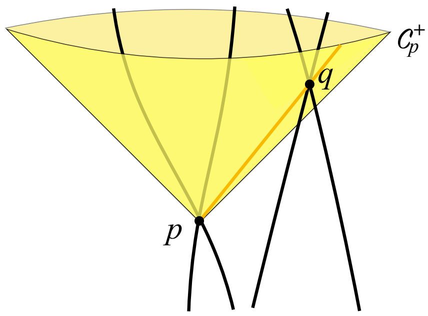

Figura 1.2: Diagrama espaço-tempo mostrando que a estrutura de cone-de-luz a partir de um evento p é independente do estado de movimento da fonte que emitiu a luz. Além disso, diferentes observadores medindo a velocidade da mesma linha-de-mundo de luz em q medem o mesmo valor c, embora os observadores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

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