e $\mathbf{U} \in \mathbb{V}^*$.

(a) Rescreva a definição acima usando a notação de índices abstratos, na qual δ é representado por δ^a_b;

(b) Calcule as componentes de δ^a_b numa base arbitrária induzida em $\mathcal{T}(1,1)$ e veja se você reconhece o resultado com outro nome;

(c) Mostre que δ^a_b pode ser interpretado como o operador *identidade* tanto em $\mathbb{V}$ quanto em $\mathbb{V}^*$;

(d) Utilizando o item anterior, mostre que $\delta^c_{b_j} T^{a_1 \ldots a_m}_{b_1 \ldots c \ldots b_n} = T^{a_1 \ldots a_m}_{b_1 \ldots b_n}$ (c colocado na j-ésima posição, com $1 \leq j \leq n$) e $\delta^{a_j}_c T^{a_1 \ldots c \ldots a_m}_{b_1 \ldots b_n} = T^{a_1 \ldots a_m}_{b_1 \ldots b_n}$ (c colocado na j-ésima posição, com $1 \leq j \leq m$);

(e) Calcule o escalar δ^a_a.

② A métrica $\mathcal{G} = g_{ab} \in \mathcal{T}(0,2)$ fornece um isomorfismo "natural" entre $\mathbb{V}$ e $\mathbb{V}^*$ — também chamado de *isomorfismo canônico* — que, na notação usando $\mathcal{G}$, é definido por[12]

$$\mathbb{V}^* \ni \mathbf{U} \leftrightarrow \mathbf{u} \in \mathbb{V} \iff \mathbf{U}(\mathbf{v}) = \mathcal{G}(\mathbf{u}, \mathbf{v}) \text{ para todo } \mathbf{v} \in \mathbb{V}.$$

[Concisamente, representa-se isso por $\mathbf{U} = \mathcal{G}(\mathbf{u}, \cdot)$.]

(a) Rescreva a definição de isomorfismo acima usando a notação de índices abstratos (inclusive para a representação concisa dada);

(b) Considere, agora, o isomorfismo *inverso* $\mathcal{F} : \mathbb{V}^* \to \mathbb{V}$, dado por

$$\mathcal{F}(\mathbf{U}) = \mathbf{u} \iff \mathbf{U} = \mathcal{G}(\mathbf{u}, \cdot).$$

Mostre que $\mathcal{F}$ pode ser visto como um tensor simétrico de posto $(2,0)$ que, em notação de índices abstratos ($\mathcal{F} = f^{ab}$), satisfaz $f^{ab} g_{bc} = \delta^a_c$.

O tensor f^{ab} é comumente denotado simplesmente por g^{ab} e a versatilidade da notação de índices abstratos nos permite definir o "abaixamento" e "levantamento" de índices tensoriais contraídos com um dos índices de g_{ab} e g^{ab}, respectivamente. Assim, ao invés de usarmos uma letra diferente para representar o covetor $g_{ab} u^b$ (digamos, U_a), simplesmente usaremos a mesma letra mas com o índice "abaixado": $u_a := g_{ab} u^b$. Note que se "levantarmos" de volta o índice de $u_a = g_{ab} u^b$, voltamos ao 4-vetor original: $g^{ab} u_b = g^{ab} g_{bc} u^c = \delta^a_c u^c = u^a$. Note, também, que a notação é consistente com o próprio levantamento de índices da métrica: $g^{ac} g^{bd} g_{cd} = g^{ac} \delta^b_c = g^{ab}$. A notação de índices abstratos permite que usemos a mesma letra para representar objetos relacionados pelos isomorfismos canônicos definidos acima, sem que haja confusão a qual espaço tensorial cada objeto pertence.

(c) Calcule as componentes de g^{ab} na base de $\mathcal{T}(2,0)$ induzida por uma base tetrada;

(d) Considere uma base tetrada $\{\mathbf{e}^a_\mu\}$ e a respectiva base dual $\{\mathbf{E}^\mu_a\}$. Definindo $(\mathbf{e}_\mu)_a := g_{ab} \, \mathbf{e}^b_\mu$ e $(\mathbf{E}^\mu)^a := g^{ab} \, \mathbf{E}^\mu_b$, relacione $(\mathbf{e}_\mu)_a$ com $\mathbf{E}^\mu_a$ e $(\mathbf{E}^\mu)^a$ com $\mathbf{e}^a_\mu$. Discuta se essas relações encontradas são válidas para uma base $\{\mathbf{x}^a_\mu\}$ (e sua dual $\{\mathbf{X}^\mu_a\}$) qualquer.

[12] A existência desse isomorfismo é garantida pelo fato da métrica ser não-degenerada. Tente entender o porquê.

③ Considere a operação de projeção ortogonal a um 4-vetor u^a (com $u_a u^a \neq 0$), definida na Sec. 1.4 através de $\mathbb{V} \ni \mathbf{v} \mapsto \mathbf{v}^{\perp} := \mathbf{v} - \mathcal{P}_{\mathbf{u}}(\mathbf{v}) \in \mathbb{V}_{\perp}(\mathbf{u})$. Mostre que essa operação é feita pelo tensor de posto (1,1) dado por

$$h^a_b = \delta^a_b - \frac{u_b u^a}{u_c u^c}.$$

Em seguida, mostre que, de fato, $h^a_b h^b_c = h^a_c$ (que o caracteriza como uma *projeção*) e $g_{ab} h^b_c = g_{cb} h^b_a$ (que, juntamente com a propriedade anterior, o caracteriza como uma projeção *ortogonal*).

④ Usando a prescrição de como um observador inercial mede distâncias e intervalos de tempo [que levou à Eq. (1.3)], pede-se:

(a) Mostre que se a velocidade de um observador inercial $\widetilde{\mathcal{O}}$ em relação a outro observador inercial $\mathcal{O}$ tem módulo V, então a velocidade de $\mathcal{O}$ em relação a $\widetilde{\mathcal{O}}$ também tem módulo V; (Sugestão: Mostre que

$$\frac{V}{c} = \sqrt{1 - \frac{1}{\left(g_{ab}\, \mathbf{e}^a_0\, \widetilde{\mathbf{e}}^b_0\right)^2}}, \tag{1.4}$$

onde $\mathbf{e}^a_0$ e $\widetilde{\mathbf{e}}^a_0$ são 4-vetores *normalizados* na direção da linha-de-mundo de $\mathcal{O}$ e $\widetilde{\mathcal{O}}$, respectivamente. Assim, argumente que a relação entre $\mathcal{O}$ e $\widetilde{\mathcal{O}}$ é simétrica.)

(b) Isole a quantidade $g_{ab}\, \mathbf{e}^a_0\, \widetilde{\mathbf{e}}^b_0$ na Eq. (1.4) e veja se o resultado lhe é familiar. (Atenção ao sinal do resultado.)

⑤ Num diagrama $c\Delta t$ por D, represente os lugares geométricos que representam valores constantes de $\mathcal{I}$ de acordo com a Eq. (1.3)

⑥ A **Fig. 1.3** mostra como construir a superfície espacial $\Sigma(p, u^a)$ para um observador inercial passando pelo evento p. Sobre a linha-de-mundo do observador, escolha dois eventos que sejam "equidistantes" (no sentido temporal) de p, um no futuro e outro no passado. A partir deles, considere linhas-de-mundo (retilíneas) de luz que se cruzam. O cruzamento dessas linhas-de-mundo de luz determina um evento q que *naturalmente* pertence a $\Sigma(p, u^a)$. Na verdade, todos os pontos de $\Sigma(p, u^a)$ podem ser gerados dessa maneira, variando-se o par de eventos (equidistantes de p) inicialmente escolhidos.

(a) Argumente o porquê é *fisicamente* natural considerar que o evento q obtido na construção acima seja considerado simultâneo a p de acordo com o observador em questão;

(b) Mostre que, de fato, $s^a := \overrightarrow{pq}$ satisfaz $g_{ab} u^a s^b = 0$ (ou seja, $\overrightarrow{pq} \in \mathbb{V}^{\perp}(u^a)$).

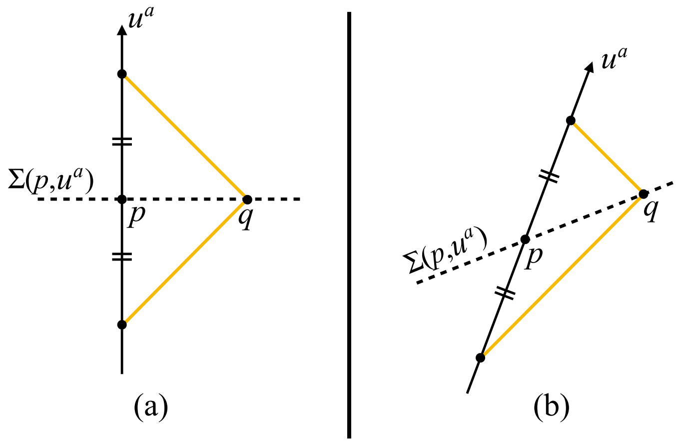

Figura 1.3: Construção da superfície espacial $\Sigma(p, u^a)$ para diferentes escolhas de observador passando pelo evento p.

⑦ Considere dois eventos p e q arbitrários (**Fig. 1.4(a)**). Agora, considere um observador inercial *qualquer* passando por um deles (digamos, p) e o cone-de-luz do outro. Seja τ_1 o intervalo de tempo, medido por esse observador, decorrido entre a intersecção de sua linha-de-mundo com o cone-de-luz passado de q e o evento p (*nessa ordem*); analogamente, τ_2 é o intervalo de tempo, medido por esse observador, decorrido entre o evento p e a intersecção de sua linha-de-mundo com o cone-de-luz futuro de q (*nessa ordem*; vide **Fig. 1.4(b)**). Com a ajuda do resultado do exercício anterior, mostre que

$$\mathcal{I}(p, q) = c^2 \tau_1 \tau_2. \tag{1.5}$$

⑧ Considere a **Fig. 1.5** abaixo, onde estão representadas as linhas-de-mundo de dois observadores inerciais passando por um mesmo evento p e parte do cone-de-luz de um evento q. A linha tracejada representa o lugar geométrico dos eventos cuja separação até o evento q possui intervalo invariante constante e positivo: $\{r \in \mathbb{M};\ \mathcal{I}(q, r) = R^2\}$. Com a ajuda do resultado do exercício anterior, mostre que

$$\tau_1 \tau_2 = \tau_3 \tau_4 = [\mathcal{I}(p, q) - R^2]/c^2. \tag{1.6}$$

(Esse resultado é o análogo espaço-temporal do chamado *teorema das cordas* em geometria euclidiana plana.)

⑨ Podemos construir um espaço-tempo que seja condizente com a física newtoniana — chamado de espaço-tempo de Galileu $\mathbb{G}$ — da seguinte

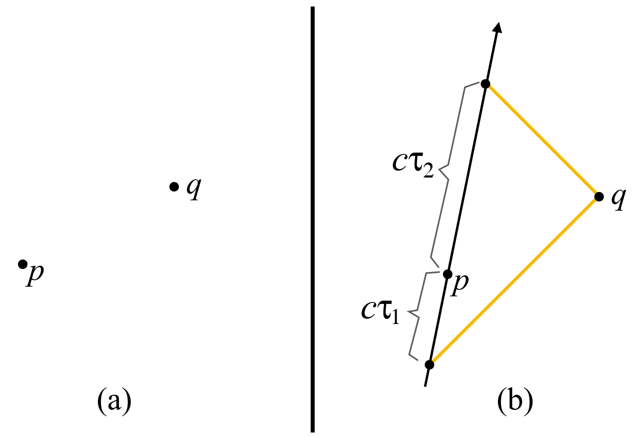

Figura 1.4: Construção para se calcular $\mathcal{I}(p,q)$ a partir de intervalos de tempo medidos por um observador inercial arbitrário.

maneira.[13] Analogamente ao caso de Minkowski, munimos o conjunto de eventos que constitui $\mathbb{G}$ de uma estrutura de espaço afim: existe um espaço vetorial $\mathbb{V}$ de dimensão 4 e um mapeamento $\psi : \mathbb{G}\times\mathbb{G} \to \mathbb{V}$ satisfazendo as propriedades listadas na Seção 1.4 — ou seja, podemos falar da separação entre eventos $p, q \in \mathbb{G}$ como sendo um 4-vetor $s^a \equiv \overrightarrow{pq} := \psi(p,q) \in \mathbb{V}$. Agora, ao invés de introduzirmos uma métrica lorenztiana (que levaria ao espaço-tempo de Minkowski), equipamos $\mathbb{G}$ com um covetor $t_a : \mathbb{V} \to \mathbb{R}$ privilegiado que tem a tarefa de, a cada par de eventos $p, q \in \mathbb{G}$, com separação $s^a := \overrightarrow{pq}$, fornecer o intervalo de tempo entre eles: $\Delta t(p,q) := t_a s^a$. Esse covetor também é o responsável por discernir direções temporais de espaciais: um 4-vetor u^a tem direção temporal se $t_a u^a \neq 0$; caso contrário ($t_a u^a = 0$), u^a tem direção espacial. Dessa forma, as direções espaciais formam um (sub)espaço vetorial de dimensão 3:

$$\mathbb{S} := \{s^a \in \mathbb{V};\ t_a s^a = 0\}\,.$$

Finalmente, $\mathbb{S}$ é equipado com um legítimo produto interno — ou seja, um tensor simétrico de posto (0,2), $\delta_{ab} : \mathbb{S}\times\mathbb{S} \to \mathbb{R}$, satisfazendo $\delta_{ab} s^a s^b \geq 0$, com a igualdade ocorrendo apenas para $s^a = 0$. Isso completa o espaço-tempo de Galileu.

(a) Mostre que o conjunto $\mathbb{S}$ definido acima é, de fato, um espaço vetorial de dimensão 3.

Dado um 4-vetor tipo-tempo u^a (ou seja, com $t_a u^a \neq 0$), o tensor de

[13] Há diferentes — mas equivalentes — maneiras de se definir o espaço-tempo de Galileu. Uma particularmente elaborada é como espaço fibrado, com espaço base $\mathbb{R}$ e fibras $\mathbb{E}^3$. Aqui, apresentaremos uma construção muito mais elementar.

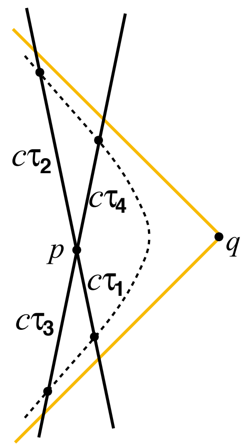

Figura 1.5: Representação de dois observadores inerciais passando pelo mesmo evento p, cone-de-luz (parcial) de um evento q e o lugar geométrico determinado por $\{r \in \mathbb{M};\ \mathcal{I}(q,r) = R^2\}$. Fixados os eventos p e q, o produto $\tau_1\tau_2$ é uma constante independente do observador escolhido. Esse resultado é o análogo do *teorema das cordas* válido em geometria euclidiana.

posto $(1,1)$ definido por

$$h^a_b := \delta^a_b - \frac{t_b u^a}{t_c u^c}$$

define uma projeção de $\mathbb{V}$ em $\mathbb{S}$. Com ele, podemos obter a distância espacial que um observador inercial — com linha-de-mundo (retilínea) na direção de u^a — atribui para um par de eventos $p, q \in \mathbb{G}$:

$$D(p,q) := \sqrt{\delta_{ab} h^a_c h^b_d s^c s^d},$$

onde $s^a := \overrightarrow{pq}$. (Ou seja, D é a norma da projeção de s^a em $\mathbb{S}$, projeção essa associada a u^a.)

(b) Mostre que, de fato, h^a_b definido acima define uma projeção e que a imagem dessa projeção é elemento de $\mathbb{S}$;

(c) Sejam $\mathcal{O}$ e $\widetilde{\mathcal{O}}$ dois observadores inerciais com linhas-de-mundo (retilíneas) nas direções dadas pelos 4-vetores u^a e $\widetilde{u}^a$, respectivamente, ambos normalizados, por coveniência, segundo $t_a u^a = 1 = t_a \widetilde{u}^a$. Mostre que o módulo da velocidade relativa entre eles é dada por

$$V = \|u^a - \widetilde{u}^a\| := \sqrt{\delta_{ab}(u^a - \widetilde{u}^a)(u^b - \widetilde{u}^b)};$$